УДК 621.391

С.Г. Рассомахин

Харьковский национальный университет им. В.Н. Каразина, Харьков

ТЕХНОЛОГИЯ ПСЕВДОСЛУЧАЙНОГО КОДИРОВАНИЯ В СЕТЕВЫХ КОММУНИКАЦИОННЫХ ПРОТОКОЛАХ КАНАЛЬНОГО УРОВНЯ

Проведен анализ потенциальных характеристик случайного кодирования и проанализированы причины, препятствующие его использованию. Предложен метод нахождения квазиоптимальных упаковок сфер в многомерном пространстве . Приведены сравнительные оценки эффективности квазиоптимальных и псевдослучайных кодов, генерируемых с использованием линейного конгруэнтного метода.

Ключевые слова: *случайное кодирование, Пуассоново поле, спектр взаимных расстояний, алгоритм Ллойда-Макса , псевдослучайные коды, линейный конгруэнтный генератор.*

Введение

Постановка проблемы. Теорема кодирования канала Шеннона [1] была доказана многими способами. Классические доказательства принадлежат Фейнстейну (комбинаторный вариант доказательства), Элайесу, Вольфовицу и Галлагеру (доказательства на основе случайного выбора кодов).

Некоторые из этих доказательств также обеспечивают получение экспоненциальной верхней границы для вероятности ошибки, как функции длины блока кода n. Доказательства на основе случайного выбора обычно называются неконструктивными (впрочем, как и неслучайное доказательство Фейнстейна), так как они не позволяют указать строго формализованный (в сложившемся стереотипе) алгоритм построения и декодирования помехоустойчивых кодов. Поэтому когда целью является построение конкретных кодов, теория кодирования, обычно, не использует доказательство теоремы кодирования канала и случайные способы, а применяет вместо этого комбинаторные и алгебраические методы.

Вместе с тем в теории кодирования и практике реализации систем передачи информации, начиная с 80-х годов наблюдается парадокс: несмотря на наличие грандиозной, наукоемкой теории помехоустойчивого кодирования, корректирующие коды применяются крайне редко. Во многом это связано с тем, что алгебраические коды, в принципе, являются лишь способом взаимного обмена между показателями удельной частотной и энергетической эффективности (при фиксированной информационной скорости для снижения требуемого отношения сигнал/шум расходуется почти в одинаковой пропорции дополнительная полоса частот канала). В этой связи актуальной становится задача отыскания приемлемых условий и методов случайного кодирования, обеспечивающих неограниченное приближение к пропускной способности физических каналов.

Анализ последних исследований. Задача конструктивного построения оптимального (для постоянного двоичного симметричного канала) алгебраического кода, имеющего при заданном объеме кодовой книги – $M = 2^k$ (где k – число информационных символов) н длине блока – n наибольшее значение минимального кодового расстояния – d_{min} в общем виде до сих пор не решена [3]. Одним из способов приближения к решению этой задачи (при достаточно большом n, как это ни парадоксально, является случайный выбор M разрешенных кодовых комбинаций из 2^n возможных. В этой связи работы, опровергающие устоявшееся мнение о том, что случайный выбор кодов является лишь способом доказательства теоремы Шеннона [1], но не практическим методом генерации кодов, начали появляться примерно 40 лет назад [2, 4]. Их конструктивизм заключается в снижении требований к случайности кодирования, что дает возможность использовать неслучайные способы получения кодовых книг. Такие коды правильнее всего называть псевдослучайными, так как для их получения используются детерминированные методы. В работе [3] показано, что при использовании случайного кодирования средняя взаимная информация между входом и выходом канала описывается выражением

$$I = C - V\frac{\log_2 n}{n}, \qquad (1)$$

где C – пропускная способность канала; V – скорость канальных символов. С увеличением n вычитаемое стремится к нулю, т.е. скорость передачи информации сколь угодно близка к пропускной способности. В работе [5] рассмотрены способы исключения из случайно получаемого кода "плохих слов" снижающих показатель d_{min}. Однако, позднее во многих публикациях было показано, что вероятность выбора "плохого" кода стремится к нулю с ростом n значительно быстрее, чем вероятность

ошибочного декодирования. Казалось бы, что полученные результаты полностью решают задачу безошибочной передачи информации со скоростью, близкой к пропускной способности. Однако практическое использование чисто случайного кода сталкивалось с непреодолимыми в то время трудностями [3, 5] – единственным методом кодирования и декодирования при случайном выборе предполагалось хранение в памяти всех разрешенных комбинаций и сравнение их с принятой комбинацией (полный перебор). При больших n такие операции становятся вычислительно не осуществимыми. В работе [6] получен конструктивно реализуемый метод декодирования псевдослучайных кодов, не обладающий экспоненциальной сложностью. Это является достаточным основанием для возрождения интереса и нового взгляда на принципы построения и применения случайных и псевдослучайных кодов.

Основной **целью данной работы** является формализация процедуры получения псевдослучайных кодов для частотно эффективных недвоичных каналов на основе линейных конгруэнтных последовательностей, а также сравнение их характеристик с характеристиками кодов квазиоптимальной пространственной упаковки, получаемых рандомизированным обобщенным алгоритмом Ллойда-Макса.

Основная часть

Рассмотрим традиционно наиболее общий случай задачи канального кодирования, при котором символы сообщений источника S с мощностью алфавита q_s, объединенные в блоки длиной k, отображаются в символы сообщений канала C с аналогичными параметрами $q_c \neq q_s$, $n \neq k$:

$$S(q_s, k) \Leftrightarrow C(q_c, n). \qquad (2)$$

Для обеспечения взаимно однозначного отображения необходимо, по крайней мере, чтобы $q_s^{\,k} = q_c^{\,n}$. Для придания канальному коду свойства помехоустойчивости необходимо потребовать выполнение неравенства

$$q_s^{\,k} < q_c^{\,n}. \qquad (3)$$

Очевидно, что при $q_s = q_c$ для выполнения (3) необходимо обеспечить $n > k$. Если $q_c > q_s$, то выполнение (3) может быть достигнуто и при $n < k$. Это означает повышение показателя удельной частотной эффективности канала – число передаваемых символов меньше числа символов в сообщении источника. Если $R = n/k$ – скорость кода, то коды с $R < 1$ будем называть частотно неэффективными, а при $R > 1$ – частотно эффективными. Получение кодов, удовлетворяющих (3) и способных обнаруживать и исправлять ошибки в канале при $R > 1$ является наиболее желаемой целью, преследуемой в современной теории кодирования. Здесь стоит подчеркнуть, что условие $R > 1$ выглядит, на первый взгляд, диссонансным с позиций устоявшихся воззрений избыточной помехоустойчивой передачи информации. Рассмотрение этого случая уместно только в совместном контексте канального и физического уровней коммуникационных протоколов – без декомпозиции передачи на этапы кодирования и модуляции. В применении к случайному кодированию, как будет показано ниже, соответствующее устройство можно с равным успехом назвать случайным кодером или случайным модулятором. Второе определение даже в большей степени отображает физическую сущность выполняемых при кодировании операций.

Процедура кодирования выглядит следующим образом. Каждому сообщению $S(q_s, k)$ ставится в соответствие кодовое слово $C(q_c, n)$. Если это слово выбирается, как случайная точка из некоторого объема, называемого емкостью канала, то получается случайный код. Если же при выборе используется детерминированный алгоритм генерации кодовых точек из этого же объема, достигающий некоторого (как правило, равномерного) распределения выходной величины, то код является псевдослучайным (ПСК). Сам по себе факт полной или псевдо случайности выбора не слишком сильно влияет на качество получаемых кодов. Для псевдослучайных методов гораздо более важным является достижение требуемого распределения выбираемых кодовых точек.

Рассмотрим энергетическое пространство случайного (псевдослучайного) кода. Пусть емкость канала определяется, как произведение трех измерений: энергии – E, времени – T и полосы частот – F. Для простоты, не снижающей общность, предположим, что каждый символ в канале передается в течение 1 секунды, а значение символа кодируется энергетическим информативным параметром гармонической квадратурной компоненты несущей частоты, например, амплитудой. Тогда в соответствии с теоремой об отсчетах (по обеим квадратурам) за время, равное 1 с., по каналу может быть передано два символа. Ограничим значение средней энергии затрачиваемой на передачу одного двоичного символа источника (в дальнейшем – бита) величиной E_b. Тогда для передачи одного символа кода в канале может расходоваться энергия $E_c = R \cdot E_b$. Нормирование длительности канального символа на 1 с дает возможность отождествления понятий средней мощности и средней энергии сигнала на интервале передачи.

Рассмотрим метод случайного формирования канальных символов, при котором кодовая книга реализуется в виде ансамбля сигналов объемной

укладки. Для каждого из M блоков символов (сообщений) двоичного источника формируется сообщение из n символов канала. Каждый символ является действительным числом, равномерно распределенным в диапазоне

$$\Delta = 2\sqrt{3 \cdot R \cdot E_b}\ , \qquad (4)$$

расположенном симметрично относительно нуля.

В этом слугае средняя энергия, расходуемая для передачи кодового слова, составит

$$E_c = \left(n \cdot \Delta^2\right) \Big/ 12 = n \cdot R \cdot E_b = k \cdot E_b\,. \qquad (5)$$

В соответствии с законом больших чисел точки кодовой книги n-символьного кода при достаточно большой длине блока располагаются вблизи поверхности n-мерной сферы, обладающей радиусом $r = \sqrt{E_c}$.

При построении кода по правилу (4) кодовая книга образует в n-мерном пространстве случайное (Пуассоново) поле точек так как выполняются следующие условия:

1) вероятность появления произвольного числа точек в любом объеме пространства не зависит от того, сколько точек попало в любые объемы, не пересекающиеся с данным;

2) вероятность попадания в элементарный объем двух или более точек пренебрежимо мала по сравнению с вероятностью попадания одной точки.

При действии в канале аддитивного гауссова белого шума (АГБШ) и использовании оптимального когерентного оценивания квадратурных амплитуд, кодовые точки на выходе канала распределяются внутри сферы неопределенности (СН) с радиусом $r_N = \sqrt{n \cdot N_0/2}$, где N_0 – спектральная плотность мощности АГБШ.

На основании свойств Пуассоново поля можно определить вероятность декодирования с ошибкой, которая произойдет, если внутри сферы с радиусом r_N окажется хотя бы одна кодовая точка, кроме истинной:

$$P_{ош} = 1 - \exp\left(-\lambda \cdot V\left(r_N\right)\right), \qquad (6)$$

где $V(r_N)$ – объем сферы неопределенности:

$$V(r_N) = \left(\sqrt{\pi \cdot n \cdot \frac{N_0}{2}}\right)^n \Bigg/ \Gamma\left(\frac{n}{2}+1\right); \qquad (7)$$

λ – плотность поля точек:

$$\lambda = 2^{R \cdot n} \Big/ V\left(r_\Sigma\right),\ V\left(r_\Sigma\right) = \frac{\left(\pi \cdot n \cdot \left(\frac{N_0}{2} + R \cdot E_b\right)\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)}; \qquad (8)$$

$V\left(r_\Sigma\right)$ – общий сферический объем существования точек поля на выходе канала с АГБШ.

Подстановка (7) и (8) в (6) дает

$$P_{ош} = 1 - \exp\left\{-\left[2^R \Big/ \left(\sqrt{1+2 \cdot R \cdot h}\right)\right]^n\right\}, \qquad (9)$$

где $h = E_b/N_0$ – отношение сигнал/шум.

Анализ поведения вероятности (9) при неограниченном возрастании длины блока n канального кода выявляет существование следующего предела:

$$\lim_{n \to \infty} P_{ош} = \begin{cases} 0, & \text{при } 2^R < \sqrt{1+2 \cdot R \cdot h}; \\ 1, & \text{при } 2^R > \sqrt{1+2 \cdot R \cdot h}. \end{cases} \qquad (10)$$

Данный результат соответствует границе Шеннона [1] для физической осуществимости сколь угодно надежной передачи информации. Следовательно, реализация случайного равномерного выбора канальных символов из диапазона (4) при ограничении на среднюю энергию передачи бита обеспечивает наилучший потенциально достижимый результат кодирования.

На практике для генерации случайных чисел кода, как правило, используется детерминированный алгоритм. Среди множества детерминированных методов генерации равномерно распределенных в заданном диапазоне чисел наиболее простым, а, следовательно, популярным является метод линейного конгруэнтного генератора (ЛКГ). При его реализации кодовое слово ПСК со скоростью R представляется точкой (вектором) n-мерного пространства: $X = \{x_0, x_1, \ldots, x_n\}$. Как указывалось выше, скорость R теоретически может быть любой неотрицательной величиной. В соответствии со свойствами ЛКГ, значения элементов X определяются правилами:

– $x_0 \in \left[0 \ldots \left(q_s{}^k - 1\right)\right]$ – порождающее число (порядковый номер) кодового слова ПСК;

– $x_i = \mathrm{mod}\left[a \cdot x_{i-1} + b, m\right], i \in 1, \ldots, n$ – числа слова ПСК, порождаемые x_0 по алгоритму ЛКГ;

– a, b, m – целые положительные константы, удовлетворяющие условиям: $m \geq q_s{}^k$, b и m – взаимно простые числа, величина $(a-1)$ кратна любому простому числу, которое меньше m и является его делителем;

– величина $(a-1)$ кратна 4, если m кратно 4.

Произвольное i-ое число ПСК связано с порождающим числом x_0 следующей зависимостью:

$$x_i = \mathrm{mod}\left[a^i x_0 + \frac{a^i - 1}{a-1} b, m\right], \quad i \in 1, \ldots, n\,. \qquad (11)$$

Если известно порождающее число x_0, то остальные числа однозначно определяются из (4) (процесс кодирования). При соблюдении указанных правил генерации чисел достигается взаимно одно-

значное отображение (2), поскольку период ЛКГ принимает максимальное значение, равное m. При этом следует наложить условие, чтобы

$$m \geq M = 2^k, \qquad (12)$$

в противном случае не удастся обеспечить выполнение требования взаимно однозначного соответствия (2). В случае обеспечения (12) с равенством, при кодировании комбинация из k двоичных символов источника может трактоваться как порождающее число x_0 соответствующего кодового слова в алгоритме ЛКГ. При этом, получаемый псевдослучайный код является полным в том смысле, что все возможные последовательности используются для представления кодовых слов. Полные ПСК более технологичны для использования не переборных методов декодирования [6].

Алгоритм ЛКГ не обладает криптографической стойкостью и, в отсутствии помех, кодовое слово кода однозначно идентифицируется по любому числу из блока длиной n. В данном случае это свойство весьма полезно, так как позволяет при наличии искажений в кодовом слове реализовать квазиоптимальный, легко реализуемый метод линейного целочисленного декодирования с исправлением ошибок, рассмотренный в [6].

Представляет интерес оценка характеристик ПСК, получаемого при помощи алгоритма ЛКГ. Такими характеристиками могут быть минимальное расстояние между кодовыми точками Пуассонова поля, а также спектр взаимных расстояний. Оценку целесообразно провести методом сравнения ПСК с обычным равномерным в Пуассоновом поле случайным кодом, а также с кодом более плотной укладки, получаемым на основе случайного кодирования. В рассматриваемых условиях потенциально достижимое минимальное расстояние идеального кода можно оценить на основе следующего равенства:

$$V(r_c) = V(r) \cdot 2^{-k}, \qquad (13)$$

где $V(r_c)$ – объем сферической области с радиусом r_c примыкающей, к каждой из точек n-мерного объема поля кода в условиях заданной скорости R и ограничения энергии E_b на передаваемый бит источника.

Решение (13) относительно r_c дает следующее значение минимального расстояния оптимального кода:

$$d = 2 \cdot r_c, \quad \text{где } r_c = 2^{-R} \sqrt{n \cdot R \cdot E_b}\,. \qquad (14)$$

Реальных способов нахождения оптимальных кодов на сегодняшний день не существует, поскольку задача плотнейшей упаковки сфер для пространств произвольной размерности в общем виде не решена. Приближение к оптимальному коду реализуемыми вычислительными методами может быть выполнено на основе применения следующего предлагаемого рандомизированного алгоритма Ллойда-Макса. В обычном виде данный алгоритм используется для векторного квантования в цифровом представлении аналоговых источников по результатам выполненных дискретных измерений [7]. В рассматриваемой задаче модификация алгоритма состоит в добавлении этапа предварительного заполнения пространства кода большим количеством равномерно распределенных случайных точек, для которых затем реализуются традиционные итерации алгоритма. Пошаговая реализация модифицированного алгоритма состоит из следующих действий.

Шаг 1. Объем n-мерного пространства кода $V(r)$ заполняется равномерно распределенными, случайно выбираемыми точками. Их число Q выбирается из условия: $Q >> 2^k$. Чем сильнее это неравенство, тем ближе будет получаемый код к оптимальному.

Шаг 2. В объеме кода, опять таки равномерно случайно выбирается 2^k опорных стартовых точек алгоритма.

Шаг 3. Имеющееся множество из Q случайных точек разбивается на 2^k подмножеств по принципу близости к каждой из имеющихся опорных точек.

Шаг 4. Для каждого из полученных подмножеств осуществляется пересчет координат опорной точки, в качестве которой принимается центр масс подмножества. Если при этом произошел сдвиг хотя бы одной из опорных точек относительно прежнего положения, то осуществляется возврат к шагу 3, переопределение подмножеств и повторение итераций. В противном случае, алгоритм считается завершенным, а код – полученным.

Виду того, что успех решения зависит от выбора опорных точек на первом шаге алгоритма, данный метод не может гарантировать строгой оптимальности кода, однако он обеспечивает, пожалуй, лучшее приближение к решению задачи плотнейшей упаковки сфер в пространстве любой размерности. К сожалению, вычислительная сложность алгоритма приемлема только для небольших длин блоков кода, а декодирование возможно только переборным способом.

Поэтому предложенный модифицированный алгоритм имеет скорее теоретическое, чем практическое значение.

Результаты сопоставительного сравнения показаны на рис. 1 для трех видов кодов:

a) равномерного случайного в кубическом объеме поля;

b) равномерного случайного, улучшенного рандомизированным алгоритмом Ллойда-Макса;

с) псевдослучайного (ПСК), полученного алгоритмом ЛКГ.

Дия простоты сравнительные расчеты сделаны для коротких длин блока $k = 4,5$ и 6 при скорости $R = 1$. Рассчитанные гистограммы спектров взаимных расстояний имеют математический смысл полученных статистическим путем плотностей вероятностей распределения расстояния между произвольной парой кодовых точек в Пуассоновом поле.

Относительно "гладкий" вид распределений а) и b) объясняется усреднением по большому числу случайно выбираемых кодов в первом случае, а также усреднением по нескольким наборам случайных стартовых точек алгоритма Ллойда-Макса – во втором. Усреднение для полных ПСК по понятным причинам невозможно. Штриховой линией на гистограммах показано значение минимального расстояния идеального кода d (14). Наименьшее отличное от нуля значение спектров соответствует минимальным расстояниям d_{min} кодов, а значение спектра в этой точке характеризует среднее число точек кода, располагающихся на расстоянии d_{min} вблизи любой произвольно выбранной точки.

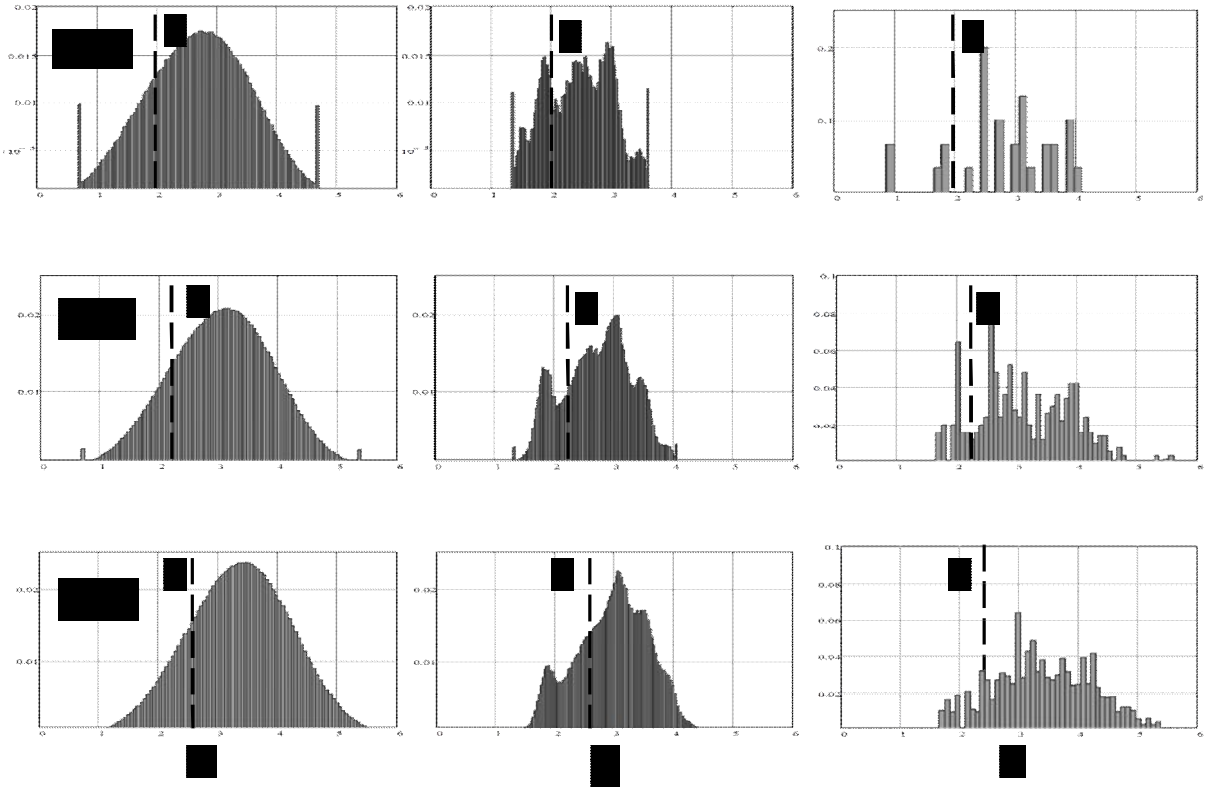

Рис. 1. Спектры взаимных расстояний

Выводы

Псевдослучайные коды, полученные по алгоритму ЛКГ, по показателю минимального расстояния выигрывают у случайных и обладают разбросом (шириной) спектра взаимных расстояний, соизмеримым с кодами Ллойда-Макса. Это делает их весьма перспективными для практического применения.

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. Shulman N. Random Coding Techniques for Nonrandom Codes / N. Shulman // IEEE Tzansactions on Information Theory. – 1999. – Vol. 45, No. 6. – P. 2101-2104.

3. Финк Л.М. Теория передачи дискретных сообщений / Л.М. Финк. – М.: Сов. радио, 1970. – 728 с.

4. Коржик В.И. Универсальное стохастическое кодирование в системах с решающей обратной связью / В.И. Коржик, С.А. Осмоловский, Л.М. Финк // Проблемы передачи информации. – 1974 .– Т. X, вып.4. – С. 25-29.

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: зб. наук. пр. – Х.: ХУПС, 2011. – Вип. 5 (95). – С. 93-98.

7. Справочник по прикладной статистике. Т. 2 / Под ред. Э. Ллойда, У. Ледермана, пер. с англ. – М.: Финансы и статистика,1990. – 526 с.

Поступила в редколлегию 2.03.2012

Рецензент: д-р техн. наук, проф. В.А. Краснобаев, Полтавский национальный технический университет им. Ю. Кондратюка, Полтава.

ТЕХНОЛОГІЯ ПСЕВДОВИПАДКОВОГО КОДУВАННЯ В МЕРЕЖЕВИХ КОМУНІКАЦІЙНИХ ПРОТОКОЛАХ КАНАЛЬНОГО РІВНЯ

С.Г. Рассомахін

Проведений аналіз потенційних характеристик випадкового кодування і проаналізовані причини, що перешкоджають його використанню. Запропоновано метод находження квазіоптимальних упаковок сфер в багатовимірному просторі. Приведені порівняльні оцінки ефективності квазіоптимальних і псевдовипадкових кодів, генерованих з використанням лінійного конгруентного методу.

Ключові слова: *випадкове кодування, Пуасоново поле, спектр взаємних відстаней, алгоритм Ллойда-Макса, псевдовипадкові коди, лінійний конгруентний генератор.*

TECHNOLOGY OF THE PSEUDOCASUAL ENCODING IS IN NETWORK OF COMMUNICATION PROTOCOLS OF LAYER OF DATA LINK

S.G. Rassomakhin

The analysis of potential descriptions of the casual encoding is conducted and reasons, impedimental to his use, are analysed. On the basis of the generalized Lloyd's algorithm of the method of being the packing of spheres is offered in multidimensional space. Comparative estimations over of efficiency of optimal and pseudocasual kodas, generated with the use of linear congruent method are brought.

Keywords: *casual encoding, Puasson the field, spectrum of mutual distances, algorithm of Lloyd's-Max, pseudorandom codes, linear congruous generator.*