УДК 621.391

ПОМЕХОУСТОЙЧИВЫЕ КОДЫ С МАКСИМАЛЬНЫМ ПРИБЛИЖЕНИЕМ К ГРАНИЦЕ ШЕННОНА

Л. А. Урывский, Е. А. Прокопенко, А. М. Пешкин

Национальный технический университет Украины “КПИ”, Киев, Украина

Разработана и исследована методика поиска наилучшего помехоустойчивого блочного кода по критерию максимального приближения к границе Шеннона. Сформулированы условия выбора блочного кода с использованием современных теорий помехоустойчивого кодирования и информации. Приведена последовательность расчета зависимостей для определения кода, исправляющего ошибки и обеспечивающего необходимую достоверность принимаемой информации при граничных значениях пропускной способности канала связи. Предложенная методика реализована в виде алгоритма выбора кода, обеспечивающего выполнение требований к достоверности декодированных символов и улучшение энергетических характеристик дискретного канала с помехами.

The best interference-immunity block code search method based on criterion of maximum approximation to the Shannon’s limit is presented and investigated. The conditions for block code choosing using recent theories of interference-immunity coding and information are formulated. The sequence of the dependences calculation is carried out for determination of code, which corrects the errors and provides the needed reliability of received information for the boundary values of channel capacity. The technique proposed is realized as the algorithm for code selection which provides the fulfillment of requirements to reliability of decoded symbols and improvement of the energy characteristics of a discrete channel with interferences.

Введение

Центральное место в теории информации занимает теорема Шеннона, согласно которой в случае, если скорость создания сообщений источником не превосходит пропускной способности канала, то существует способ кодирования и декодирования, при котором можно осуществить передачу сообщений по каналу с помехами со сколь угодно малой вероятностью ошибки. Однако теорема Шеннона не определяет параметры помехоустойчивого кода и условия его получения. Поэтому при синтезе помехоустойчивых кодов теорема Шеннона могла быть применена лишь с определенными ограничениями. Это обусловило необходимость усовершенствования теории Шеннона. Развитие теоремы Шеннона стало основой создания так называемой теории помехоустойчивого кодирования. В рамках этой теории синтезировано большое количество различных кодов, способных исправлять ошибки в передаваемых по каналам связи сообщениях. Однако в теории помехоустойчивого кодирования не решается задача определения кода, который бы не только исправлял ошибки в заданном канале связи и обеспечивал необходимую достоверность передачи сообщений, но и чтобы при этом скорость кодирования максимально приближалась к значению пропускной способности такого канала. Именно такой код претендует на статус оптимального помехоустойчивого кода в теории информации. В свою очередь, качество канала связи с позиций достоверности принимаемых символов оценивается с помощью методов теории потенциальной помехоустойчивости.

Целью работы является создание методики поиска помехоустойчивого кода с заданной длиной блока в канале связи при заданных требованиях к достоверности приема, исходя из предельных корректирующих возможностей этого кода при максимальном приближении скорости передачи к границе Шеннона.

Для достижения данной цели решается задача выбора блочного кода, оптимального по критериям теории информации, на основе объединения методов теории помехоустойчивого кодирования и теории потенциальной помехоустойчивости.

Постановка задачи

Рассмотрим модель дискретного канала с помехами, показанную на рис. 1. В данном канале от источника сообщений передаются символы со скоростью S, которые в кодере подвергаются преобразованию в блочный код с длиной блока n, после чего скорость передачи становится равной $V > S$. Каждый блок такого кода содержит $k < n$ информационных символов. Полученная информация преобразуется путем модуляции и затем передается по линии связи с помехами. Эти информационные символы на приемной стороне демодулируются с веро-

ятностью ошибки P_{er} и декодируются с достоверностью приема сообщений источника P_b. В предложенной модели требуемая достоверность приема сообщений источника обозначена P_b, скорость передачи символов источником S, бит/с, скорость передачи символов в канале связи $V > S$, бит/с. К основным параметрам используемого блочного кода отнесем также [1]: длину кода n, количество информационных символов k, скорость кодирования $r_к = k / n$, минимальное кодовое расстояние d, которое определяет исправляющую способность кода, т. е. способность исправлять $t \le (2d-1)/2$ ошибок в блоке из n символов.

Определим в рассматриваемой задаче границу Шеннона как кривую, ограничивающую максимальную скорость передачи символов источника, при которой с помощью избыточного кода еще можно исправить ошибки в канале при заданных значениях отношения сигнал / шум. Для дискретного канала с помехами, в случае передачи двоичных символов, значение пропускной способности C определяется формулой [1, 2]:

$$C = V[1 + P_{er}\log P_{er} + (1-P_{er})\log(1-P_{er})] = VE_1, \quad (1)$$

где $V \ge C$ — скорость передачи символов в канале, бит/с; P_{er} — вероятность ошибочного приема одиночного символа в канале.

Множитель в квадратных скобках соотношения (1) численно совпадает с показателем взаимной энтропии E_1 одного передаваемого символа источника, т. е. с тем количеством информации $E_1 \le 1$, которое сохранилось после передачи по каналу связи одного двоичного символа источника на его входе при воздействии помех. Из (1) следует, что скорость передачи символов источника S не может превышать пропускной способности канала: $S \le C \le V$. Эта избыточность определяет принцип помехоустойчивого кодирования, призванного обеспечить достоверность приема символов источника с вероятностью $P_b < P_{er}$. Важно заметить, что показатель E_1, называемый также удельной пропускной способностью, численно совпадает с допустимой долей символов источника в общем потоке символов в канале связи. Реальную долю символов источника в общем потоке символов в канале определяет показатель $r_к$. Отсюда следует, что

$$k / n = r_к \le E_1 = C / V = 1 + P_{er}\cdot\log P_{er} + (1-P_{er})\cdot\log(1-P_{er}) \quad (2)$$

Таким образом, в предельном случае, когда $E_1 = r_к$ в соотношении (2), пропускная способность C во столько раз меньше скорости передачи символов в канале связи V, во сколько раз число информационных символов k отличается от длины блока n избыточного кода. Стремление $r_к$ к значению E_1 при условии, что код с указанной скоростью кодирования $r_к$ обеспечивает требуемую достоверность P_b в канале с заданной энергетикой, и означает достижение кодом границы Шеннона.

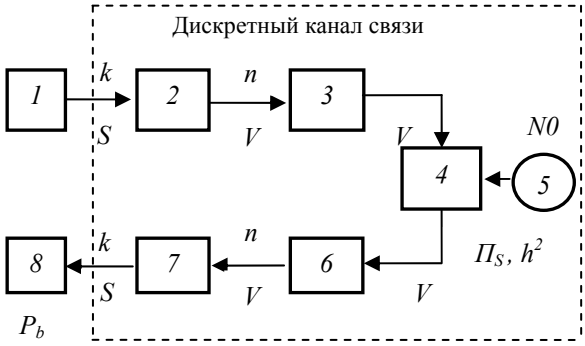

Рис. 1. Модель дискретного канала с помехами при использовании помехоустойчивого кодирования:(*1*) источник информации; (*2*) кодер (n,k); (*3*) модулятор; (*4*) линия связи; (*5*) помеха; (6) демодулятор; (*7*) декодер (n,k); (*8*) получатель информации

Взаимная энтропия E_1 зависит от значения P_{er} и, следовательно, от энергетического параметра h^2 дискретного канала, в котором действует помеха со спектральной плотностью N_0:

$$h^2 = \Pi_S / (N_0 V), \quad (3)$$

где Π_S — мощность сигнала в точке приема.

В дальнейшем будем рассматривать дискретный канал с двоичной фазовой манипуляцией ФМ-2, для которого вероятность ошибки определяется формулой:

$$P_{er}(h^2) = 0{,}5[1 - \Phi(\sqrt{2h^2})], \quad (4)$$

где $\Phi(x)$ — функция Крампа.

Из соотношений (2)—(4) следует, что с возрастанием энергетического параметра h^2 увеличивается также и величина взаимной энтропии E_1. Результаты расчета зависимости взаимной энтропии E_1 от энергетического параметра h^2 приведены на рис. 2. Полученную кривую можно рассматривать как границу Шеннона для двоичного символа, передаваемого от источника по каналу с помехами.

Как видно из рисунка, при безграничном увеличении энергетического параметра взаимная энтропия стремится к единице. В реальных каналах связи значение энергетического параметра ограничено,

что наряду с уменьшением пропускной способности снижает помехоустойчивость линии связи. В результате возникает необходимость обращаться к методам помехоустойчивого кодирования.

Для определения пропускной способности канала связи выражение (1) должно использоваться совместно с формулой (2), которая сближает характеристики канала P_{er} с параметрами кода $r_к$. Для оценки эффективности использования корректирующего кода необходимо выявить взаимосвязь между энергетическим состоянием канала связи и границей Шеннона. Это дает возможность определить условия возникновения ошибок в таком канале и отыскать параметры кода, который в состоянии их исправить.

Вероятность того, что в последовательности из n символов в канале с энергетическим параметром h^2 появится ровно m ошибок, определяется формулой [1—3]:

$$P_m(m,n,h^2) = \frac{n!}{m!(n-m)!} P_{er}^m (1-P_{er})^{n-m}, \qquad (5)$$

где P_{er} является функцией h^2.

Тогда на основе теории помехоустойчивого кодирования для вероятности P_b правильного приема одиночного символа на выходе декодера, который надежно исправляет t ошибок в блоке из n символов, можно получить следующее выражение:

$$P_b = 1 - P_t(m,n,h^2) = \\ = 1 - \sum_{m=0}^{t} \frac{n!}{m!(n-m)!} P_{er}^m (1-P_{er})^{n-m}, \qquad (6)$$

где $m \le t$, а P_{er} то же, что и в (5).

Выражение (6) связывает достоверность принимаемых символов источника с исправляющей способностью кода. В соответствии с (2) и (6) искомый код должен одновременно исправлять t ошибок и иметь скорость кодирования $r_к$, максимально приближенную к значению E_1.

Задача заключается в том, чтобы среди известных помехоустойчивых кодов с заданной длиной блока n найти код, который бы позволял передавать сообщения от источника по каналу связи с заданной достоверностью декодирования символов источника P_b и со скоростью кодирования $r_к$, максимально приближающейся к границе Шеннона.

В итоге задача нахождения оптимального помехоустойчивого кода со скоростью кодирования $r_к$ сводится к поиску минимума следующего функционала:

$$\min_{(c,\,k/n)} (E_1 - r_к) \,\Big|\, n, h^2, P_b, \qquad (7)$$

где $c = d/(2n)$ является аргументом функций, определяющих необходимые и достаточные условия существования кодов с заданными корректирующими свойствами

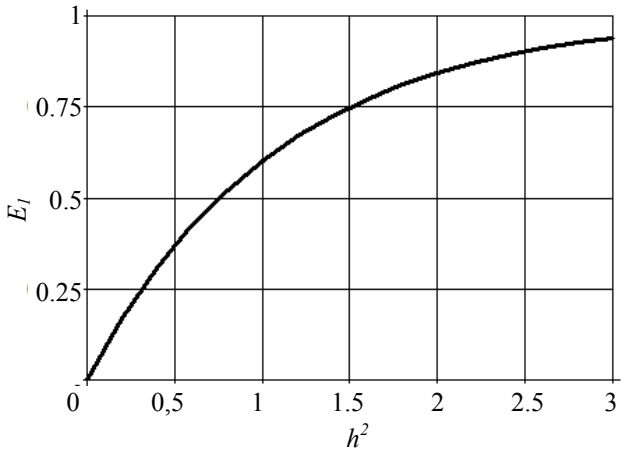

Рис. 2. Зависимость взаимной энтропии E_1 от энергетического параметра h^2

Результаты

В [2] получены результаты для случая целочисленных значений t количества подлежащих исправлению ошибок в зависимости от энергетического параметра h^2 при фиксированной длине блока n. Чтобы получить зависимость $t(h^2)$ в общем виде, применяем аппроксимацию гамма-функции:

$$\Gamma(z) = \int_0^\infty t^{z-1} \exp(-t) dt.$$

Результаты расчета зависимостей $t(h^2)$ при $P_b = 10^{-5}$ для нескольких значений n представлены на рис. 3. Приведенные зависимости отображают количество символов, которые необходимо исправить в канале с заданными энергетическими параметрами для обеспечения требуемой достоверности $P_b = 10^{-5}$ при использовании двоичной фазовой манипуляции ФМ-2. Количество возникающих ошибок как функция длины блока имеет монотонную, но не линейную зависимость. Значительное увеличение длины n блока незначительно увеличивает число вероятных ошибок t в блоке.

Преобразуем координату ошибок t на рис. 3 в координату c, связанную с характеристиками корректирующих кодов, используя соотношение $d \ge 2t + 1$. Данное преобразование позволяет перейти к осям координат зависимостей, показанных на

рис. 4. Они позволяют оценить пределы корректирующих возможностей кодов.

Приведенные на рис. 4 зависимости имеют ту же тенденцию изменения, что и кривые, показанные на рис. 3, отличие состоит лишь в количественных соотношениях. Здесь параметры t и d растут медленнее, чем значения n, поэтому линия для случая $n = 50$ на рис. 3 проходит выше, чем линия для случая $n = 100$, а на рис. 4 — ниже.

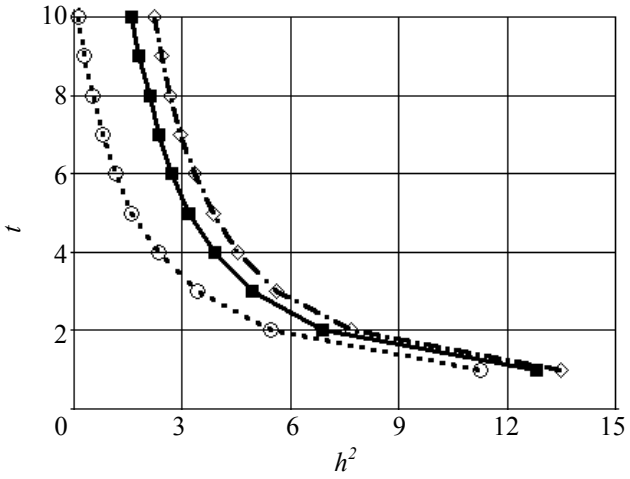

Рис. 3. Вероятное количество ошибок t в блоке заданной длины n при достоверности $P_b = 10^{-5}$: $n = 10$ (пунктирная линия); $n = 50$ (сплошная линия); $n = 100$ (штрихпунктирная линия)

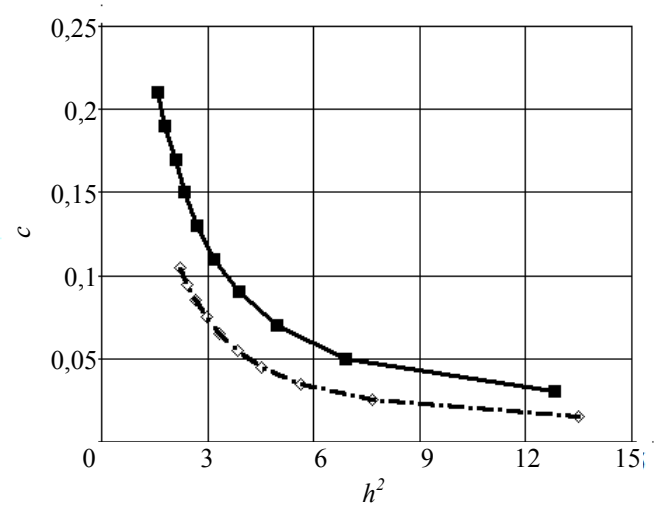

Рис. 4. Приведенное количество ошибок $c = d/(2n)$ в блоке заданной длины n при достоверности $P_b = 10^{-5}$: $n = 50$ (сплошная линия); $n = 100$ (штрихпунктирная линия)

Применим полученные результаты для определения характеристик того кода, который может обеспечить заданную достоверность принимаемой информации при скорости кода $r_к$, приближающейся к границе Шеннона E_1.

Для этого обратимся к структурам кодов, исправляющих ошибки в передаваемых по каналу связи сообщениях, на границе их корректирующих возможностей. Необходимые условия существования кодов с заданными корректирующими свойствами определяет граница Плоткина, которая устанавливает предел мощности двоичного кода, имеющего длину n и минимальное кодовое расстояние d.

Граница Плоткина применима для кодов с большими значениями $n >> 1$ и определяется следующими условиями. При длине кодового блока $n \geq 2d - 1$ число проверочных символов $r = n - k$, необходимых для того, чтобы минимальное расстояние линейного кода достигало значения d, должно быть не меньше, чем $2d - 2 - \log_2 d$. Следовательно,

$$r_к \leq 1 - (2d - 2 - \log_2 d)/n \qquad (8)$$

С учетом соотношения (8) можно построить границу Плоткина в координатах $(c, r_к)$, которая показывает, что достижимая скорость кода ограничена длиной блока n и минимальным кодовым расстоянием d.

Достаточное условие существования кода определяет граница Варшамова — Гильберта [1], в соответствии с которой существует такой (n, k) код с длиной блока n и минимальным расстоянием d, для которого справедливо следующее неравенство:

$$n - k \leq \log_2 \sum_{i=0}^{d-2} C_{n-1}^{i}. \qquad (9)$$

Выражение (9) показывает, что при фиксированных значениях n, k и t должен существовать код, который гарантированно исправляет все ошибки, и количество избыточных символов которого не превышает значения, определяемого правой частью неравенства.

Границы Плоткина и Варшамова — Гильберта отображены кривыми на рис. 5. Видно, что граница Плоткина всегда находится выше границы Варшамова — Гильберта по оси $r_к$, поскольку при фиксированном значении параметра c гарантированная скорость кодирования ниже потенциально достижимой скорости кодирования. Соответственно, при фиксированном значении скорости кодирования $r_к$ длина блока гарантированного границей Варшамова — Гильберта кода больше потенциально достижимой минимальной длины блока, определяемой границей Плоткина.

Условие (2) определяет взаимосвязь величин $r_к$ и E_1, как показывают построенные зависимости $E_1(h^2)$ на рис. 2, $c(h^2)$ на рис. 4 и $r_к(c)$ на рис. 5. Из рис. 4 следует, что заданной величине h^2 при $P_b = \text{const}$ соответствует единственное значение $c \le 0,25$. А одному значению c, как показывает рис. 5, соответствуют два значения $r_к$, определяемые координатами пересечения линии $c = \text{const}$ с границами Плоткина и Варшамова — Гильберта. Например, на рис. 4 величине $h^2 = 3,5$ соответствует одно значение $c = 0,6$, а на рис. 5 этому значению $c = 0,6$ отвечают две проекции на ось $r_к$: пересечение с границей Плоткина определяет величину $r_к = 0,77$, пересечение с границей Варшамова — Гильберта определяет значение $r_к = 0,55$.

Изменяя непрерывно значение параметра $h^2 \ge 0$, можно получить две зависимости в координатах $(r_к, h^2)$, соответствующие границам Плоткина и Варшамова — Гильберта, как показано на рис. 6*г* (кривые *1* и *2*). Видно, что обе зависимости проходят ниже границы Шеннона, показанной на рис. 6*г* кривой *3*. Следовательно, при известной достоверности символов в канале P_{er} и заданных требованиях к достоверности символов на выходе декодера P_b нельзя приблизиться к границе Шеннона $E_1(h^2)$ ближе, чем это определяет линия $r_к(h^2)$, соответствующая границе Плоткина.

Характеристики реальных кодов могут оказаться не хуже показателей $r_к(h^2)$, соответствующих границе Варшамова — Гильберта. Таким образом, выбор кода должен осуществляться на плоскости $(c, r_к)$ по признаку пребывания в области, ограниченной границами Плоткина, Варшамова — Гильберта и линией $c = \text{const}$ (на рис. 5 эта область показана заштрихованным треугольником).

Изложенные условия при известных параметрах канала связи позволяют произвести выбор кода, скорость которого была бы максимально приближенной к границе Шеннона. Эти условия определяют последовательность операций, необходимых для отыскания кода, который бы не допускал превышения пропускной способности канала связи и обеспечивал заданную достоверность приема декодированных символов.

Рассмотрим более подробно предлагаемую методику нахождения оптимального кода в соответствии с полученными соотношениями и графической иллюстрацией, показанной на рис. 6. В рамках методики считаем заданными: длина блока кода n; требуемая достоверность приема символов источника P_b; энергетический параметр канала связи h^2; способ формирования и обработки символов.

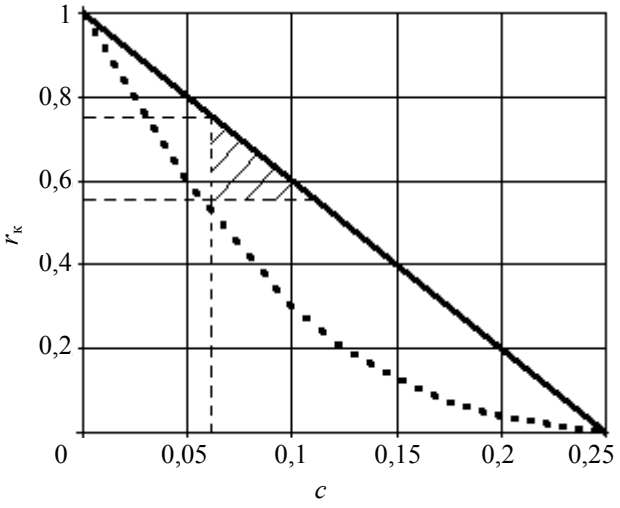

Рис. 5. Граница помехоустойчивого кодирования Плоткина (сплошная линия) и Варшамова — Гильберта (штриховая линия)

По этим исходным данным строим зависимости: вероятности ошибки в канале связи P_{er} от параметра h^2 (рис. 6*а*); количества ошибок t в блоке длиной n от параметра h^2 (рис. 6*б*); приведенного количества ошибок c от параметра h^2 (рис. 6*в*); взаимной энтропии E_1 от параметра h^2 (кривая *3* на рис. 6*г*). Рассчитываем и наносим: границы Плоткина (кривая *1* на рис. 6*д*) и Варшамова — Гильберта (кривая *2* на рис. 6*д*); проекции границ Плоткина (кривая *1* на рис. 6*г*) и Варшамова — Гильберта (кривая *2* на рис. 6*г*); параметры известных кодов с длиной блока n в координатах $(c, r_к)$ (например, точка *4* на рис. 6*д*).

В соответствии с предложенной методикой выбор оптимального кода с длиной блока n проводим в следующей последовательности. По заданному значению h^2 с помощью рис. 6*а* — рис. 6*в* определяем соответствующее значение c. По проекции найденного значения c на плоскости $(c, r_к)$ определяем граничные значения $r_к$ для корректирующего кода, как показано на рис. 6*д*. Строим проекции границ Плоткина и Варшамова — Гильберта на плоскости $(r_к, h^2)$, как показано на рис. 6*г*. Оцениваем предельные возможности кода по достижению границы Шеннона:

$$\Delta_{\max} = r_к / E_1 \le 1,$$

где $r_к$ есть проекция c от границы Плоткина.

Далее формируем треугольник параметров на плоскости $(c, r_к)$ с использованием проекции найденного значения c, как показано на рис. 6*д* в виде заштрихованной области. Выбираем внутри данного треугольника код, соответствующий необходимой

достоверности передачи и имеющий кодовую скорость $r_к$, максимально приближенную к границе E_1, например, обозначенный точкой *3* на рис. 6*д*, и определяем параметр k для блока длиной n согласно формуле $k = nr_к$. Оцениваем степень приближения выбранного кода (n,k) к границе Шеннона:

$$\Delta_к = r_к / E_1 \le \Delta_{max}.$$

Параметры кодов, которые попадают в заштрихованную область или же находятся вблизи нее, представлены в табл. 1 [3].

Задаваясь длиной блока $n = 127$, допустимой вероятностью ошибки $P_b = 10^{-6}$ и значением энергетического параметра $h^2 = 3,5$, требуемый код выбираем с помощью рис. 6 и табл. 1. Из рис. 6 находим значение $t = 11$ и соответствующий ему диапазон кодовых скоростей $r_к = 0,39$—$0,7$. Как следует из табл. 1, найденный код (127,57) с параметрами $t = 11$ и $r_к = 0,45$ обеспечивает $P_b = 1,55 \cdot 10^{-6}$. При этом скорость кода $r_к$ составляет около 50 % от значения, определяемого границей Шеннона $E_1 = 0,89$.

Таблица 1. Коды с длиной блока $n = 127$

n	k	T	$r_к$	P_b
127	71	9	0,559	0,014
127	64	10	0,504	$3,67 \cdot 10^{-3}$
127	57	11	0,449	$1,55 \cdot 10^{-6}$
127	50	13	0,394	$2,4 \cdot 10^{-8}$
127	43	14	0,339	$2,7 \cdot 10^{-9}$

Следующий код $(127,50)$, соответствующий достоверности передачи $P_b = 2,4 \cdot 10^{-8}$, находится значительно правее границы, задающей необходимую исправляющую способность кода, и имеет еще меньшую кодовую скорость $r_к = 0,39$. Удаление от границы Шеннона стало платой за увеличение достоверности. Приближение к границе Шеннона может быть достигнуто существенным увеличением длины кода $n > 127$.

Описанный алгоритм позволяет найти блочный код, обеспечивающий передачу информации с любыми требованиями к достоверности, приближая скорость кодирования $r_к$ к границе Шеннона.

Заключение

В статье с использованием современных достижений теории помехоустойчивого кодирования и теории информации сформулированы условия выбора наилучшего блочного кода по критерию максимального приближения его скорости кодирования к границе Шеннона.

Возможность достижения высокой достоверности передачи символов в канале при использовании избыточных кодов доказана К. Шенноном в рамках теории информации. Существующая теория помехоустойчивого кодирования позволяет оценить пределы корректирующих возможностей известных кодов без оценки пропускной способности реальных каналов связи. В статье предложен механизм объединения элементов указанных теорий с целью решения задачи поиска оптимального помехоустойчивого кода.

Приведена последовательность расчета для определения кода, который бы исправлял ошибки в заданном канале связи и обеспечивал необходимую достоверность принимаемой информации на границе его пропускной способности.

Исследованы корректирующие свойства помехоустойчивых кодов на границах их предельных возможностей, определяемых границами Плоткина и Варшамова — Гильберта. Расширение диапазона корректирующих свойств помехоустойчивых кодов до границы Шеннона подтверждает справедливость утверждения теоремы Шеннона о существовании кода, который можно использовать для передачи информации по каналу связи с как угодно малой ошибкой. Однако, как показывают исследования, реальные коды не позволяют достичь границы Шеннона в силу ограниченности длины кодового блока.

Предложен универсальный алгоритм поиска оптимального помехоустойчивого кода в широком диапазоне изменений энергетических параметров канала связи и требований к достоверности. Определено, что увеличение требований к достоверности приводит к отдалению кода от границы Шеннона. Приближение кодового пространства к границе Шеннона достигается увеличением длины блока.

Литература

1. Теория электрической связи: Учебник для вузов / А. Г. Зюко, Д. Д. Кловский, В. И. Коржик, М. В. Назаров // Под ред. Д. Д. Кловского. — М.: Радио и связь, 1999. — 432 с.

2. Уривський Л. О. Прокопенко К. А. Умови вибору завадостійкого блокового коду в каналі з заданими показниками достовірності // Вісник інженерної академії України. — 2011. — № 1. — С. 151—154.

3. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки / Перевод с анг. Под ред. Р. Л. Добрушина, С. И. Самойленко. — М.: Мир, 1976. — 575 с.

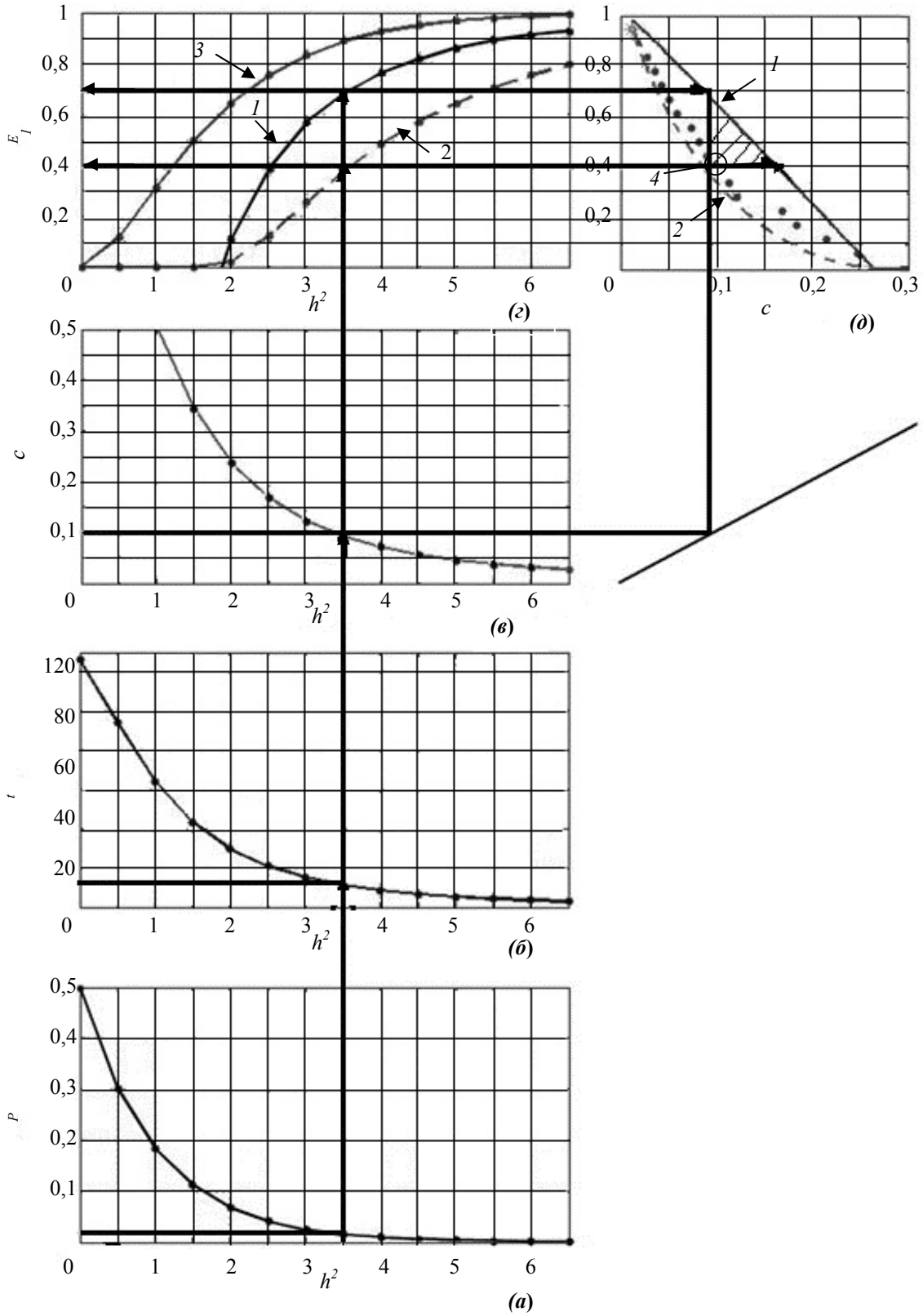

Рис. 6. Графическая иллюстрация последовательности нахождения параметров кода и границ для длины кодового блока $n = 127$ при достоверности передачи $P_b \le 10^{-6}$

Поступила после переработки 16.11.2010