УДК 621.391.

Л.О. Урившський, К.А. Прокопенко
ІТС НТУУ “КПІ”

ВИЗНАЧЕННЯ ГРАНИЦЬ КОРИГУВАЛЬНИХ ВЛАСТИВОСТЕЙ БЛОКОВИХ КОДІВ

Обґрунтовано методику побудови границі коригувальних властивостей блокових кодів в просторі тверджень Шеннона щодо безпомилкового передавання дискретної інформації по каналу із шумами. У результаті Гамма-аппроксимації отримано неперервні залежності між кількістю виникаючих помилок і енергетичними параметрами каналу. Завдяки поєднанню співвідношень теорії інформації та теорії завадостійкості встановлені границі коригувальних властивостей завадостійких кодів кінцевої довжини на підставі їх граничних можливостей, що визначаються границею Плоткіна.

Ключові слова: границя Шеннона, границя Плоткіна, швидкість кода, швидкість передачі інформації в каналі.

This article contains techniques for determination of block codes correcting properties limits in space of Shannon axioms for error-free transmission of discrete information in the noisy channel. As a result of gamma-approximation indiscrete dependence of errors amount from channel power parameters is got. Due to the combination of correlations from information theory and the theory of noise immunity, the block codes correcting properties limits with given length are defined by Plotkin bound.

Key words: Shannon’s limit, Plotkin bound, code rate, channel rate.

Вступ

Теорія інформації, сформульована К. Шенноном через постулати і засадничі теореми більше 60 років тому, відкрила епоху яскравих досягнень в галузі підвищення достовірності зв'язку на основі завадостійкого кодування.

Відповідно до теореми Шеннона у разі, якщо швидкість створення повідомлень джерелом не перевищує деякої величини, що називається пропускною спроможністю каналу, то, вибираючи спосіб кодування і декодування, можемо здійснити передачу цих повідомлень по каналу з шумом із скільки завгодно малою вірогідністю помилки [1]. Отже, уперше було визначено, що в каналі з шумом можна добитися скільки завгодно малої вірогідності помилки на приймальній стороні, якщо обмежувати швидкість передачі символів джерела.

Побудова границь завадостійкого коду на основі аксіоматики Шеннона

Конструктивним результатом звернення до теорії Шеннона потрібно розглядати поняття границі Шеннона – кривої, що визначає максимальну швидкість передачі символів джерела, за якої за допомогою надмірного коду ще можливо виправити помилки в каналі при заданому відношенні сигнал/шум.

Для дискретного каналу з завадами, у разі передачі двійкових символів, значення пропускної спроможності C, біт/с визначається за формулою

$$C = V \cdot \{1 + P_{\mathrm{er}} \cdot \log P_{\mathrm{er}} + (1 - P_{\mathrm{er}}) \cdot \log(1 - P_{er})\}, \tag{1}$$

де V, біт/с – швидкість передачі символів в каналі; P_{er} – вірогідність помилкового приймання у поодинокого символу в каналі.

Множник у фігурних дужках співвідношення (1) числово збігається з показником взаємної ентропії E_1 одного переданого символу джерела, тобто з тією кількістю інформації $E_1 \leq 1$ біт, яка збереглася після передавання по каналу одного двійкового символу в результаті дії завад.

Отже, швидкість передачі символів джерела v, біт/с не може перевищувати значення пропускної спроможності : $v \leq C \leq V$. Надмірне кодування має забезпечити достовірність приймання символів джерела з вірогідністю $P_b << P_{er}$.

У межах теореми Шеннона створена величезна кількість різноманітних кодів, здатних виправляти помилки в каналах з різною якістю.

Проте, параметри коду, за допомогою якого можна отримати результати, про які йдеться в теоремі, не визначені. Тому склалася думка про те, що наведена теорема Шеннона, загалом, неконструктивна для практики синтезу завадостійких кодів.

Метою статті є встановлення границь коригувальних властивостей завадостійких кодів із заданою довжиною блока на підставі їх граничних можливостей, які визначаються границею Плоткіна, і оцінювання міри їх наближення до границі Шеннона.

Реалізація мети повинна надати конструктивні засади аксіоматиці Шеннона.

Розглянемо модель дискретного каналу, в якому передаються закодовані блоковим кодом довжини n і модульовані символи джерела, які на приймальній стороні піддаються демодуляції і декодуванню.

Енергетичний стан дискретного каналу, в якому потужність сигналу в точці приймання дорівнює P_s і діє перешкода із спектральною щільністю N_0 визначається показником :

$$h^2 = \frac{P_S}{N_0 \cdot V} \quad . \qquad (2)$$

Для прикладу в цих залежностях за основу взятий дискретний канал з фазовою маніпуляцією (ФМ-2), де вірогідність помилки визначається за формулою

$$P_{er}(h^2) = 0,5[1-\Phi(\sqrt{2h^2})] \ , \qquad (3)$$

де Ф (x) – функція Крампа.

Залежність взаємної ентропії E_1 від енергетичного параметра h^2 відповідно до (1)–(3) наведена на рис. 1, де збільшення h^2 пов'язане тільки зі зменшенням швидкості V передавання символів.

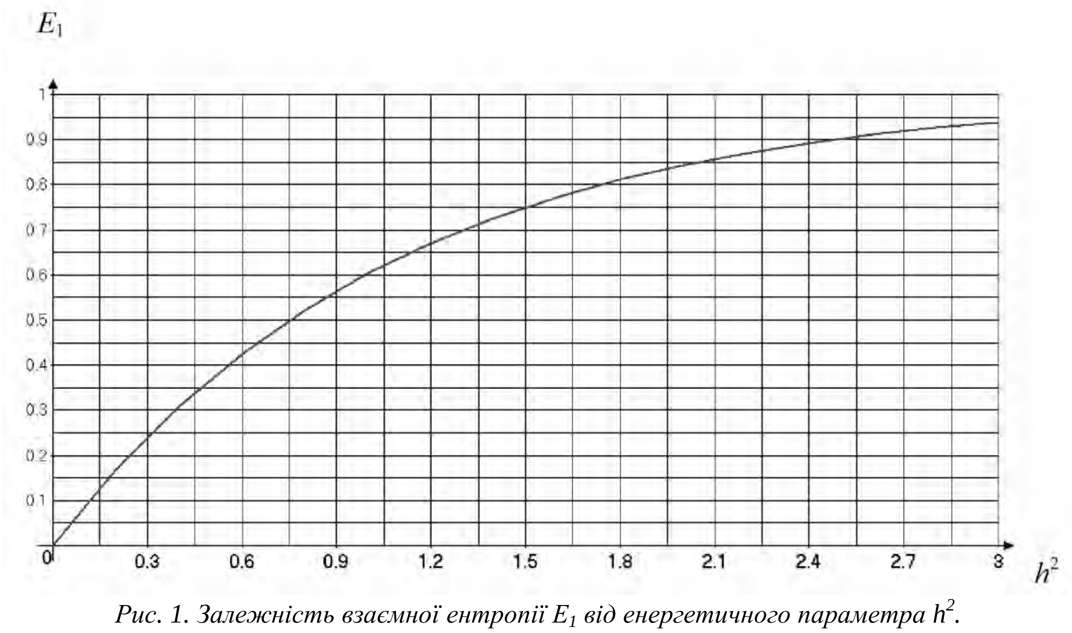

Рис. 1. Залежність взаємної ентропії E_1 від енергетичного параметра h^2.

Визначивши за допомогою співвідношень (1)–(3) взаємозв'язок між енергетичним станом каналу і границею Шеннона, звернемося до умов виникнення і виправлення помилок у такому каналі.

Знаючи характеристики завадостійкої каналу, можна визначити вірогідність виникнення заданої кількості помилок m у блоці довжини n за формулою

$$P_m(m,n,h^2) = \frac{n!}{m!(n-m)!} P_{er}(h^2)^m (1-P_{er}(h^2))^{n-m} \qquad (4)$$

Результати розрахунку вірогідності появи різної кількості помилок для блокового коду при зміні співвідношення параметра h^2 на підставі виразів (2)–(4) наведені на рис.2.

З рис.2 видно, що при фіксованій довжині блока (n = 100) можлива поява різної кількості помилок з різною вірогідністю $P_m(m)$, залежною від параметра h^2. Кожному значенню m помилок відповідає деяке значення h^2, за якого досягається максимальне значення $P(m, h^2) = P_{max}(m)$. У разі зменшення значень h^2 максимуми зміщуються в область великих значень m, що відповідає фізичній природі передавання сигналів. Відповідно, кількість помилок зменшується за збільшення співвідношення сигнал/шум.

Сформулюємо ключове твердження для формулювання вимог до коригувальних властивостей надмірного коду: для досягнення необхідної достовірності приймання декодованих символів джерела P_b необхідно, щоб надмірний код дозволяв виправити t помилок, що виникають у блоці довжини n з вірогідністю $P_n\ (t, h^2) = P_b$.

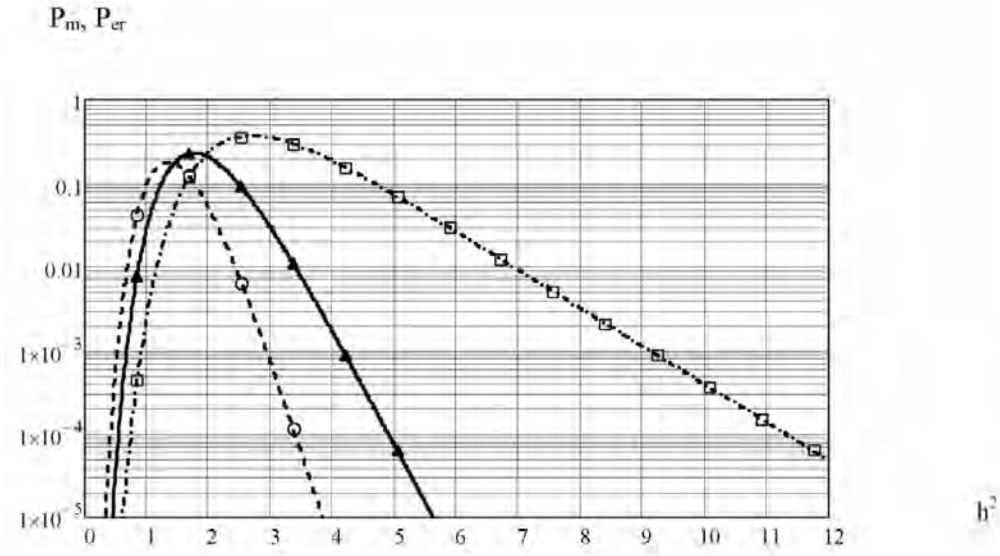

Рис. 2. Залежності вірогідності появи різної кількості помилок від співвідношення сигнал/шум у блоці з n=100 символів m = 1, 3, 5 помилок від h^2.
(– n=100, m=1, – n=100, m=3, – n=100, m=5)

На рис. 3 наведені залежності кількості помилок $t = \arg\{P_{max}(n, h^2)\}$ для заданого значення $P_b = 10^{-5}$ і $n = 100$ як функцій параметра h^2.

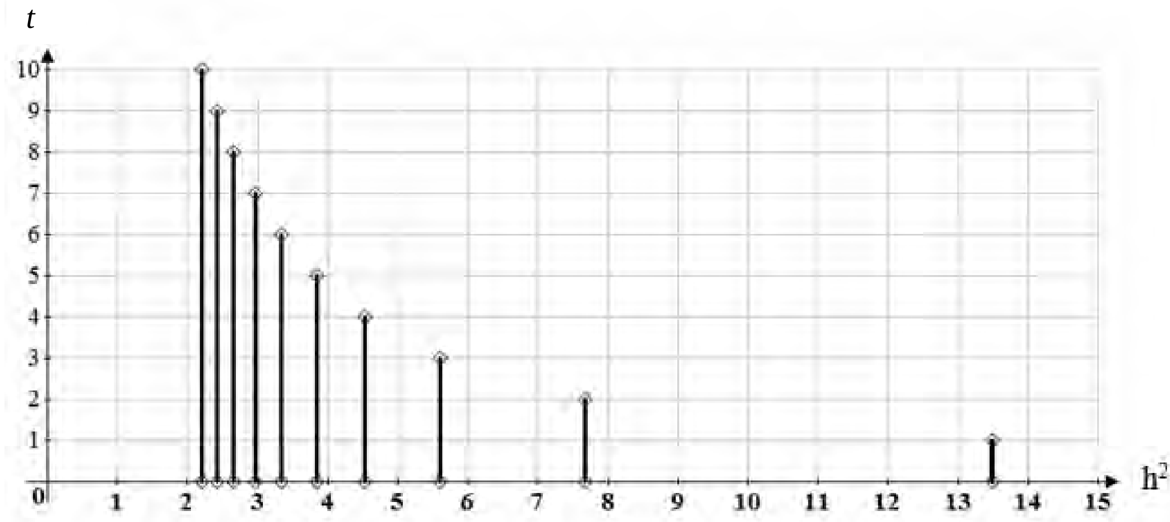

Рис. 3. Вірогідна кількість помилок у блоці заданої довжини n за заданої достовірності P_b=10^{-5} для випадку n=100

На рис. 3 відображені цілочислові значення t кількості помилок, які необхідно виправляти за різних значень h^2 при фіксованій довжині блока. Щоб отримати безперервну залежність $t\ (h^2)$ для не цілих значень t, звернемося до категорії гамма-функція $\Gamma(z)$, яка розширює поняття факторіалу у виразі (4) на дробові значення аргументу z [2]. Гамма-функція не виражається через елементарні функції, але може бути зображена у вигляді

$$\Gamma(z) = \int_0^\infty t^{z-1}\exp(-t)dt\ . \tag{5}$$

Результати розрахунку відповідних залежностей $t\ (h^2)$ при P_b=10^{-5} для декількох значень n наведені на рис.4.

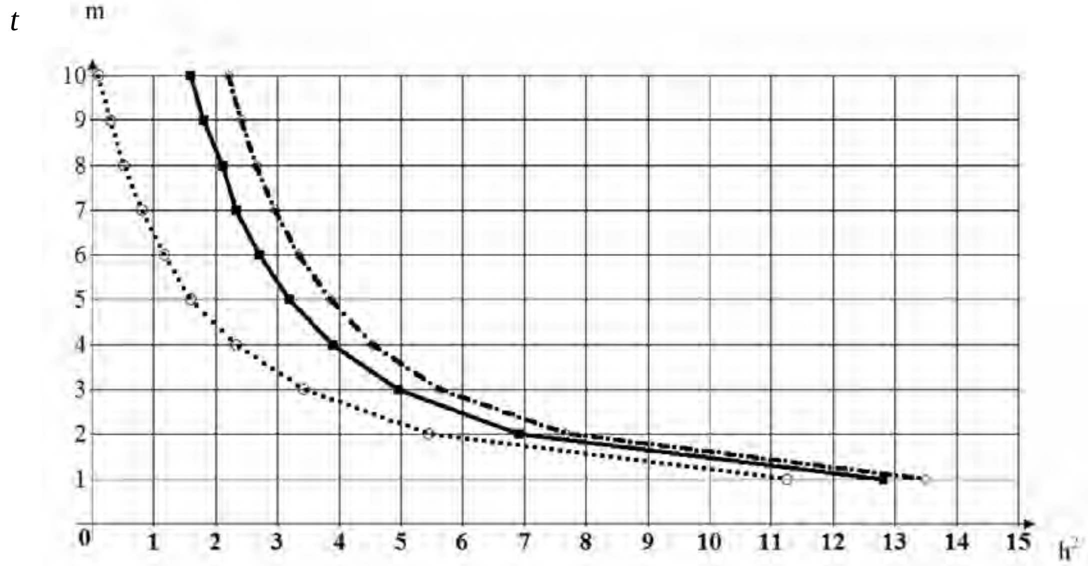

Рис. 4. Вірогідна кількість помилок t у блоці заданої довжини n при заданій достовірності P_b=10^{-5}.
– n=100, – n=50, – n=10

Визначивши умови виникнення помилок у блоці з n символів, звернемося до границі коригувальних можливостей n-значного блокового коду.

Основними параметрами блокових кодів є: довжина коду n, кількість інформаційних символів k, швидкість кодування $R = k / n$, мінімальна кодова відстань d і, як наслідок, коригувальна здатність коду – здатність виправити $t \le (2\ d - 1)/2$ помилок у блоці з n символів.

Необхідною умовою для існування кодів із заданими коригуючими властивостями виступає границя Плоткіна, яка визначає межу потужності двійкового коду довжини n і мінімальної кодової відстані d.

Границю Плоткіна можна застосовувати для кодів з великими значеннями n >> 1 і задають такими умовами: при довжині кодового блока $n \ge 2d$-1 кількість контрольних символів для перевірки $r = n - k$, які необхідні для того, щоб мінімальна відстань лінійного коду досягала значення d, має бути не меншою $2d\text{-}2\text{-}\log_2\ d$ [3]:

$$k \le n - (2d - 2 - \log_2 d). \tag{5}$$

Тоді, як наслідок, необхідна швидкість кодування становить

$$R \le 1 - \frac{2d - 2 - \log_2 d}{n}\ . \tag{6}$$

Простір границь кодів, що реалізовуються, відображається в координатах $\{\frac{k}{n}; \frac{d}{2n}\}$, де k – кількість інформаційних символів у блоці довжини n; $\frac{k}{n} = R$ – швидкість коду; d – відстань Хеммінга для обраного коду, яка пов'язана з коригувальними властивостями коду співвідношенням: $t \le (d - 1)/2$ (рис. 5).

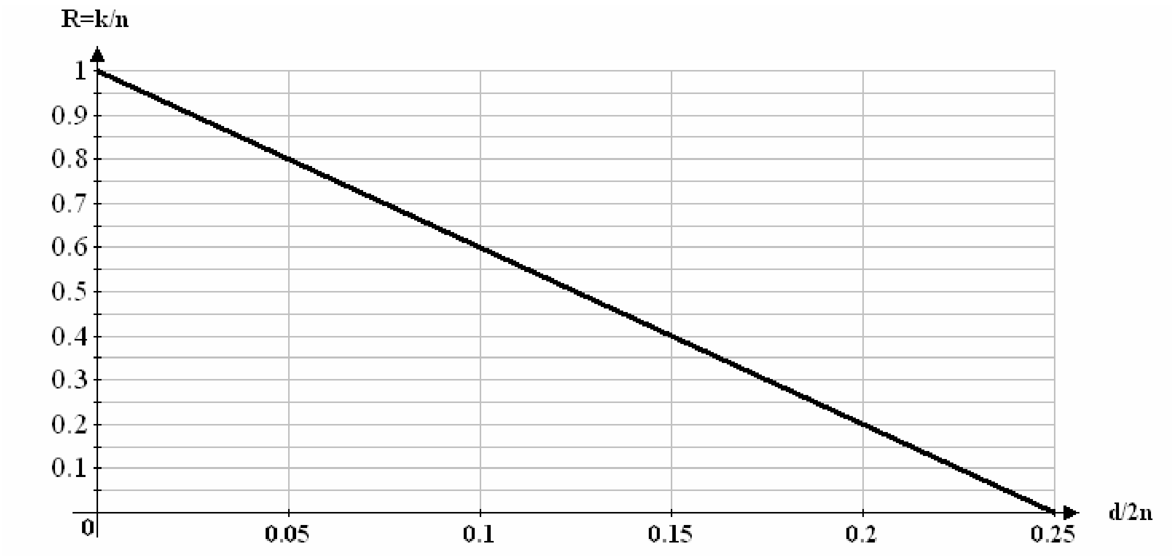

Рис. 5. Границя завадостійкого кодування Плоткіна

Для узгодження координати помилок t на рис. 4 з координатою коригуючих властивостей кодів d / $2n$ на рис.5 трансформуємо графік рис. 4 до координат $d/2n$ (h^2), використовуючи співвідношення $d = 2t + 1$. Отримані залежності наведені на рис.6.

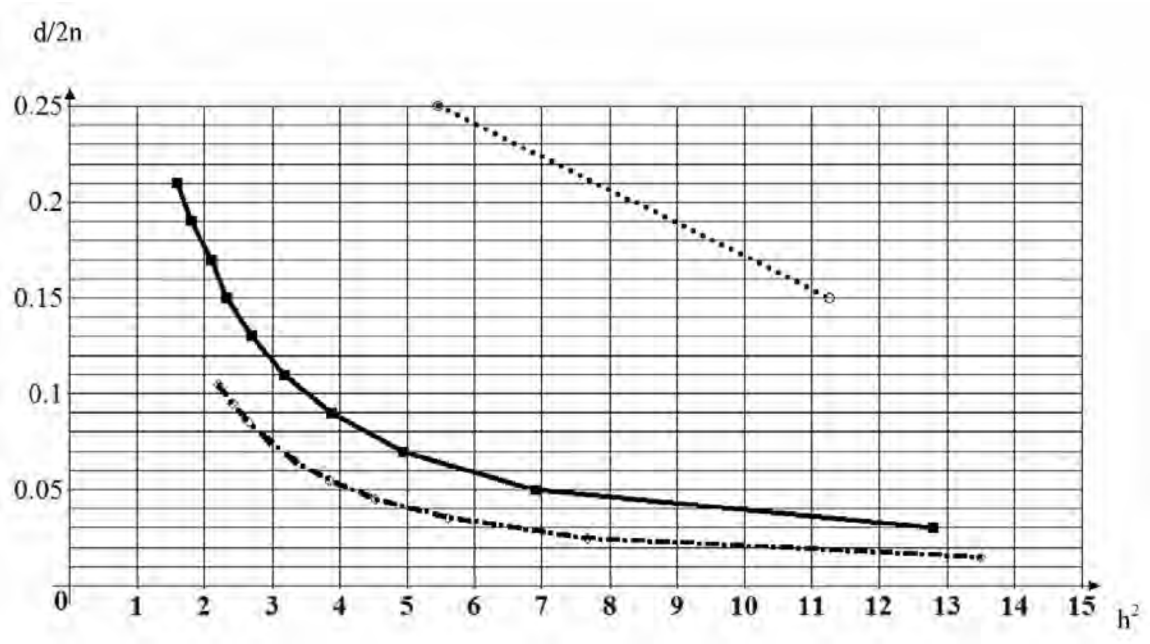

Рис. 6. Наведена кількість помилок у блоці заданої довжини n при заданій достовірності $P_b=10^{-5}$ в координатах d/2n. – n =100, – n = 50, – n = 10

Наступний крок до побудови методики побудови границь коригувальних властивостей блокових кодів полягає в твердженні, що швидкість надмірного коду $R = \frac{k}{n}$, яка забезпечує передачу кожних k символів джерела в послідовності з n закодованих символів, числово збігається з показником $E_1 = \frac{C}{V}$ (рис. 1)*:* $R = E_1$.

Інакше кажучи, в ідеальному випадку пропускна спроможність C в стільки разів менша від швидкості передавання символів у каналі V, у скільки разів кількість інформаційних символів відрізняється від довжини блока n надмірного коду. Ідеальним потрібно вважати випадок, що відповідає доведеному Шеноном твердженню про те, що, за швидкості передавання повідомлень джерела меншій за пропускну спроможність каналу зв'язку ($v < C$), існують коди і методи декодування такі, що середня вірогідність помилки декодування прагне до нуля ($P_b \rightarrow 0$), коли довжина блока прагне до нескінченності ($n \rightarrow \infty$).

Оскільки реальні коди мають кінцеву довжину коду, то останнім етапом методики є встановлення границі коригувальних властивостей цих кодів у просторі параметрів $\{n;h^2;d\}$. Для цього об'єднаємо залежності, наведені вище на рис.1, 5 і 6, в тривимірному просторовому відображенні (рис.7).

Одна площина простору є залежністю швидкості кодування від енергетичного параметра h^2 (рис. 1), друга є границею потужності завадостійкого кодування Плоткіна (рис. 5), третя – залежність найвірогіднішої кількості помилок у блоці заданої довжини n при заданих параметрах каналу (рис. 6).

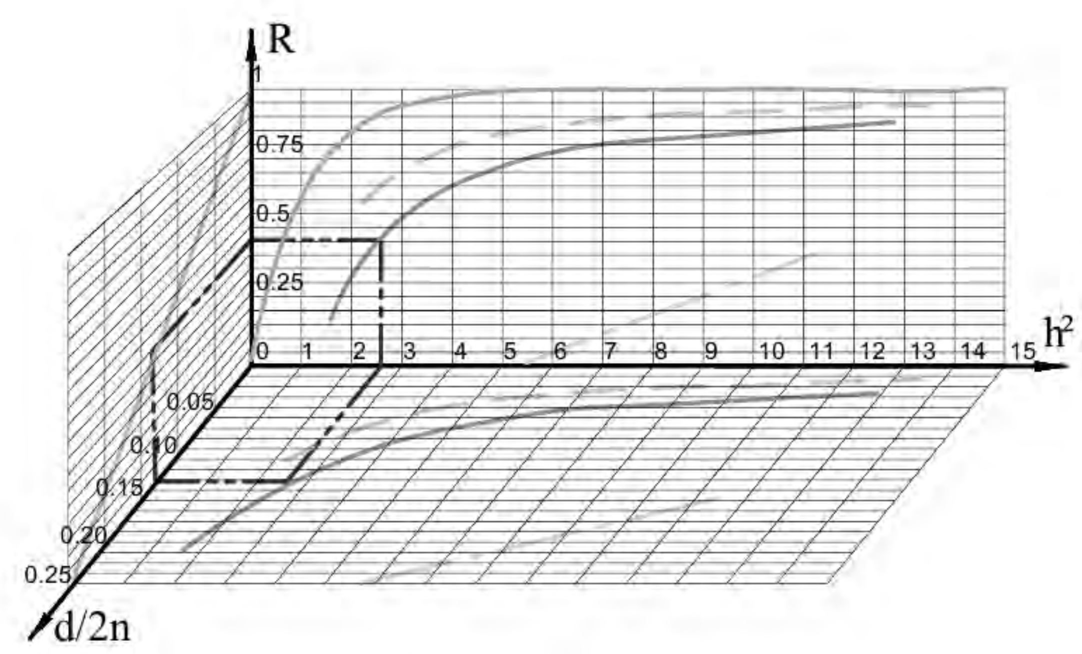

Рис. 7. Знаходження границь коригувальних властивостей блочних кодів

Пропонується така послідовність дій в тривимірному просторі для формування границі коригувальних властивостей блокових кодів.

1. Для коду з вибраним значенням довжини n і необхідною достовірністю приймання символів P_b, враховуючи спосіб формування і оброблення символів у каналі із заданим енергетичним параметром h^2, визначається значення t через вірогідність $P_n(t, h^2) = P_b$ із співвідношень (4) і (5). Після цього параметр t перераховується в параметр $d/2n$.

Для змінних значень h^2 і фіксованих n криві $d/2n\{h^2; n\}$ наведені на горизонтальній площині рис. 7. Жовта штрих-пунктирна лінія відповідає n=10, коричнева суцільна для n=50, зелена переривчаста для n=100.

2. Використовуючи проекцію значення $d/2n$ на лінію границі Плоткіна (ліва вертикальна площина на рис.7), визначаємо граничне значення параметра R – швидкості кодування.

3. На перетині проекцій осей R і h^2 фронтальної площини на рис. 7 визначаємо те значення ентропії $E_1\{h^2; n\} = R$, яке відповідає гранично досяжному співвідношенню між пропускною спроможністю каналу C (рожева суцільна лінія) і швидкістю передавання джерела v.

Для змінних значень h^2 і фіксованих n (10,50 і 100) криві $E_1\{h^2; n\} = R$ наведені на фронтальній площині рис. 7.

Траєкторія для отримання однієї точки, бажаної залежності, показана червоною лінією.

Головний висновок з отриманих залежностей полягає в тому, що реальні коди довжини $n<\infty$ не дозволяють досягти границі Шеннона. До того ж для блокових кодів довжини n, на які накладається обмеження межі Плоткіна, завжди можна визначити таке значення h^2 в заданому каналі, за якого вони забезпечують задану достовірність через значення P_b. Чим більше n, тим менше h^2 при фіксованому P_b .

Крім того, отримані результати демонструють, що у разі збільшення довжини коду n границя коригувальних властивостей блокових кодів наближається до границі Шеннона.

Висновки

Встановлені границі коригувальних властивостей завадостійких кодів на підставі їх граничних можливостей, що визначаються границею Плоткіна. Отримані границі надають конструктивності теоремі Шеннона, в якій стверджувалося існування коду, який можливо використати для передавання інформації по каналу з помилкою, що прагне до нуля. Проте реальні коди довжини $n<\infty$ не дозволяють досягти границі Шеннона.

Запропонована методика дає змогу виявити нові границі співвідношення між швидкістю джерела і пропускною спроможністю каналу при заданих параметрах каналу і відомій довжині блокового коду.

Для блокових кодів довжини n, з урахуванням границі Плоткіна, завжди можна визначити те значення h^2 в заданому каналі, за якого коди забезпечують задану через значення P_b достовірність. Чим більше n, тим менше h^2 для фіксованого P_b .

Отримані результати демонструють, що за збільшення довжини коду n границя коригувальних властивостей блокових кодів наближається до границі Шеннона.

1. W.E. Ryan and Shu Lin, Channel Codes Classical and Modern. – New York, NY: Cambridge University Press, 2009. 2. Бронштейн И.Н. Справочник по математике для инженеров и учащихся втузов./ Бронштейн И.Н., Семендяев К.А. – М.: Наука. Главная редакция физико-математической литературы, 1981. – 706 с. 3. Зюко А.Г. , Кловский Д.Д. и др. 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.