

ОПТИМІЗАЦІЯ АЛГОРИТМУ ОБЧИСЛЕННЯ ОЦІНОК ІНФОРМАЦІЙНОЇ ЕНТРОПІЇ ВИПАДКОВИХ СИГНАЛІВ ДЛЯ ЦИФРОВИХ ЗАСОБІВ ОБМІНУ ДАНИМИ

Одним з невирішених науково-технічних завдань, пов'язаних з реалізацією апаратного та алгоритмічного забезпечення засобів обміну даними на основі використання оцінок інформаційної ентропії, є дослідження шляхів оптимізації процедури опрацювання широкосмугових сигналів з маніпульованим значенням ентропії.

Розглянуто особливості реалізації методу опрацювання випадкових сигналів з маніпульованим значенням ентропії в комунікаційних каналах комп'ютерних систем. Запропоновано один з варіантів оптимізації алгоритму розрахунку оцінок ентропії, що не потребує реалізації алгоритму розрахунку ряду Тейлора при обчисленні функції логарифму.

В результаті проведених досліджень запропоновано спрощені алгоритми розрахунку оцінок інформаційної ентропії при опрацюванні випадкових сигналів з маніпульованим значенням ентропії.

Ключові слова: оцінки ентропії, алгоритм, випадкові сигнали, маніпуляція.

STEPAN IVANOVYCH MELNYCHUK, MARIYA VOLODYMYRIVNA KOROPETSKA,
IRYNA ZINOVIVNA MANULYAK

Ivano-Frankivsk National Technical University of Oil and Gas

OPTIMIZATION ALGORITHM FOR COMPUTATION ENTROPY ESTIMATIONS OF RANDOM SIGNALS FOR DIGITAL COMMUNICATION DEVICES

Abstract. One of the outstanding scientific and technical problems associated with implementation of hardware and software algorithmic means of data exchange on the basis of information entropy estimation, is to study ways to optimize processing procedures broadband signals manipulated entropy value.

Implementation features of the method of processing random signals manipulated entropy value in computer systems and communication channels. One of the options proposed optimization algorithm of entropy estimates that the algorithm does not require calculation of Taylor series when calculating the logarithm function.

As a research result the proposed simplified algorithms for calculating the estimates in the information processing random signals manipulation entropy value.

Keywords: entropy estimation algorithm, random signals manipulation.

ВСТУП

Однією з умов реалізації якісних розподілених комп'ютерних систем є наявність надійних комунікаційних каналів [1]. Традиційно в сучасних комп'ютерних системах реалізація комунікаційного устаткування ґрунтується на використанні сигналів з одиначною базою. Такий підхід зумовлений швидкою мініатюризацією обладнання комунікаційних каналів та інтенсивним зниженням їх вартості. Слід зазначити, що згадані засоби мають низьку завадостійкість, що неприйнятно в умовах дії інтенсивних завод промислового характеру, а також потребують монопольного використання частотного ресурсу, що суттєво знижує ефективність їх групового використання.

Одним з варіантів покращення згаданих характеристик є перехід до сигналів з розширеною базою, такий підхід дозволяє забезпечити надійний обмін даними при малих співвідношеннях сигнал/завада і, за умови використання кодового розділення, забезпечує суттєве ущільнення інформаційного каналу. Проте реалізація апаратних засобів для формування та оброблення широкосмугових сигналів характеризується значною технічною та алгоритмічною складністю, що обмежує, на даному етапі розвитку, їх широке використання. Таким чином ефективно вирішення окреслених завдань може бути отримане за умови успішного розв'язання науково-технічних проблем, які пов'язані із створенням та дослідженням методів формування і оброблення широкосмугових сигналів, оптимізації відповідних схемотехнічних та алгоритмічних рішень.

1. АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Практично більшість запропонованих способів реалізації обміну даними на основі сигналів з розширеною базою ґрунтується на кореляційному опрацюванні. Такий підхід зумовлює необхідність застосування складних схемотехнічних чи алгоритмічних рішень при формуванні сигналів із заданими кореляційними властивостями, а також значних обчислювальних затрат при їх опрацюванні [2, 3]. Крім того зберігання взірців псевдовипадкових сигналів у кореляторах приводить до необхідності залучення додаткових об'ємів пам'яті.

В роботах [4, 5] розглянуто варіант реалізації обміну даними на основі широкосмугових сигналів з відомим значенням ентропії. В запропонованому підході інформаційною ознакою двійкового елементу цифрового повідомлення є відповідна оцінка ентропії фрагменту випадкових сигналів. Фактично тут реалізовано маніпуляцію ентропією, коли у відповідність до біту інформаційного повідомлення ставиться фрагмент випадкового сигналу із заданим значенням ентропії. Зокрема значення "0" може бути

представлено фрагментом сигналу (амплітуда приймає одне значення) з нульовою оцінкою ентропії, а значення "1" представляється фрагментом сигналу (амплітуда приймає задану кількість значень) з наперед відомою, ненульовою оцінкою ентропії. Основні етапи описаного методу маніпуляції для двійкового інформаційного повідомлення: 10110 подано на рисунку 1.

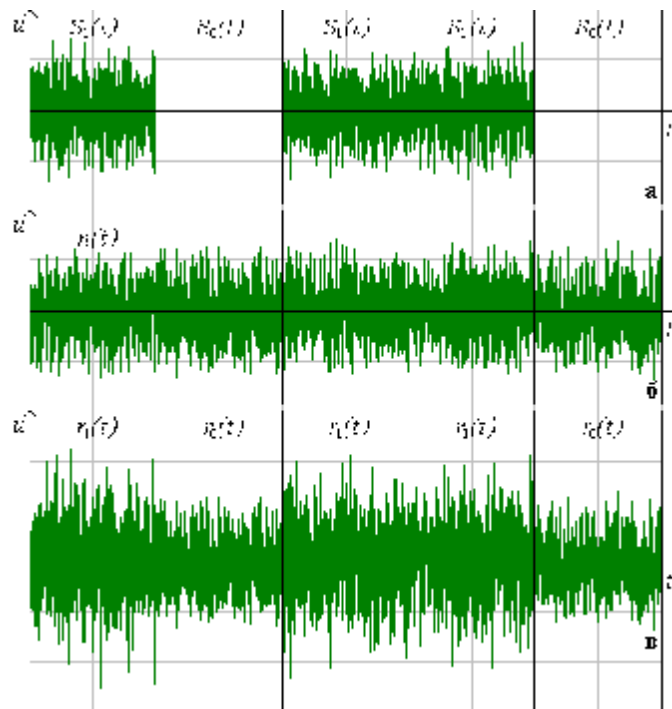


Рис. 1. Сигнальне представлення методу маніпуляції ентропії випадковими сигналами різної потужності

Для представлення нульових елементів інформаційного повідомлення використано пасивну, детерміновану реалізацію сигналу $S_0(t)$ з потужністю $P_0 = 0$ і значенням ентропії $H_0 = 0$. Для представлення одиничних елементів задіяна реалізація випадкового процесу $S_1(t)$ фіксованої потужності P_1 та ентропії H_1 відповідно (рис.1,а).

При проходженні через канал обміну даними сигнали зазнають спотворень, внаслідок дії адитивних завад $n(t)$, які характеризуються власним розподілом, потужністю P_n та ентропією H_n (рис.1,б). Така взаємодія приводить до утворення реалізацій $r_0(t)$ та $r_1(t)$, (рис.1,в). Доцільно зазначити, що в залежності від ступеня кореляції $S_1(t)$ та завади $n(t)$ потужність результуючого сигналу $r_1(t)$ може приймати значення нуля (для інверсних абсолютно корельованих сигналів), суми потужностей (для некорельованих сигналів) тощо. Оцінки ентропії реалізацій таких сигналів змінюються нелінійно, за умови нормального розподілу можливих станів – за логарифмічним законом.

За характеристиками завадостійкості такий метод обміну даними поступається некогерентній модуляції FSK та когерентній модуляції PSK, проте значно простіший з погляду апаратної реалізації [6]. Крім того, при однакових часових затратах, такий спосіб забезпечує опрацювання випадкових сигналів з більшою базою і, як наслідок, дозволяє отримати покращення завадостійкості до 2 дБ. Доцільно також зазначити, що використання випадкових сигналів забезпечує більш рівномірне використання частотної смуги інформаційного каналу і не потребує залучення додаткових об'ємів пам'яті для збереження взірців псевдовипадкових сигналів.

2. ФОРМУЛЮВАННЯ ЦІЛЕЙ РОБОТИ

Одним з невирішених раніше науково-технічних завдань, пов'язаних з реалізацією апаратного та алгоритмічного забезпечення засобів обміну даними на основі згаданого вище способу, є дослідження шляхів оптимізації процедури оброблення широкосмугових сигналів з відомим значенням ентропії. Тобто об'єктом дослідження є спосіб опрацювання випадкових сигналів, зокрема процедура розрахунку оцінок ентропії, а зменшення обчислювальних затрат на реалізацію алгоритму такого опрацювання – основною метою роботи.

3. АПЕРТУРА ОЦІНОК ЕНТРОПІЇ

Дослідження проводились шляхом моделювання. Основним чинником, що діє в каналі обміну даними, вважають стаціонарний адитивний білий гаусів шум AWGN [2, 3]. Отже відповідні сигнали та завади, в межах поданого дослідження, формувались, зберігались та оброблялись в цифровому форматі з такими параметрами: кількість рівнів квантування 65535, частота дискретизації 48 кГц, ширина частотної смуги сигналів та завад 24 кГц, швидкість передачі війкових інформаційних символів від 2 до 4800 біт/с.

Визначення апертури оцінок ентропії \hat{H}_i (рис.2) послідовних фрагментів випадкового сигналу проведено шляхом багаторазових розрахунків для різних розмірів вибірок на основі змодельованих тестових вибірок, узагальнені результати подано в таблицях 1 та 2. Як можна побачити, з отриманих результатів, апертури оцінок ентропії випадкових сигналів поставлених у відповідність “0”-го та “1”-го інформаційних символів не перевищують 0,08 та 0,09 відповідно.

Таблиця 1

Зміна характеристик \hat{H}_i в залежності від розміру вибірки для випадкових сигналів “0”-го інформаційного символу

Розмір вибірки (відліків)	$\max \hat{H}_i$	$\min \hat{H}_i$	$\Delta \hat{H}_i$
512	9,0010	8,9261	0,0752
1024	9,9842	9,9091	0,0750
2048	10,9340	10,8700	0,0640
4096	11,8370	11,7800	0,0570
8192	12,6640	12,6113	0,0527
16384	13,3560	13,3130	0,0430
32768	13,8594	13,8251	0,0343
65535	14,1740	14,1520	0,0220

Фактично флуктуація в околі усередненого значення складає 0,04 та 0,09 і є не меншою за 0,02 та 0,015 для оцінок ентропії відповідних інформаційних ознак.

Таблиця 2

Зміна характеристик \hat{H}_i в залежності від розміру вибірки для випадкових сигналів “1”-го інформаційного символу

Розмір вибірки (відліків)	$\max \hat{H}_i$	$\min \hat{H}_i$	$\Delta \hat{H}_i$
512	8,9981	8,9090	0,0890
1024	9,9672	9,8822	0,0850
2048	10,9030	10,8200	0,0830
4096	11,7722	11,6950	0,0772
8192	12,5320	12,4750	0,0570
16384	13,1360	13,0862	0,0498
32768	13,5421	13,5072	0,0349
65535	13,7760	13,7570	0,0190

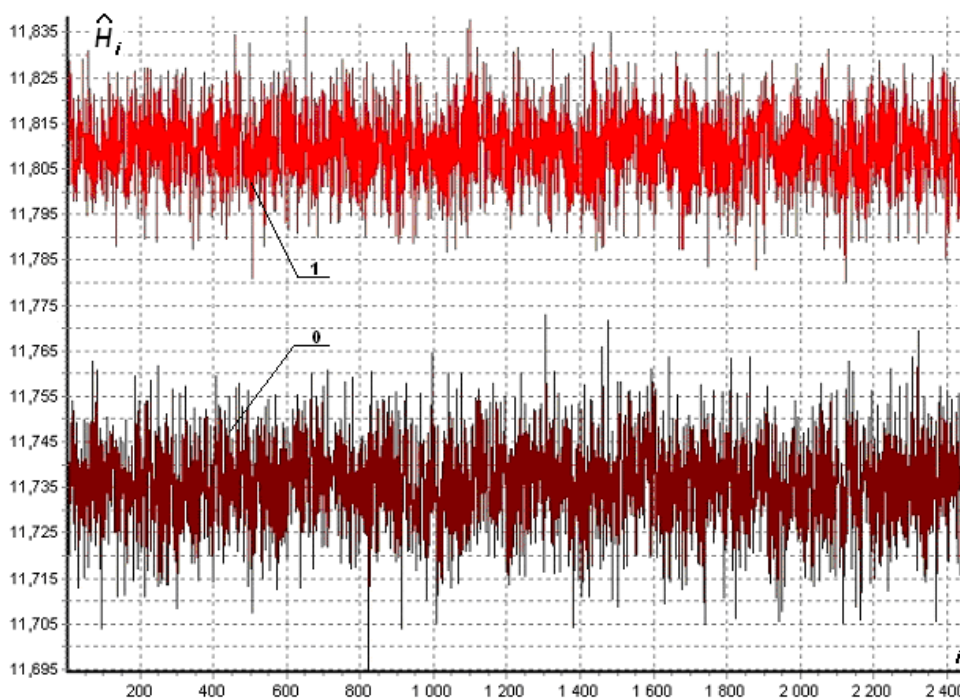


Рис. 2. Коливання \hat{H}_i послідовних фрагментів випадкового сигналу, які відповідають бітам інформаційного повідомлення

Оцінка дисперсії для поданих в таблицях 1 та 2 складає 0.0007 та 0.0004 відповідно. З огляду на отримані результати доцільно зазначити, що при реалізації процедури обчислення оцінок ентропії розглянутих вище типів сигналів достатньо забезпечити точність розрахунку не нижче 10^{-3} чи 10^{-4} .

Таким чином оптимізація алгоритму обчислення може бути реалізована за рахунок допустимої втрати точності розрахунку \hat{H}_i .

4. ОСОБЛИВОСТІ РЕАЛІЗАЦІЇ АЛГОРИТМУ РОЗРАХУНКУ ЕНТРОПІЇ

Традиційно [6], розрахунок оцінок ентропії може проводитись за формулами Шеннона

$$\hat{H} = -\sum_{i=1}^m p(r_i) \cdot \log_2 p(r_i), \text{ де } p(r_i) - \text{імовірність появи стану сигналу } r_i; \text{ Хартлі } \hat{H} = \log_2 m, \text{ де } m -$$

кількість станів сигналу; диференційної ентропії $\hat{H} = \log_2 \sqrt{2pes^2}$. Доцільно зазначити, що використання формули диференційної ентропії забезпечує прийнятну ефективність тільки у випадку нормально розподілених ймовірностей станів вхідного сигналу, що практично не зустрічається в експлуатаційних умовах.

Розглянувши подані вирази можна відзначити, що елементом, який потребує найбільше обчислювальних ресурсів є функція логарифму, оскільки її розрахунок ґрунтується на розкладі в ряд Тейлора [7]:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots = \sum_{n=0}^{\infty} \frac{(-1)^n \cdot x^{n+1}}{n+1} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} \cdot x^n}{n}$$

для усіх $|x| < 1$.

Однією з неприйнятних особливостей такої реалізації є суттєве зростання необхідної кількості обчислень членів коли $x \rightarrow 0$. Результати підрахунку необхідної кількості математичних операцій (додавання, віднімання, множення та ділення), які необхідно виконати при обчисленні логарифму для діапазону від 1 до 0, що представляє ймовірності появи відповідних станів випадкового сигналу, подано на рисунку 3.

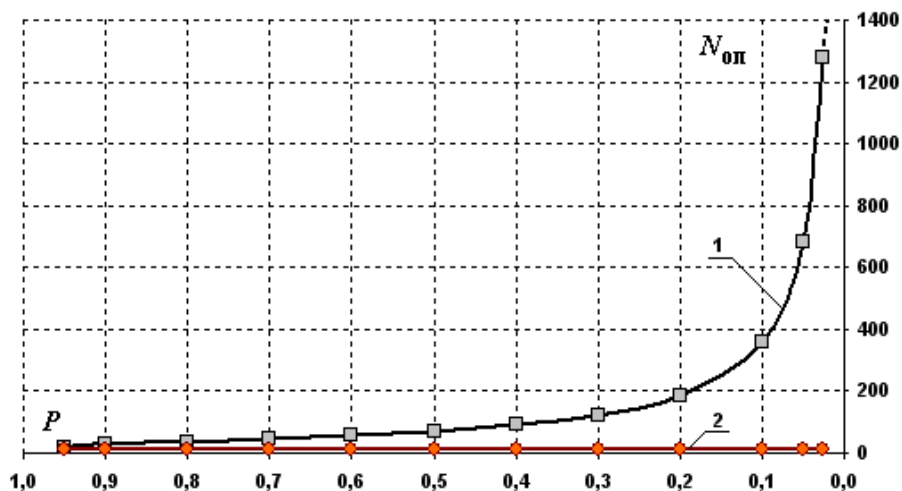


Рис. 3. Залежність кількості математичних операцій від ймовірності появи станів при розрахунку за: 1 - формулою Шеннона через ряд Тейлора, 2 - системою поліноміальних функцій 3-го степеня

Як можна побачити, при наближенні ймовірності до 0 кількість операцій, які необхідні для обчислення ентропії з використанням ряду Тейлора зростає непропорційно, що практично унеможливує ефективну апаратну чи програмну реалізацію.

Одним із шляхів оптимізації розрахунку є відмова від здійснення розрахунків, що реалізуються на основі ряду Тейлора, що неминуче приведе до втрати точності. Оскільки ймовірність появи окремих станів r_i оброблюваного сигналу охоплює інтервал $[0,1]$ то графічна інтерпретація для складових формули Шеннона $p(r_i) \cdot \log_2 p(r_i)$ матиме вигляд, наведений на рисунку 4.

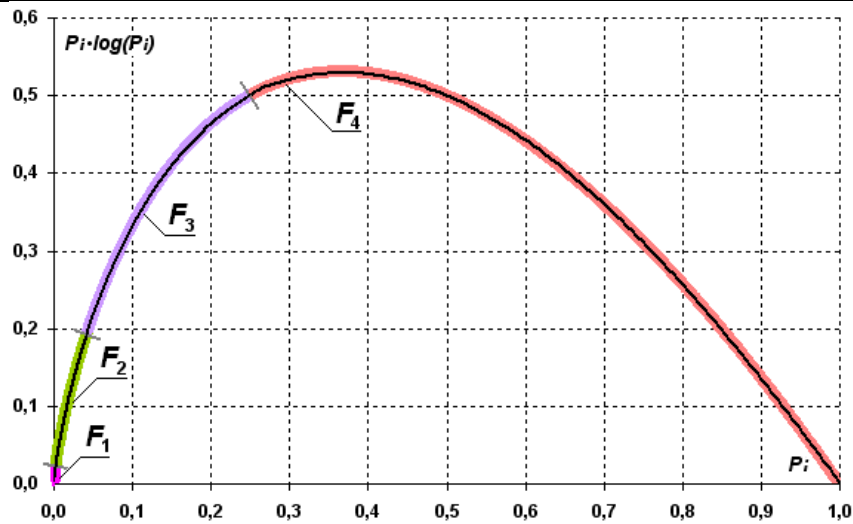


Рис. 4. Залежність $p(r_i) \cdot \log_2 p(r_i)$ формули Шеннона від зміни $p(r_i)$ з розділенням на чотири сплайни

На основі проведеного регресійного аналізу встановлено, що одним із найпростіших варіантів заміни складової $p(r_i) \cdot \log_2 p(r_i)$ формули Шеннона для обчислення ентропії є апроксимація сплайнами третього порядку з нерівномірним розділенням інтервалу $[0,1]$:

$$\begin{cases} 0.0000 \leq P_i \leq 0.0030; & F_1(P_i) \\ 0.0030 < P_i \leq 0.0445; & F_2(P_i) \\ 0.0445 < P_i \leq 0.2550; & F_3(P_i) \\ 0.2550 < P_i \leq 1.0000; & F_4(P_i) \end{cases} \quad (1)$$

де $F_1(P_i) = 332272.3468 \cdot P_i^3 - 2068.47238 \cdot P_i^2 + 11.73700 \cdot P_i + 0.00002$

$$F_2(P_i) = 651.89677 \cdot P_i^3 - 81.79213 \cdot P_i^2 + 6.70684 \cdot P_i + 0.00002$$

$$F_3(P_i) = 13.15949 \cdot P_i^3 - 11.14588 \cdot P_i^2 + 3.74297 \cdot P_i + 0.05596$$

$$F_4(P_i) = 0.69757 \cdot P_i^3 - 2.52898 \cdot P_i^2 + 1.57670 \cdot P_i + 0.25690$$

Результати оцінок абсолютної похибки, що виникає при обчисленні оцінок ентропії за виразом 1, подано на рисунку 5.

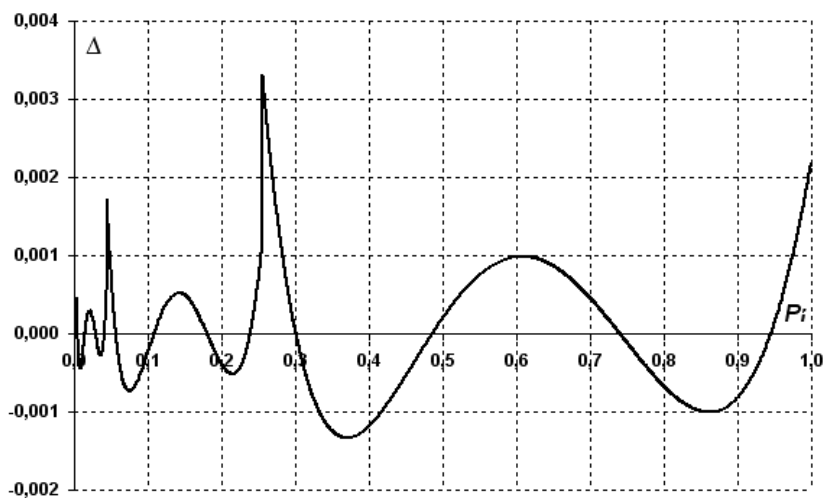


Рис. 5. Абсолютна похибка розрахунку ентропії за виразом 1

Як можна побачити реалізований аналітичний опис 1 дозволяє забезпечити абсолютну похибку обчислення практично на порядок меншу по відношенню до флуктуації, що спостерігається в околі усередненого значення оцінок ентропії відповідних інформаційних ознак.

Блок-схему реалізації відповідного алгоритму на основі аналітичного виразу 1 для обчислення ентропії за формулою К.Шеннона подано на рисунку 6.

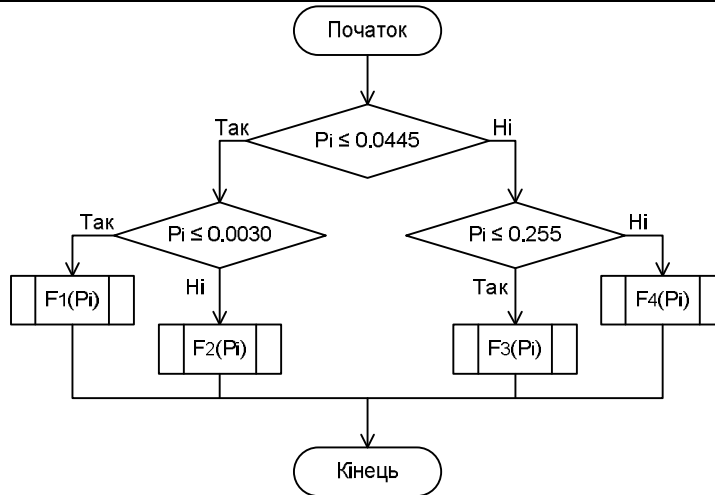


Рис. 6. Блок-схема алгоритму обчислення ентропії за виразом (1)

Наступне зменшення абсолютної похибки розрахунку можна досягнути шляхом збільшення степеня сплайнів та їх кількості. Результати регресійного аналізу для сплайнів четвертого степеня подано на рисунку 7.

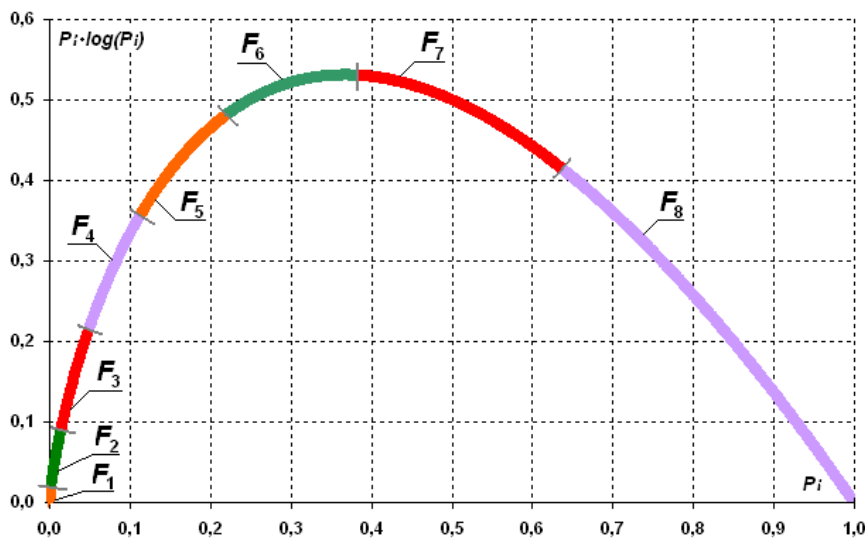


Рис. 7. Залежність $p(r_i) \cdot \log_2 p(r_i)$ формули Шеннона від зміни $p(r_i)$ з розділенням на 8 сплайнів

$$\left\{ \begin{array}{l} 0.0000 \leq P_i \leq 0.0025; \quad F_1(P_i) \\ 0.0025 < P_i \leq 0.0155; \quad F_2(P_i) \\ 0.0155 < P_i \leq 0.5000; \quad F_3(P_i) \\ 0.5000 < P_i \leq 0.1150; \quad F_4(P_i) \\ 0.1150 < P_i \leq 0.2240; \quad F_5(P_i) \\ 0.2240 < P_i \leq 0.3895; \quad F_6(P_i) \\ 0.3895 < P_i \leq 0.6410; \quad F_7(P_i) \\ 0.6410 < P_i \leq 1.0000; \quad F_8(P_i) \end{array} \right. \quad (2)$$

де

$$\begin{aligned} F_1(P_i) &= -186569246.460937 \cdot P_i^4 + 1265118.366985 \cdot P_i^3 - 3501.057660 \cdot P_i^2 + 12.403319 \cdot P_i + 0.000002 \\ F_2(P_i) &= -239489.921162 \cdot P_i^4 + 1234.313603 \cdot P_i^3 - 295.789851 \cdot P_i^2 + 8.368012 \cdot P_i + 0.002379 \\ F_3(P_i) &= -4061.107343 \cdot P_i^4 + 780.766753 \cdot P_i^3 - 72.532461 \cdot P_i^2 + 6.297200 \cdot P_i + 0.010342 \\ F_4(P_i) &= -234.353792 \cdot P_i^4 + 114.664654 \cdot P_i^3 - 27.545219 \cdot P_i^2 + 4.886321 \cdot P_i + 0.027784 \\ F_5(P_i) &= -26.183010 \cdot P_i^4 + 26.426930 \cdot P_i^3 - 13.178840 \cdot P_i^2 + 3.817640 \cdot P_i + 0.058489 \end{aligned}$$

$$F_6(P_i) = -4.338916 \cdot P_i^4 + 7.943741 \cdot P_i^3 - 7.211768 \cdot P_i^2 + 2.945031 \cdot P_i + 0.107304$$

$$F_7(P_i) = 0.908594 \cdot P_i^4 + 2.796879 \cdot P_i^3 - 4.275817 \cdot P_i^2 + 2.189704 \cdot P_i + 0.181280$$

$$F_8(P_i) = -0.223510 \cdot P_i^4 + 1.096563 \cdot P_i^3 - 2.675455 \cdot P_i^2 + 1.512272 \cdot P_i + 0.290126$$

Результати оцінок абсолютної похибки, що виникає при обчисленні оцінок ентропії за виразом 1, подано на рисунку 8.

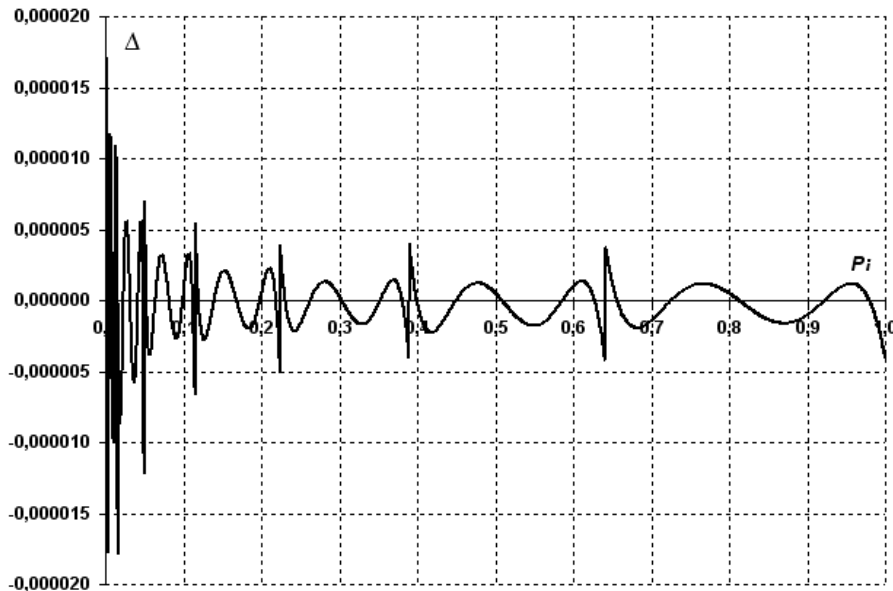


Рис. 8. Абсолютна похибка розрахунку ентропії згідно аналітичного виразу (2)

Як можна побачити отриманий аналітичний опис 2 дозволяє забезпечити абсолютну похибку обчислення практично на два порядки меншу по відношенню до флуктуації, що спостерігається в околі усередненого значення оцінок ентропії відповідних інформаційних ознак.

Блок-схему реалізації відповідного алгоритму на основі 2 для обчислення ентропії за формулою К.Шеннона подано на рисунку 9.

При реалізації обчислень за розробленими алгоритмами, рисунки 6 та 9, виникає необхідність виконання двох і трьох операцій порівняння відповідно, проте реалізації деревоподібної структури забезпечує суттєве зниження обчислювальних затрат, рис.3 крива 2.

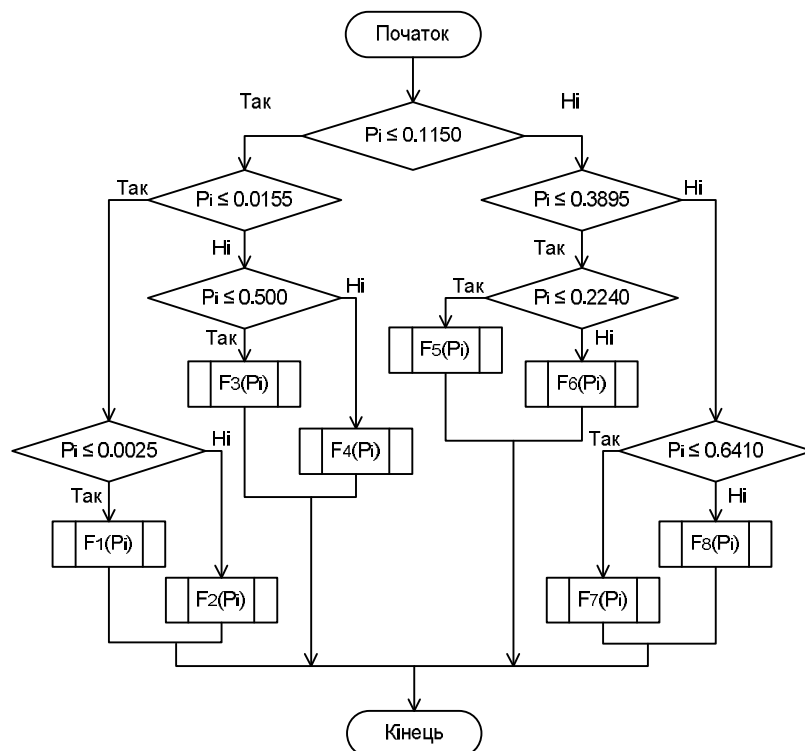


Рис. 9. Блок-схема алгоритму обчислення ентропії згідно аналітичного виразу 2

Висновки

У результаті проведених досліджень запропоновано спрощені алгоритми розрахунку оцінок ентропії при опрацюванні випадкових сигналів з відомим значенням ентропії, які дозволяють забезпечити прийнятну точність при суттєвому зниженні обчислювальних затрат. Коректність отриманих аналітичних виразів підтверджується тим, що величина абсолютної похибки менша по відношенню до флуктуації, що спостерігається в околі усередненого значення оцінок ентропії відповідних інформаційних ознак менша на 1 і 2 порядки відповідно.

Література

1. Бойко Ю.М. Аналіз енергетичної та спектральної ефективності сигналів у цифрових телекомунікаційних системах. / Ю.М. Бойко, О.І. Єрьоменко, Ю.Б. Ісаєнко // Вісник Хмельницького національного університету. Технічні науки. – Хмельницький. – 2013. – №6. – С. 147-160.
2. Скляр Бернард. Цифровая связь. Теоретические основы и практическое применение. / Бернард Скляр; пер с англ. – 2-е изд., испр.). – М. : Издательский дом «Вильямс», 2003. – 1004 с.
3. Прокис Дж. Цифровая связь / Дж. Прокис ;пер. с англ.; под ред. Д.Д. Кловского. – М.: Радио и связь, 2000. – 598 с.
4. Пат. 81017 Україна, МПК(2006) Н04В 1/69. Спосіб передавання та приймання інформації / Мельничук С.І., Козленко М.І. (Україна). – заявка № а 2005 08893; заявл. 19.09.2005; опубл. 26.11.2007, Бюл. № 19.
5. Мельничук С.І. Теорія ентропійних методів маніпуляції сигналів / С.І. Мельничук // Тези доповідей проблемно-наукової міжгалузевої конференції: Інформаційні проблеми комп'ютерних систем, юриспруденції, економіки та моделювання (ПНМК-2009). – Бучач: Інститут менеджменту і аудиту, 2009. – 6 с.
6. Козленко М.І. Дослідження завадостійкості способу передавання та приймання інформації, зокрема, способів формування та обробки широкосмугових сигналів. / М.І. Козленко, С.І. Мельничук // Електроніка та зв'язок. . – 2007. – № 2(37). – С. 82-92. – ISSN 1811-4512.
7. Жураковський Ю.П. Теорія інформації кодування: підручник / Ю.П. Жураковський, В.П. Полторак. – К.: Вища школа., 2001. – 255с.

References

1. J.M. Boiko, O.I. Eromenko, J.B. Isaenko. Analiz enerhetuchnoi ta spektralnoi efektyvnosti syhnaliv u cyfrovyykh telekomunikaciynnykh systemakh. Visnyk Khmelnytskoho natsionalnoho universytetu. Technical sciences. Khmelnytsky. 2013. Volume 6. Issue 1. pp. 147-160.
2. Sklyar Bernard. Cifrovaya svyaz'. Teoreticheskie osnovy i prakticheskoe primeneniye. Izd. 2-e, ispr. Moskow, Izdatelskij dom «Vil' yams», 2003, 1004 p.
3. Prokis Dzh. Cyfrovaya svyaz'. Ed. by D.D. Klovs'koho, Moskow, Radio i svyaz', 2000, 598 p.
4. Pat. 81017 Ukraina, MPK (2006) N04V 1/69. Sposib peredavannia ta pryimannia informacii / Melnychuk S.I., Kozlenko M.I. (Ukraina). – zaivka №a 2005 08893; zaivl. 19.09.2005; opubl. 26.11.2007, Biul. №19.
5. Melnychuk S.I. Teoriia entropiinyx metodiv manipuliacii syhnaliv. Tezu dopovidei problemno-naukovoi mizhhaluzevoi konferencii. Informaciini problemu komp'iuternyx system, jursprudenicii, ekonomiky ta modeliuvannia (PNMK-2009). Buchach. 2009. – p. 6.
6. M.I. Kozlenko, S.I. Melnychuk. Doslidzhennia zavadostiikosti sposobu peredavannia ta pryimannia informacii, zokrama sposobiv formuvannia ta obrobky shyrokosmuhovyyx suhnaliv. Elektronika ta zv'iazok. Kyiv. 2007. №2(37). pp. 82-92. ISSN 1811-4512.
7. Zhurakovskiy J.P., Poltorak V.P. Teoriia informacii koduvannia: [pidruchnik]. Kyiv, Vushcha shk., 2001, 255 p.

Рецензія/Peer review : 12.5.2015 р.

Надрукована/Printed :15.5.2015 р.

Рецензент: д.т.н., проф. Заміховський Л.М.