



МЕТОД ЗМЕНШЕННЯ НАДЛИШКОВОСТІ ГАРМОНІЧНИХ СИГНАЛІВ НА ОСНОВІ КОДІВ ПОЛЯ ГАЛУА

Наталія Яцків

Інститут комп'ютерних інформаційних технологій,
Тернопільська академія народного господарства,
вул. Львівська, 11, 46004, Тернопіль, Україна,
e-mail: ny@tanet.edu.te.ua

Резюме: в статті запропонований та досліджений метод зменшення надлишковості гармонічних сигналів на основі адаптивного кодування в базисі Галуа, який базується на реєстрації $1/8$ періоду та аналізі параметрів сигналу. Даний метод дозволяє зменшити об'єми даних, які реєструються на об'єктах електроенергетики.

Ключові слова: стиснення даних, гармонічний сигнал, біти Галуа, кодування.

1. ВСТУП

Швидкий розвиток інформаційних систем і мереж, до яких належать розподілені комп'ютерні системи, обчислювальні мережі, системи збору та обробки даних, зумовлює істотне збільшення інформаційних потоків між територіально розподіленими джерелами та користувачами повідомлень, а також ставить важливу задачу зберігання постійно зростаючих об'ємів інформації.

Для підвищення ефективності використання комунікаційних та інформаційно-обчислювальних ресурсів вказаних систем застосовують методи зменшення надлишковості даних.

Пошуку шляхів стиснення даних на основі ефективного кодування та теоретико - числових перетворень (ТЧП) присвячено багато наукових робіт вітчизняних та зарубіжних вчених. У більшості робіт розв'язання задачі зменшення надлишковості даних відбувається з використанням базису Радемахера.

Розробка і дослідження ефективних методів стиснення цифрових даних, які краще адаптовані до системних характеристик квазістаціонарних джерел інформації та забезпечують меншу чутливість стиснених даних до збоїв та помилок в процесі їх передавання, зберігання та цифрової обробки є актуальною науково-технічною задачею.

2. МЕТОДИ СТИСНЕННЯ ТЕХНОЛОГІЧНИХ ДАНИХ

В даний час широко досліджуються методи стиснення звукових повідомлень, зображення, відеоданих про що свідчать чисельні публікації, але значно менше уваги приділяється розробці спеціалізованих методів стиснення технологічних даних, що є необхідною умовою для вдосконалення автоматизованих комп'ютерних систем керування технологічними об'єктами.

При первинній обробці та стисненні технологічних даних в основному використовують методи на основі приростів [1, 2]:

- логарифмічна імпульсно-кодова модуляція (ІКМ);
- адаптивна ІКМ;
- адаптивна диференціальна ІКМ;
- блочна ІКМ;
- дельта модуляція (ДМ);
- адаптивна дельта-модуляція.

При дельта-модуляції в канал зв'язку передається не абсолютне значення сигналу, а різниця між вихідним аналоговим сигналом і апроксимуючим значенням, внаслідок чого ДМ є чутливою до завад. Основний недолік ДМ полягає в тому, що дельта-кодер не встигає

відслідковувати швидкі зміни рівня сигналу, внаслідок чого виникає переваження по крутизні.

При обробці оцифрованих значень технологічних сигналів використовують алгоритми (Лемпеля–Зіва, Хафмана, арифметичне кодування), які лежать в основі архіваторів (pkzip, arj, lha) і програм динамічного стиснення дисків (Stacker, DoubleSpace).

Основна ідея: друге і наступне входження деякого рядка символів у повідомлення замінюються посиланням на його першу появу в повідомленні [3].

Використання універсальних методів стиснення даних не забезпечує високих коефіцієнтів стиснення, так як вони не враховують структуру повідомлення.

Перспективним в цьому напрямі є розробка методів стиснення даних з використання дискретних базисів та теоретико-числових перетворень.

В той же час, застосування відомих базисів для реалізації методів стиснення даних, в тому числі базисів Радемахера, Крейга, Уолша, перетворень Адамара та інших нерекурентних ТЧП, на сьогодні вже майже вичерпало свої можливості по їх вдосконаленню та розширенню сфери застосування.

Практика розробки теоретичних основ вертикальної інформаційної технології на основі рекурентних ТЧП в базисі Галуа відкриває нові можливості створення ефективних методів зменшення надлишковості цифрових даних та розвитку сфери їх застосування для ряду інформаційних потоків, що характеризуються специфічними властивостями. Наприклад, зі стрибкоподібними змінами амплітуди спектральних, кореляційних та ентропійних характеристик, що, як відомо, приводить до різкого зниження ефективності кодування існуючими методами, а в окремих випадках, до збільшення об'ємів даних по відношенню до початкових.

Крім того, відомі методи стиснення даних практично не захищені від збоїв і помилок, особливо вставок і стирань в біт-орієнтованих потоках, що ускладнює і знижує ефективність їх застосування в системах реального часу, які передбачають керування складними технологічними об'єктами.

3. ТЕОРЕТИЧНІ ОСНОВИ ЗАПРОПОНОВАНОГО МЕТОДУ

Особливо гостро проблема використання ресурсів пам'яті постала при розробці цифрового

реєстратора миттєвих значень струмів та напруг у аварійних режимах роботи електричної мережі.

В запропонованому методі ми працюємо з масивом цифрових даних але при цьому враховуємо структуру даних до процесу аналого-цифрового перетворення.

Значення контрольованих параметрів електроенергетичних систем та мереж описуються гармонічними сигналами.

Нехай маємо один період синусоїди (рис.1), яка описується рівнянням

$$x_1(t) = A_0 \cdot \sin(\omega \cdot t),$$

де A_0 - амплітуда; ω - частота; t - час.

Побудуємо на цьому ж графіку (рис.1) косинусоїди

$$x_2(t) = A_0 \cdot \cos(\omega \cdot t),$$

$x_3(t) = A_0 \cdot \cos(\omega \cdot t + \pi)$ і синусоїду

$$x_4(t) = A_0 \cdot \sin(\omega \cdot t + \pi).$$

Таким чином, отримаємо систему рівнянь:

$$\begin{cases} x_1(t) = A_0 \cdot \sin(\omega \cdot t) \\ x_2(t) = A_0 \cdot \cos(\omega \cdot t) \\ x_3(t) = A_0 \cdot \cos(\omega \cdot t + \pi) \\ x_4(t) = A_0 \cdot \sin(\omega \cdot t + \pi) \end{cases}$$

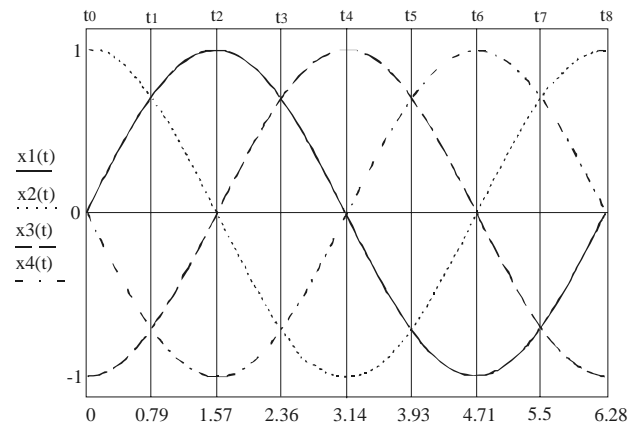


Рис. 1—Чотирьохфазний гармонічний сигнал

Представлені функції мають ряд особливих точок (рис.1), залежно від параметру t :

$$\begin{aligned} t = 0, & \quad x_2(t) = A_0 \\ t = 1, & \quad x_1(t) = x_2(t) \\ t = 2, & \quad x_1(t) = A_0 \\ t = 3, & \quad x_1(t) = x_3(t) \\ t = 4, & \quad x_3(t) = A_0 \\ t = 5, & \quad x_3(t) = x_4(t) \\ t = 6, & \quad x_4(t) = A_0 \\ t = 7, & \quad x_4(t) = x_2(t) \end{aligned} \quad (1)$$

Оскільки функція є періодичною, то в точці $t = 8$ маємо повторення $t = 0$.

При подачі на вхід порогової схеми (рис.2) вхідного синусоїдального сигналу з довільною фазою і пороговим приймачем, описаного системою рівнянь (1), на виході отримаємо послідовність імпульсів (рис.3). Порогова схема спрацьовує в момент, коли значення сигналу дорівнює значенню похідної, це і є перша особлива точка. В цей момент починається кодування сигналу.

В результаті гармонічний синусоїдальний сигнал можна закодувати, використовуючи сумісне представлення в базисі Радемахера і Галуа наступним чином:

$$\{A_0, t_0, t_1, t_2, t_3, t_4, t_5, t_6, t_7\}.$$

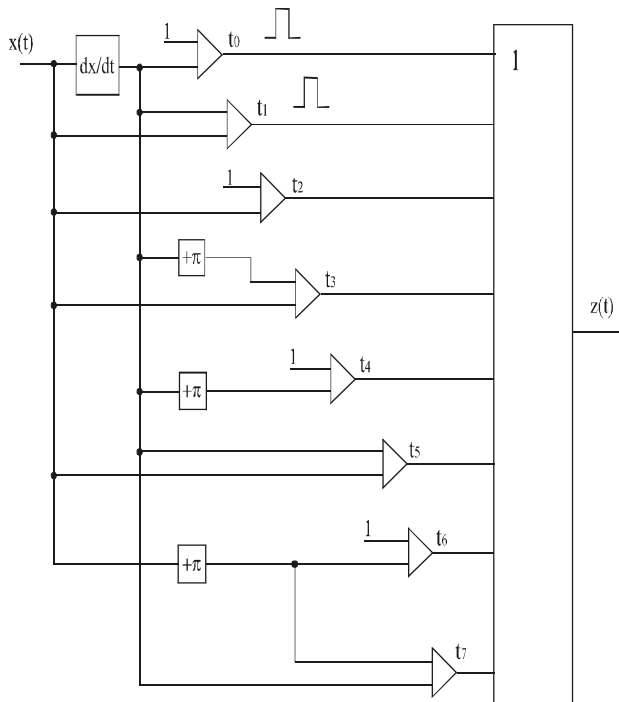


Рис. 2 – Порогова схема:

$z(t)$ – реакція порогової схеми на вхідний синусоїдальний сигнал.

У кожній з восьми особливих точок спрацьовує один із компараторів порогової схеми порівняння. Це вказує на те, що амплітуда та частота контрольованого параметру не змінилися.

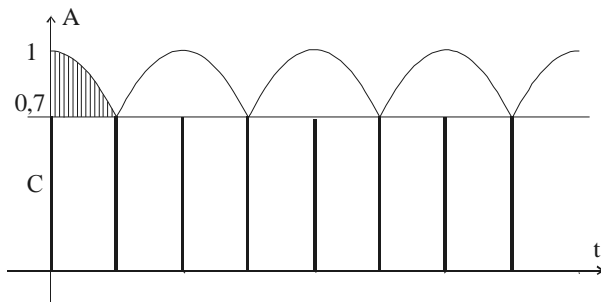


Рис. 3 – Реакція порогової схеми на вхідний синусоїдальний сигнал

Декодування гармонічного сигналу відбувається на основі алгоритму (рис.4):

$$x(t) = \begin{cases} t_0 \leq t < t_1, & f_0(t) = \int_0^t (z_0(t) + C) dt; \\ t_1 \leq t < t_2, & f_1(t) = \overrightarrow{C + z_0(t)}; \\ t_2 \leq t < t_3, & f_2(t) = C + z_0(t); \\ t_3 \leq t < t_4, & f_3(t) = \frac{d(z_0(t) + C)}{dt}; \\ t_4 \leq t < t_5, & f_4(t) = \frac{d(z_0(t) + C)}{dt}; \\ t_5 \leq t < t_6, & f_5(t) = 0 - \overrightarrow{(z_0(t) + C)}; \\ t_6 \leq t < t_7, & f_6(t) = 0 - (z_0(t) + C); \\ t_7 \leq t < t_8, & f_7(t) = 0 - \left(\int_0^t \overrightarrow{(z_0(t) + C)} dt \right). \end{cases} \quad (2)$$

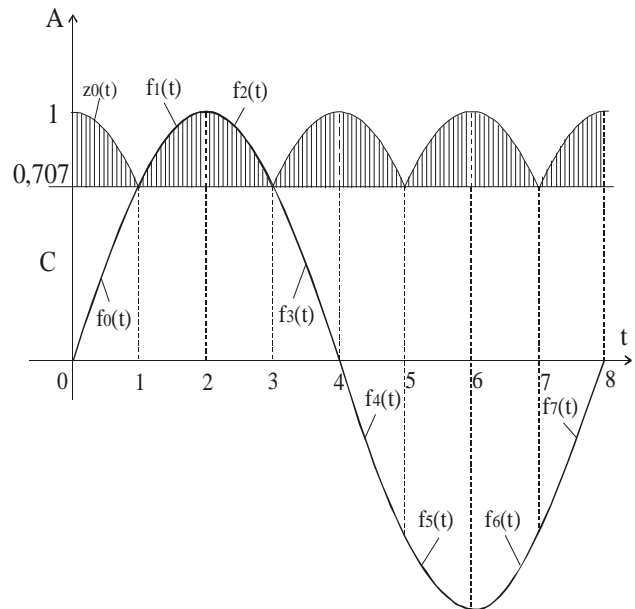


Рис.4. Декодування синусоїди на основі фрагменту $z_0(t)$.

В результаті, всі ці функції можуть бути зведені до одного фрагменту (рис.5), за допомогою якого можна однозначно відновити гармонічний сигнал.

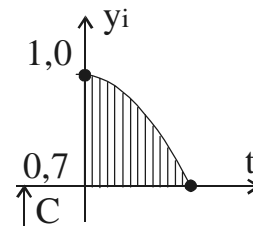


Рис.5 – Фрагмент сигналу

$$x_i(t) = C \pm F(y_i),$$

де $C=0,707$.

Інформація на виході кодера на інтервалі періоду синусоїди представляється двома масивами: A_{z_0} – амплітуда синусоїди, яка вважається незмінною на протязі періоду, і масив кодів $\{t_i\}$, що відповідають моментам зміни функцій в алгоритмі (1).

На основі системи рівнянь (2) однозначно відновлюється значення синусоїди, якщо в пам'яті записано наступний масив:

$$\{A_{z_0}, t_i\}.$$

Оцінимо об'єм даних джерела при описаному способі кодування:

$$I_x = \hat{E}[\log_2 z_{i\max}] + 7 \cdot \hat{E}[\log_2 T],$$

де $z_{i\max}$ – максимальне значення фрагмента сигналу $x(t)$; T – максимальна тривалість одного періоду синусоїди.

Наприклад, при передаванні сигналів стандартами ІКМ $A_{0,x} = 256$, частота дискретизації 8 кГц.

При кодуванні описаним способом

$$z_{i\max} = (1 - C) \cdot A_{0,x},$$

де $C = \frac{\sqrt{2}}{2}$, $z_{i\max} = 75$.

Звідки $\hat{E}[\log_2 z_{i\max}] = \hat{E}[\log_2 75] = 7$ біт.

Якщо частота сигналу F_c змінюється від 10 Гц до 8000 Гц, відповідно період T змінюється від 100000 до 125 мкс.

Для кодування t_i необхідно 17 біт.

$$\hat{E}[\log_2 100000] = 17 \text{ біт.}$$

Таким чином об'єм даних буде дорівнювати

$$I = 7 + 17 \cdot 8 = 143 \text{ біт.}$$

Аналіз запропонованого методу кодування показує, що реєструвати всі інтервали t_i недоцільно, їх можна буде обчислити в процесорі декодера, якщо прийняти допущення, що час t_i є базовим, і на протязі одного періоду не змінюється, тобто $t_1 = t_2 = t_3 = \dots = t_8$.

В даному випадку код синусоїди буде представлений двома кодами:

$\{z_{i\max}, t_1\}$ а об'єм даних, який однозначно описує параметри синусоїди буде дорівнювати:

$$I_2 = 7 + 17 = 24 \text{ біт.}$$

Дане представлення має три недоліки:

1) маніпуляція сигналу синусоїди чотирма фазами не дозволяє синхронізуватися (визначити перехід функції через 0);

2) кодування моментів часу t в базисі Радемахера приводить до значної надлишковості;

3) не дозволяє відслідковувати зміни сигналу на протязі періоду.

Тому пропонується метод кодування синусоїдальних сигналів, який базується на наступному: при переході синусоїди через 0, і додатній похідній, що відповідає рівнянням:

$$\begin{cases} x(t) = 0 \\ dx/dt > 0 \end{cases}$$

відбувається реєстрація значень сигналу і, паралельно, відбувається обчислення похідної. В момент рівності зареєстрованого значення сигналу і обчисленого значення похідної спрацьовує перший компаратор порогової схеми порівняння і реєструється або передається в канал зв'язку біт Галуа. Обчислені значення похідної до цього моменту формують фрагмент

$Z_i, Z_i = \frac{dx(t)}{dt}$ який передається в канал зв'язку,

після біту Галуа. Далі відбувається порівняння реальних значень сигналу з обчисленими згідно алгоритму декодування, що описується системою рівнянь (2). При спрацюванні кожного наступного компаратора порогової схеми в канал зв'язку поступає біт Галуа. Якщо не спрацював один із компараторів – біт Галуа інвертується. Після інвертованого біту в канал зв'язку передають l – реальних значень вимірюваного параметру. Величина l – визначається частотою

дискретизації $l = \frac{l_d}{8}$, l_d – кількість точок

дискретизації на період гармонічного сигналу. З зареєстрованих значень за допомогою перетворень кодування – декодування формується новий фрагмент, який подається на наступний компаратор схеми порівняння. Якщо даний компаратор спрацьовує то зареєстрований фрагмент вважається базовим для всіх наступних етапів кодування і в канал подається прямий біт Галуа. Якщо схема не спрацьовує, тобто форма сигналу відрізняється від форми гармонічного сигналу то біт Галуа інвертується і за ним в канал зв'язку поступають l значень реального сигналу.

4. ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ КОДУВАННЯ

Експериментальні дослідження запропонованого методу кодування гармонічних сигналів проведено на основі даних знятих цифровим реєстратором аналогових сигналів розробленим в Інституті мікропроцесорних систем керування

об'єктами електроенергетики (м. Львів), який забезпечує реєстрацію (осцилографування) миттєвих значень електричних параметрів (струмів, напруг) на первинному електричному обладнанні в передавальних і аварійних режимах.

Перетворення аналогового сигналу в цифровий здійснюється за допомогою 12-розрядного АЦП з частотою дискретизації $f_d = 2000$ Гц.

Оцінимо об'єм даних джерела при заданих параметрах:

$$I = n \cdot l_d,$$

де n – розрядність АЦП; l_d – кількість вибірок сигналу, $l_d = f_d \cdot T$; T – період сигналу в електричній мережі.

$$\text{При } n = 12, l_d = 2000 \cdot 0.02 = 40.$$

Для кодування одного періоду гармонічного сигналу необхідно:

$$I = 12 \cdot 40 = 480 \text{ біт.}$$

При кодуванні вказаних сигналів запропонованим методом для $n = 12$:

$$Z_i = \left(1 - \frac{\sqrt{2}}{2}\right) \cdot 2^n = 0.293 \cdot 4096 = 1200.$$

Звідки

$$n_1 = \hat{E}[\log_2 Z_i] = \hat{E}[1200] = 11 \text{ біт.}$$

Об'єм даних фрагмента (рис.5) буде дорівнювати:

$$I_1 = \left(\hat{E}\left[\frac{l_d}{8}\right] \cdot n_1\right) = 55 \text{ біт.}$$

При застосуванні запропонованого методу для потокового кодування, параметр часу t_i можна представити послідовністю Галуа типу: 11101000 [4].

Якщо на протязі одного періоду гармонічний сигнал не змінив свою форму, тобто спрацювали всі вісім компараторів ми отримаємо наступний масив даних [5]:

$$G Z_1 Z_2 Z_3 Z_4 Z_5 G G G G G G G,$$

перший прямий біт Галуа G вказує на те, що далі слідує масив значень кодового фрагменту Z_i , за допомогою якого відбувається декодування сигналу. Наступні прямі біти Галуа вказують на те, що декодування відбувається згідно алгоритму.

Об'єм даних для вищевказаних параметрів сигналу дорівнює:

$$I_x = 1 + 5 \cdot 11 + 1 + 1 + 1 + 1 + 1 + 1 + 1 = 63 \text{ біт.}$$

Коефіцієнт стиснення дорівнює

$$k_c = \frac{I}{I_x} = \frac{480}{63}.$$

Застосування даного алгоритму дозволяє реєструвати гармонічний сигнал із зменшенням надлишковості в 7.6 рази на протязі одного періоду порівняно з початковим об'ємом даних без втрат точності.

Осцилограми комплексного значення напруги $x(t)$ (рис.6, рис.7), описують реальний перехідний процес, що відбувається під час аварійних станів на об'єктах електроенергетики. Рис.6 відображає початок перехідного процесу, а рис. 7 момент стабілізації перехідного процесу і переходу до нормального режиму роботи.

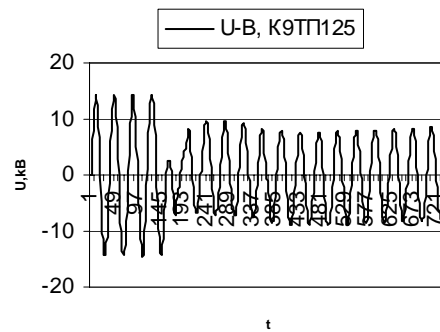


Рис. 6 – Осцилограма сигналу в момент аварії на об'єкті електроенергетики

В результаті кодування сигналу представленого на рис.6. отримаємо потік даних: $G\{Z_i\}G \dots \bar{G}\{y_i\}\bar{G}\{y_i\} \dots G\{Z_i\}G \dots \bar{G}\{y_i\}$.

Розрахуємо об'єм даних при кодуванні запропонованим методом.

Для кодування одного масиву $\{Z_i\}$ використовується $I_1 = 55$ біт, одного масиву $\{y_i\}$ – $I_2 = n \cdot l = 12 \cdot 5 = 60$ біт.

Кількість бітів Галуа обчислюється за формулою:

$$I_G = \frac{t_r}{T} \cdot 8, \quad I_G = 152 \text{ біт,}$$

де t_r – час реєстрації.

Об'єм даних для кодування перехідного процесу (рис.6) обчислюється за формулою:

$$I_\Sigma = c_1 \cdot I_1 + c_2 \cdot I_2 + I_G,$$

c_1 – кількість фрагментів $\{Z_i\}$ в закодованій послідовності, $c_1 = 3$;

c_2 – кількість фрагментів $\{y_i\}$ в закодованій послідовності, $c_2 = 96$.

$$I_\Sigma = 6076 \text{ біт.}$$

Початковий об'єм даних $I_0 = 8892$ біт.

Коефіцієнт стиснення дорівнює:

$$k_c = \frac{I_0}{I_\Sigma}; k_c = 1,46.$$

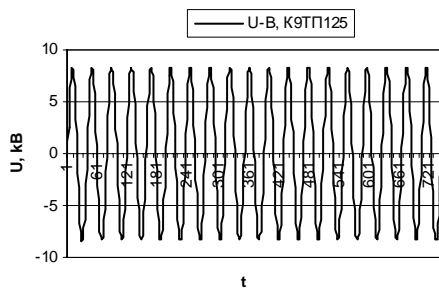


Рис. 7– Осцилограма сигналу після аварії на об'єкті електроенергетики

При кодуванні сигналу представленого на рис.7 отримаємо потік даних:

$$G\{Z_i\}GGGGGGG\dots GGGGGG\dots GGGGGG,$$

його об'єм обчислюється за формулою

$$I_\Sigma = c_1 \cdot I_1 + I_G, I_\Sigma = 207,$$

при $c_1 = 1$.

Коефіцієнт стиснення дорівнює

$$k_c = \frac{I_0}{I_\Sigma}, k_c = 42,96.$$

5. ВИСНОВКИ

Приведені розрахунки коефіцієнта стиснення для різних фрагментів зареєстрованого сигналу показали, що коефіцієнт стиснення залежить від кількості періодів синусоїди з однаковими параметрами, чим таких періодів більше тим більший коефіцієнт стиснення.

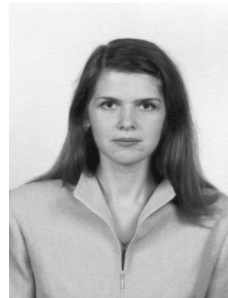
Кодування бітами Галуа гарантує цілісність пакету даних і дає пікову синхронізацію синусоїди, тому що ця послідовність має особливу автокореляційну функцію.

При впливі завад типу “стирання” або вставок інформаційних бітів та кодування часу розрядно-позиційним кодом є неефективним, в той час, як кодування шляхом інверсії бітів Галуа дозволяє однозначно прив'язати суттєві відліки Z_i до фактичних значень часу, незалежно від попередніх помилок.

СПИСОК ЛІТЕРАТУРИ

- [1] Орищенко В.И. Санников. В.А., Свириденко В.А. Сжатие данных в системах сбора и передачи информации / Под ред. В.А. Свириденко. – М.: Радио и связь, 1985. – 184 с.
- [2] Свириденко В. А. Анализ систем со сжатием данных. М.: Связь, 1977. –744 с.

- [3] J.Ziv , A.Lempel. Compression of individual sequences via variable rate coding. IEEE transactions on Information Theory. Vol. IT-24. N.5, September 1978, pp. 530-535.Ватолин Д., [4] Ратушняк А., Смирнов М., Юкин В. Методи сжатия данных. Устройство архиваторов, сжатие изображений и видео. – М.: Диалог-Мифи, 2002. –384 с.
- [5] Муттер В.М. Основы помехоустойчивой телепередачи информации. – Л.: Энергоатомиздат. Ленингр. отд-ние, 1990.- 268 с.
- [6] Яцків Н.Г. Методи стиснення даних в інформаційно-керуючих системах // Розвідка і розробка нафтових і газових родовищ. Серія: Технічна кібернетика та електрифікація об'єктів паливно-енергетичного комплексу. №37. – Івано-Франківськ.: ІФДТУНГ, 2001. – С. 183-186.



Наталія Яцків у 1997р. закінчила Івано-Франківський державний технічний університет нафти і газу за спеціальністю “Методи та прилади неруйнуючого контролю для технічної і медичної діагностики”.

У 2002 р. закінчила аспірантуру при Івано-Франківському національному технічному університеті нафти і газу за спеціальністю “Обчислювальні машини, системи та мережі”.

У даний час працює старшим викладачем кафедри спеціалізованих комп'ютерних систем Тернопільської академії народного господарства.

Наукові інтереси: компресія даних, теоретико-числові перетворення.