

Національний університет “Львівська політехніка”

Яцків Наталія Георгіївна

УДК 004. 627+004.415

**МЕТОДИ ТА ЗАСОБИ СТИСНЕННЯ ДАНИХ
В РОЗПОДІЛЕНИХ КОМП’ЮТЕРНИХ СИСТЕМАХ
НА ОСНОВІ КОДІВ ПОЛЯ ГАЛУА**

Спеціальність 05.13.13 – обчислювальні машини, системи та мережі

АВТОРЕФЕРАТ

дисертації на здобуття наукового ступеня

кандидата технічних наук

Львів – 2003

Дисертацією є рукопис.

Робота виконана в Івано-Франківському Національному технічному університеті нафти і газу Міністерства освіти і науки України.

Науковий керівник: доктор технічних наук, професор
Николайчук Ярослав Миколайович,
Тернопільська академія народного господарства,
завідувач кафедри спеціалізованих комп'ютерних систем.

Офіційні опоненти: доктор технічних наук, професор
Черкаський Микола Вячеславович,
Національний університет “Львівська політехніка”,
професор кафедри обчислювальної техніки;

доктор технічних наук, старший науковий співробітник **Русин Богдан Павлович**,
фізико-механічний інститут ім. Г. В. Карпенка НАН України, завідувач відділу методів і систем обробки, аналізу й ідентифікації зображень.

Провідна установа: Інститут кібернетики ім. В. М. Глушкова НАН України, м. Київ, відділ мікропроцесорної техніки.

Захист відбудеться “26” вересня 2003 р. о 16 годині на засіданні спеціалізованої вченої ради Д 35.052.05 при Національному університеті “Львівська політехніка” за адресою: 79013, м. Львів, вул. С. Бандери, 12.

З дисертацією можна ознайомитись у науково-технічній бібліотеці Національного університету “Львівська політехніка” за адресою: 79013, м. Львів, вул. Професорська, 1.

Автореферат розісланий “_____” серпня 2003 р.

*Вчений секретар
спеціалізованої вченої ради,
д. т. н., професор*

Федасюк Д. В.

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Швидкий розвиток інформаційних систем і мереж, до яких належать розподілені комп'ютерні системи (РКС), обчислювальні мережі, системи збору та обробки даних, зумовлює істотне збільшення інформаційних потоків між територіально розподіленими джерелами та користувачами повідомлень, а також породжує задачу зберігання постійно зростаючих обсягів інформації.

Для підвищення ефективності використання комунікаційних та інформаційно-обчислювальних ресурсів вказаних систем застосовують методи зменшення надлишковості даних. Це дає змогу значно розвантажити канали зв'язку та системи обробки і зберігання даних.

Пошуку нових шляхів стиснення даних на основі ефективного кодування та теоретико-числових перетворень (ТЧП) присвячено багато наукових робіт. Зокрема необхідно відзначити значний внесок відомих вчених: Харкевича О. О., Вошні О. Г., Макса Ж., Хартлі Р., Рабінера Л., Рейдера У., Голда Б., Акушського І. Я., Николайчука Я. М., Ольховського Ю. Б.

В той же час застосування найпоширеніших базисів для реалізації методів зменшення надлишковості даних, в тому числі базисів Радемахера, Крейга, Уолша, перетворень Адамара та інших нерекурентних ТЧП, на сьогодні вже майже вичерпало свої можливості щодо їх вдосконалення та розширення сфери застосування.

Практика розробки теоретичних основ вертикальної інформаційної технології на основі рекурентних ТЧП у базисі Галуа відкриває нові можливості створення ефективних методів зменшення надлишковості цифрових даних і 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, призводить до різкого зниження ефективності кодування існуючими методами, а в окремих випадках до збільшення обсягів даних стосовно початкових. Крім того, відомі методи стиснення даних практично не захищені від збоїв і помилок, особливо вставок і стирань у біт-орієнтованих потоках, що ускладнює і знижує ефективність їх застосування в системах реального часу, які передбачають керування складними технологічними об'єктами.

Тому розробка, дослідження і розвиток ефективних методів стиснення цифрових даних, котрі краще адаптовані до системних характеристик специфічних квазістаціонарних джерел інформації (ДІ) та забезпечують меншу чутливість стиснених даних до збоїв і помилок у процесі їх передавання, зберігання і цифрової обробки, а також створення відповідних апаратних та програмних засобів стиснення даних, є актуальною науково-технічною задачею.

Зв'язок роботи з науковими програмами, планами, темами.

Дисертаційна робота виконана в рамках плану наукових досліджень, що проводяться кафедрою комп'ютерних технологій і систем управління (КТ і СУ) Івано-Франківського Національного технічного університету нафти і газу (ІФНТУНГ) та Карпатським державним центром інформаційних засобів і технологій НАН України (КД ЦІЗІТ) за держбюджетними темами ІФНТУНГ:

– Д-60 НДР “Розробка теоретичних основ вертикальної інформаційної технології та вдосконалення методів автоматизованого контролю розподілених об'єктів нафтогазовидобувних підприємств”, № 01940001278;

– НДР “Розробка методологічних основ (теорії, моделей, алгоритмів, процедур і технічних засобів) діагностування і автоматизованого управління об'єктами нафтогазового комплексу України” (ТЗ ИГАУ-573/99);

– кафедрою спеціалізованих комп'ютерних систем (СКС) Тернопільської академії народного господарства (ТАНГ): НДР “Розробка теорії, методології побудови та технічних засобів спеціалізованих комп'ютерних систем”, № 0102U00576А;

– науково-технічною програмою КД ЦІЗІТ та інституту мікропроцесорних систем керування об'єктами електроенергетики (ІМСКОЕ): “Створення концепції інформаційно-керуючих систем контролю та обліку енергоносіїв, систем керування технологічними процесами електроенергетичних об'єктів та захисту їх обладнання”, Львів, 2000 р.

Усі вищезазвані роботи виконувались за безпосередньою участю автора.

Мета і задачі дослідження. Метою роботи є вирішення актуальної науково-технічної задачі розвитку методів стиснення даних на основі розробки теоретичних положень, методів, технічних і програмних засобів зменшення надлишковості цифрових даних у базисі Галуа.

Для досягнення мети необхідно вирішити наступні задачі:

– проаналізувати існуючі методи та інформаційні технології стиснення даних у комп'ютерних системах на основі представлення потоків повідомлень в різних базисах: унітарному, Крестенсона, Хаара, Крейга, Радемахера та Галуа;

– визначити взаємозв'язок між системними параметрами та станами джерел інформації, що характеризують ентропію;

– провести теоретичне та експериментальне дослідження, вдосконалити існуючі та розробити нові методи стиснення даних на основі кодів поля Галуа;

– дослідити можливість використання нових принципів стиснення даних для зменшення надлишковості в розподілених комп'ютерних та телекомунікаційних системах;

– здійснити апробацію запропонованих методів в діючих розподілених комп'ютерних системах;

– розробити і впровадити технічні засоби та інформаційні технології для

реалізації розроблених методів стиснення даних в промисловості.

Об'єкт досліджень – низові розподілені комп'ютерні мережі автоматизованих систем зі стаціонарними та квазістаціонарними джерелами інформації.

Предмет дослідження – методи та засоби стиснення даних в базисі Галуа, які забезпечують ефективне зменшення надлишковості цифрових даних в РКС.

Методи дослідження. Для розробки методів і засобів стиснення даних у комп'ютерних системах використовувались методи теорії інформації, статистичного аналізу, теорії кодування у різних базисах і прикладної теорії цифрових автоматів.

Дослідження інформаційних технологій кодування даних у базисі Галуа, проведені з метою визначення оптимальних умов їх застосування, ґрунтуються на методах ентропійних досліджень.

Під час проведення експериментальних досліджень та обробки їх результатів використовувались методи теорії ймовірності, а також методи цифрового моделювання і програмної емуляції роботи цифрових процесорів обробки даних. Розробка технічних засобів здійснювалась з використанням методів схемотехніки.

Наукова новизна одержаних результатів:

– вдосконалено логіко-статистичні інформаційні моделі шляхом заміни унітарного коду послідовностями Галуа, що дало змогу кодувати мітки часу без введення надлишковості;

– вперше запропоновано базисні функції Галуа нульового та першого порядку, на основі яких розроблено нові методи стиснення цифрових даних, які забезпечують зменшення обсягів даних у 3–8 разів при кодуванні процесів, для яких класичні методи практично непрацездатні, що створює умови їх ефективного використання у стандартних системах передавання, архівації і зберігання даних;

– вперше запропоновано метод стиснення даних, представлених гармонічними сигналами, що базується на реєстрації $1/8$ періоду сигналу, та кодуванні повторів бітами Галуа, коефіцієнт стиснення дорівнює 7,6 раза;

– розроблено модифікований метод адаптивного кодування даних в базисі Галуа, що забезпечує стиснення даних у 1,6-2 рази при 50% активних відліків і не вводить надлишковості на відміну від існуючих при 90% активних відліків;

– розроблено новий метод кодування багатовимірних джерел інформації на основі перетворення Крестенсона–Галуа, який дає змогу перейти від багатовимірного до одновимірного представлення даних, забезпечує зменшення службових даних при передаванні інформації, криптографічне кодування і захист від помилок.

Практичне значення одержаних результатів:

– розроблено та впроваджено технічні засоби і технології, які базуються на

використанні запропонованих методів стиснення даних, що дало змогу підвищити ефективність роботи РКС;

– розроблено алгоритми та програми стиснення даних, що суттєво зменшують надлишковість і тим самим знижують собівартість формування, зберігання і передавання інформації у комп'ютерних мережах;

– розроблено та виготовлено дослідний взірець інтегрально-імпульсного перетворювача з розширеними функціональними можливостями, що дало можливість розширити сферу його застосування.

Теоретичні і практичні результати дисертаційної роботи використані та впроваджені:

– у дослідно-конструкторських роботах, що проводились КД ЦІЗІТ НАН України та ІМСКОЕ “Створення концепції інформаційно-керуючих систем контролю та обліку енергоносіїв, систем керування технологічними процесами електроенергетичних об'єктів та захисту їх обладнання”;

– на ВАТ “Івано-Франківськцемент” у комп'ютерній системі контролю та обліку витрати енергоносіїв “ALFIYA”;

– у навчальному процесі кафедри СКС Тернопільської академії народного господарства при проведенні лекційних, практичних і лабораторних занять з курсів “Компресія даних”, “Низові комп'ютерні мережі” та “Системи передавання даних”.

Особистий внесок здобувача. Основні положення і результати дисертаційної роботи отримані автором самостійно. У публікаціях, написаних у співавторстві, здобувачеві належить: [2] – архітектура комп'ютерної системи ідентифікації семантичних, технологічних та інформаційних станів; [3] – аналітичні залежності коефіцієнта стиснення даних багатоканальних комп'ютерних систем збору технологічних даних; [4] – обґрунтування і класифікація ДІ та визначення їхніх системних характеристик з точки зору квазістаціонарних властивостей, розрахунок ентропії квазістаціонарних ДІ для різних методів зменшення надлишковості даних; [6] – структура спецпроцесора обробки даних на основі перетворення Крестенсона – Галуа ; [7] – аналіз методів цифрового кодування даних на фізичному рівні комп'ютерних систем ; [8] – дослідження методу стиснення даних на основі кодів поля Галуа в багатоканальних системах; [9] – метод кодування багатовимірних ДІ; аналітичні залежності коефіцієнта стиснення при різних методах кодування багатовимірних ДІ в базисі Крестенсона та Галуа.

Апробація результатів дисертації. Основні результати дисертаційної роботи доповідались та обговорювались на: Міжнародній конференції “Сучасні проблеми радіоелектроніки, телекомунікацій та комп'ютерної інженерії” TCSET`2002 – Львів-Славсько, 2002; Міжнародній науково-практичній конференції “Автоматизація виробничих процесів” МНПК АВТ-2002 – Хмельницький, 2002; Науково-технічній конференції “Вимірювальна та

обчислювальна техніка в технологічних процесах” – Хмельницький, 2002; Міжнародній конференції “Досвід проектування і застосування САПР в мікроелектроніці” CADSM 2003 – Львів-Славсько, 2003, а також доповідалось та обговорювалось на наукових семінарах кафедр КТ і СУ ІФНТУНГ та СКС ТАНГ.

Публікації. Результати досліджень опубліковані у 9 друкованих працях автора, з них 7 статей у фахових журналах та збірниках наукових праць, 2 доповіді у збірниках науково-технічних конференціях.

Структура дисертації. Дисертаційна робота складається із вступу, чотирьох розділів, висновків, списку використаних джерел і додатків. Загальний обсяг роботи становить 177 сторінок, з них 139 сторінок основного тексту, що включає 66 рисунків і 8 таблиць. Список використаних джерел – 83 найменування, додатки – на 29 сторінках.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі обґрунтовано актуальність роботи, визначено її зв'язок з науковими програмами, темами. Сформульовано наукову новизну та практичне значення одержаних результатів, а також показана їх апробація і публікації.

У першому розділі проведено аналіз трирівневої зірково-магістральної архітектури РКС, яка включає інформаційну систему контролю та ідентифікації станів об'єкта керування (ОК), і складається з: сенсорів (C_p); логічних та евристичних аналізаторів семантичних (S_i), технологічних (X_j) та інформаційних (I_k) станів ОК; промислових контролерів (ПрК), низової комп'ютерної мережі (НКМ) та базового комп'ютер-сервера диспетчерського центру (рис. 1). Розроблена архітектура системи ідентифікації станів ОК забезпечує на низовому рівні можливість зменшення обсягів даних на 1–2 порядки за рахунок кодування відповідного стану об'єкта та системи формальних параметрів, які описують стан об'єкта, замість інформації, що формується усіма сенсорами.

Проведені дослідження нафтогазових технологічних установок, як ДІ, показали, що вони належать до класу квазістаціонарних ОК, причому в кожному конкретний момент часу такі об'єкти класифікуються трьома характеристиками станів:

- а) семантичні (S_i ; $i = 1, 2, \dots, s$);
- б) технологічні (X_j ; $j = 1, 2, \dots, q$);
- в) інформаційні (I_k ; $k = 1, 2, \dots, h$),

де s, q, h – відповідна кількість класифікованих станів ОК.

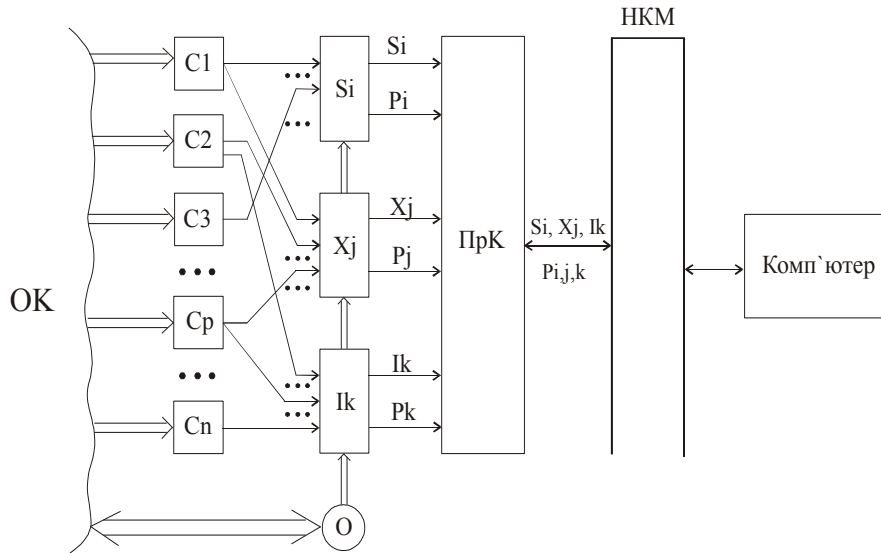


Рис. 1. Архітектура системи ідентифікації семантичних, технологічних та інформаційних станів ОК:

О – оператор технологічної установки; P_i, P_j, P_k – формалізовані параметри станів ОК; P_{ijk} – характеристичний параметр екологічного стану ОК.

Розроблена класифікація методів стиснення даних за область застосування: універсальні, методи стиснення звуку, зображення, технологічних даних (рис. 2).

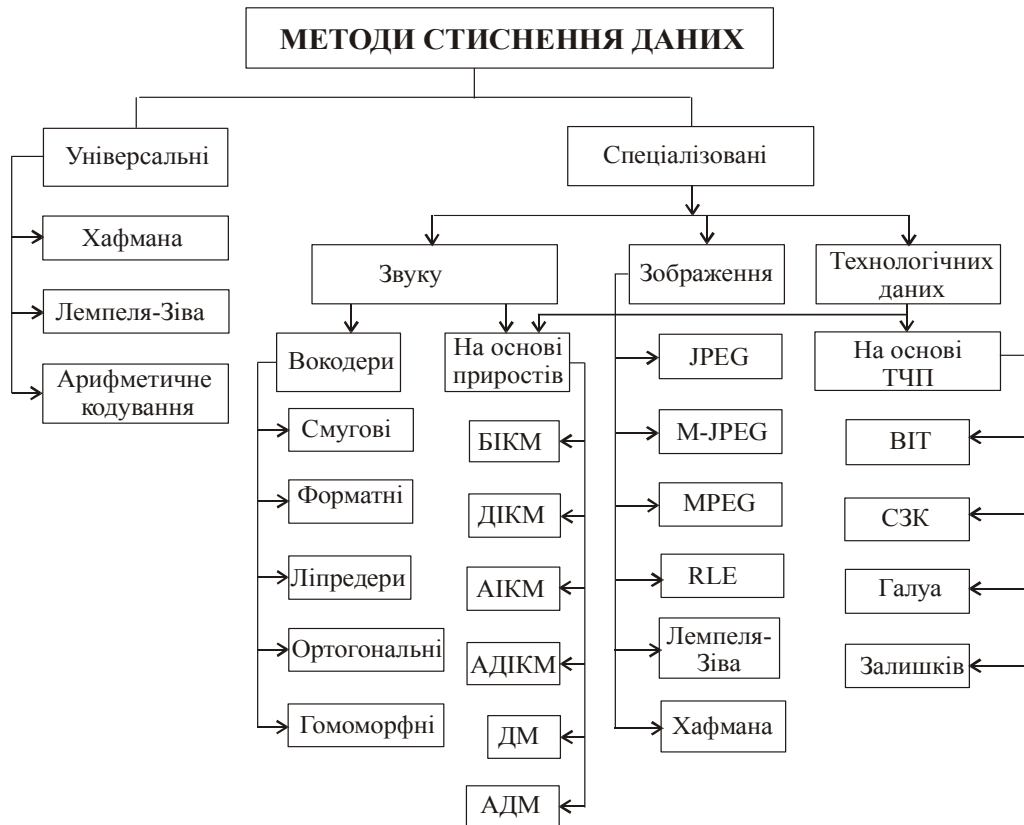


Рис. 2. Класифікація методів стиснення даних: БІКМ – блокова імпульсно - кодова модуляція; ДІКМ – диференціальна ІКМ; АІКМ – адаптивна ІКМ; АДІКМ – адаптивна диференціальна ІКМ; ДМ – дельта-модуляція; АДМ – адаптивна ДМ; RLE – кодування довжин повторів; ВІТ – вертикальна інформаційна технологія; СЗК – система залишкових класів.

У даний час широко досліджуються методи стиснення звукових повідомлень, зображення, але значно менше уваги приділяється розробці спеціалізованих методів стиснення технологічних даних, що є необхідним для вдосконалення автоматизованих комп'ютерних систем керування технологічними об'єктами.

Перспективним у цьому напрямку є розробка методів стиснення даних з використанням дискретних базисів і теоретико-числових перетворень.

На основі розробленої класифікації показано, що методи стиснення даних у базисі Галуа, які становлять предмет дослідження дисертаційної роботи, належать до класу нелінійних ТЧП і найкраще придатні для кодування технологічних даних.

У **другому розділі** вдосконалено логіко-статистичні інформаційні моделі (ЛСІМ) шляхом заміни унітарного коду послідовністю Галуа, що дало змогу фіксувати час без введення надлишковості (рис. 3). ЛСІМ призначені для первинної обробки повідомлень в реальному часі, контролю відхилень та ідентифікації станів ОК, їх застосування дає змогу зменшити обсяг даних до величини $I = m$, де m – кількість відліків, оскільки для кодування кожного відліку використовується один біт.

Запропонована ЛСІМ, призначена для реєстрації відхилення ε контрольованого гармонічного сигналу від передбачуваного ε_0 .

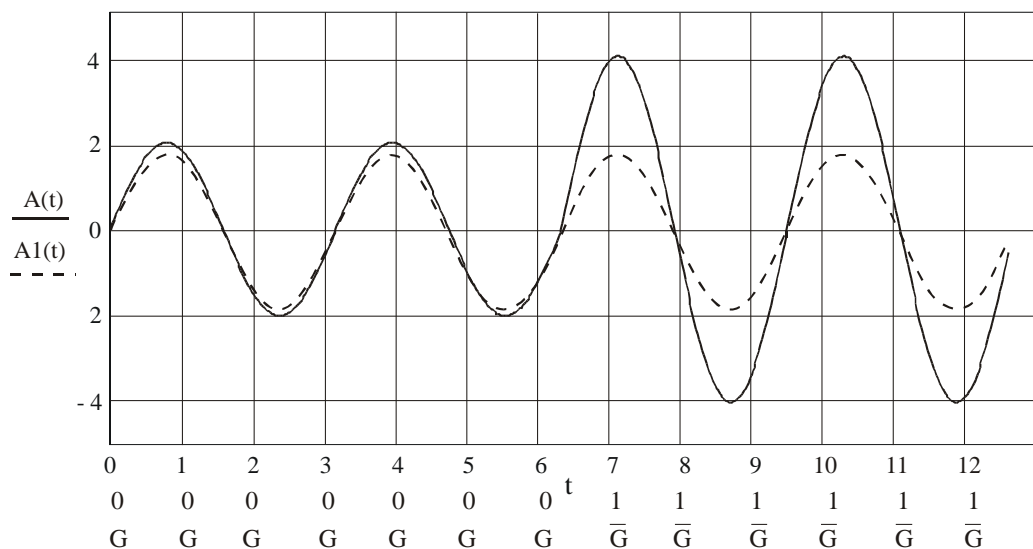


Рис. 3. ЛСІМ: $A(t)$ – вимірний сигнал; $A1(t)$ – передбачуваний сигнал.

Значення наведеної ЛСІМ визначаються за формулою:

$$L_6 = \begin{cases} G, & \varepsilon < \varepsilon_0 \\ \bar{G}, & \varepsilon \geq \varepsilon_0 \end{cases},$$

де G , \bar{G} – відповідно прямий та інвертований біт Галуа.

З метою забезпечення рівномірної швидкості формування повідомлень при використанні вертикальної інформаційної технології розроблені та досліджені

базисні функції Галуа нульового й першого порядку. Кожна з функцій нульового порядку кодується послідовністю Галуа, що генерується фазовим зсувом вихідної послідовності. Квазістаціонарний процес (рис. 4) можна однозначно представити біт-орієнтованою базисною функцією Галуа $y_i = G_i$, що описується лінією.

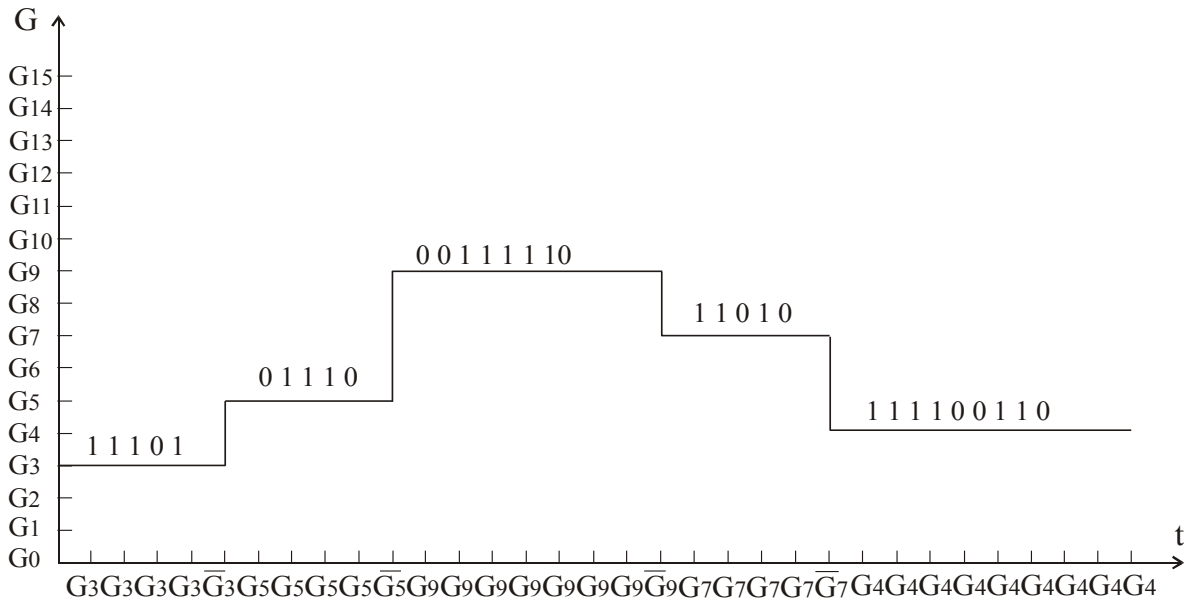


Рис. 4. Кодування на основі базисних функцій нульового порядку.

Запропонований метод дає змогу кодувати інформаційний відлік одним бітом послідовності Галуа при умові, що контрольований об'єкт перебуває у визначеному стані протягом $n_g + 1$ тактів, де n_g – розрядність кодону Галуа.

Даний метод кодування інформаційних потоків у базисі Галуа забезпечує коефіцієнт стиснення даних:

$$k_1 = \hat{E}[\log_2 A],$$

де A – діапазон квантування при $A = 1024$, $k_1 = 10$ (рис. 5).

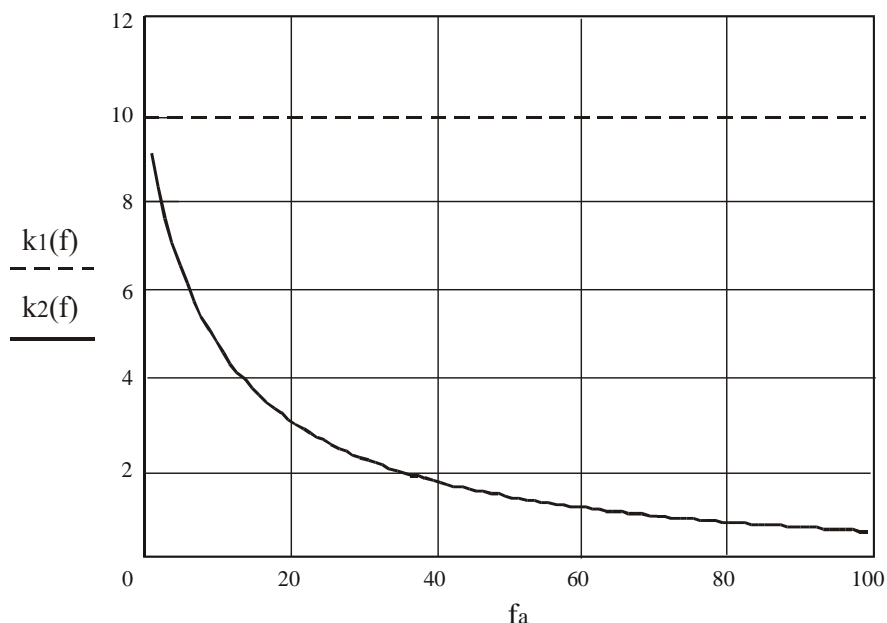


Рис. 5. Залежність коефіцієнта стиснення від кількості активних відліків.

При невиконанні вказаної умови коефіцієнт стиснення визначається за формулою:

$$k_2 = \frac{n \cdot m}{(n+1) \cdot f_a + (m - f_a)},$$

де m – кількість відліків; n – розрядність відліків; f_a – кількість активних відліків.

Розроблені базисні функції Галуа першого порядку (рис. 6), кожна з яких відповідає визначеному градієнту наростання значень функції.

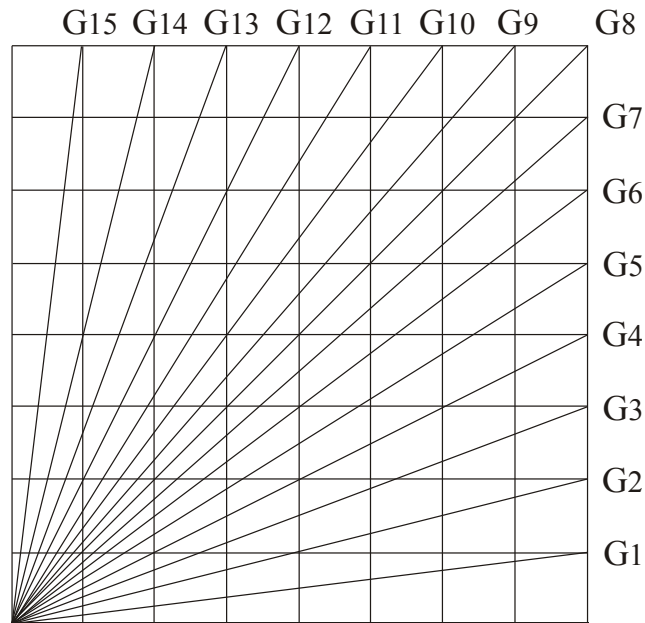


Рис. 6. Базисні функції Галуа першого порядку.

При цьому біт-орієнтоване кодування дискретних функцій в базисі Галуа однозначно можливе для стрибкоподібних і лінійно наростаючих функцій (рис. 7).

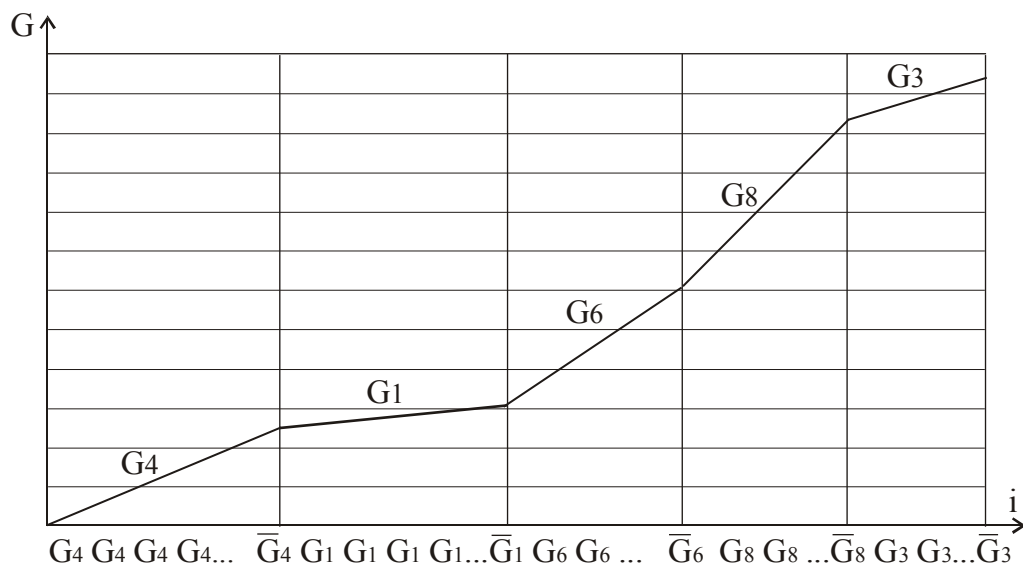


Рис. 7. Кодування на основі базисних функцій першого порядку.

Метод стиснення даних на основі базисних функцій Галуа першого порядку доцільно використовувати для кодування інтегрованих значень параметрів ОК. Кут нахилу лінії визначає швидкість наростання інтеграла. Визначеному куту нахилу однозначно відповідає кодон Галуа, що дає змогу декодувати значення інтеграла функції. Запропонований метод не має недоліку, котрий полягає у постійному зростанні розрядності відліків, а також нерівномірного слідування бітів Галуа.

У третьому розділі розроблено метод стиснення синусоїдальних сигналів, що базується на наступному: при переході синусоїди через 0 і при додатній похідній, що відповідає рівнянням:

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

обчислюється похідна контрольованого сигналу. У момент рівності сигналу і похідної формується кодовий фрагмент Z_i (рис. 8). Для кодування одного відліку фрагмента потрібно $Z = \hat{E}[\log_2 A_0]$ біт, $0 \leq A_0 \leq \frac{\sqrt{2}}{2} \cdot A$, тобто на 1 біт менше, ніж

для кодування амплітуди сигналу $Z = \hat{E}[\log_2 A] - 1$.

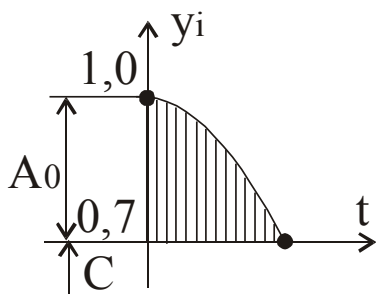


Рис. 8. Фрагмент сигналу $z_0(t)$.

Кодовий фрагмент визначає параметри вимірюваного сигналу протягом періода, якщо вони незмінні, реєструються прямі біти Галуа G і значення фрагмента Z_i . В результаті період гармонічного сигналу представляється двома кодами $\{G, Z_i\}$, на основі яких можна однозначно відновити сигнал.

Якщо відбулися зміни характеристик сигналу, реєструється інвертований біт Галуа \bar{G} і значення реального сигналу.

Коли амплітуда і частота наступного періоду синусоїди не змінилися (похідна, яка визначає амплітуду в нульовій точці $\frac{dx(t)}{dt_{01}} = \frac{dx(t)}{dt_{02}}$), то реєстрація Z_i

не відбувається, що кодується прямим бітом Галуа G .

Декодування гармонічного сигналу (рис. 9) відбувається на основі алгоритму:

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

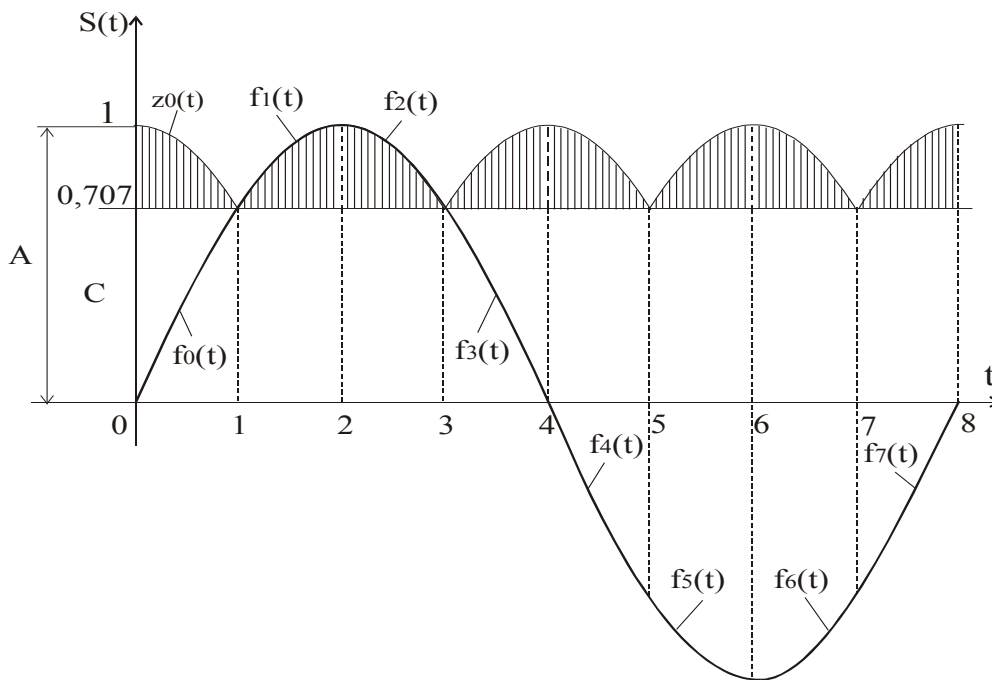


Рис. 9. Декодування синусоїди на основі фрагменту $z_0(t)$.

Коефіцієнт стиснення даного методу рівний $k_c = 7,6$ раз на період і залежить від кількості періодів гармонічного сигналу, що повторюються. Розроблений метод стиснення гармонічних сигналів дав змогу збільшити час реєстрації перехідних процесів на об'єктах електричної мережі без збільшення обсягів пам'яті запам'ятовуючих пристроїв.

З метою визначення ефективних методів кодування в РКС проведено дослідження методів стиснення даних одновимірних і багатовимірних ДІ.

Початковий обсяг даних розраховується за формулою:

$$I_0 = m \cdot \hat{E}[\log_2 A].$$

Отримані аналітичні вирази та графічні залежності (рис. 10) коефіцієнтів стиснення даних при різних методах зменшення надлишковості:

а) однопараметричне адаптивне кодування:

$$K_{ct1} = \frac{m \cdot \hat{E}(\log_2 A)}{f_a \cdot \left(\hat{E}[\log_2 A] + \hat{E}[\log_2 m] \right)};$$

б) адаптивне кодування у базисі Галуа:

$$K_{ct2} = \frac{m \cdot \hat{E}(\log_2 A)}{m + f_a \cdot \hat{E}[\log_2 A]};$$

в) модифікований метод адаптивного кодування у базисі Галуа:

$$K_{ct3} = \frac{m \cdot \hat{E}[\log_2 A]}{\hat{E}[\log_2 m] + f_a \cdot \hat{E}[\log_2 A] + (f_a - 1)}.$$

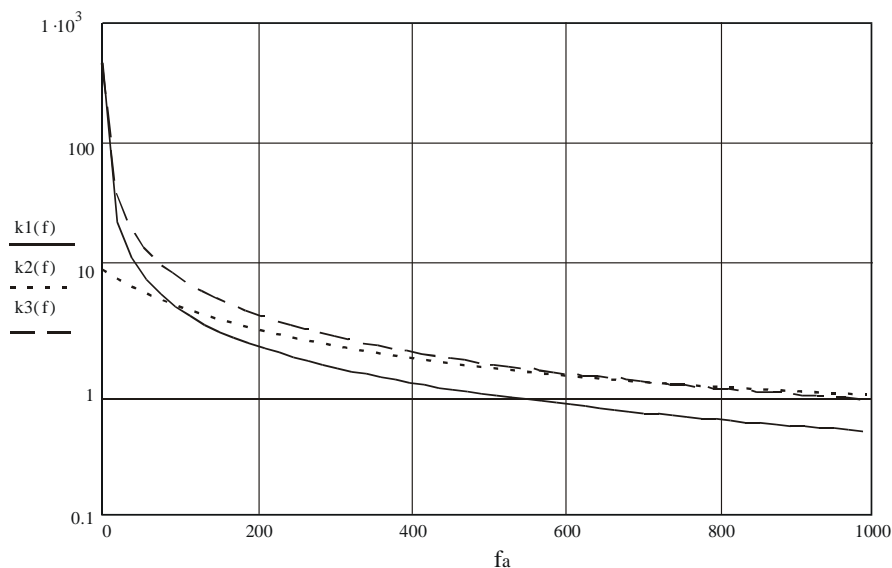


Рис. 10. Залежності коефіцієнта стиснення від кількості активних відліків.

Отже, запропонований метод доцільно використовувати для стиснення даних, що містять понад 50% активних відліків, оскільки відомий метод дає зростання обсягів даних, а не стиснення (рис. 10). За рахунок рекурентних властивостей бітів Галуа формується потік даних з блоковою синхронізацією. Це дає можливість використовувати даний метод у багатоканальних системах збору інформації.

Розроблений метод кодування багатовимірних ДІ в базисі Крестенсона дає можливість виключити надлишковість кодування у тривимірному просторі (рис. 11).

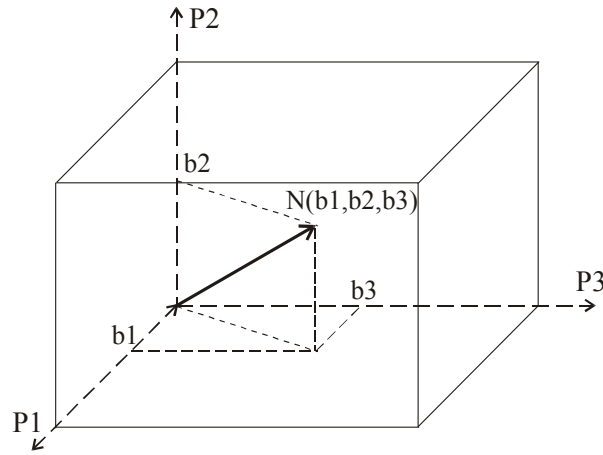


Рис. 11. Кодування тривимірного джерела в базисі Крестенсона.

При цьому вибирається така система взаємопростих модулів P_i , для яких виконуються умови:

$$b_i \leq P_i - 1,$$

де b_i – значення технологічного параметра.

Застосувавши пряме перетворення системи залишкових класів, отримаємо одновимірний вектор:

$$N_k = \text{res} \sum_{i=1}^n B_i \cdot b_i \pmod{\wp},$$

де res – операція отримання залишку; $B_i = \frac{\wp}{P_i} \cdot m_i \equiv 1 \pmod{P_i}$; $\wp = \prod_{i=1}^n P_i$,

$$0 \leq m_i \leq P_i - 1.$$

Важливою задачею є кодування ДІ, що є багатовимірними векторами у вузлах двовимірної площини (рис. 12).

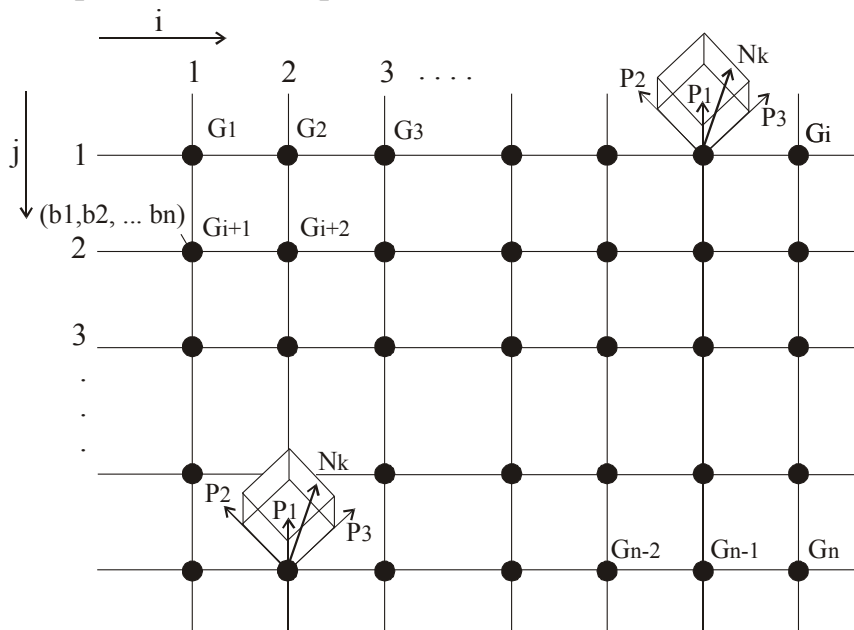


Рис. 12. Модель тривимірного джерела інформації.

До таких задач належать задачі картографії, екологічного моніторингу, розподілу характеристик в перерізах нафто-газопроводів, метеорологічні задачі.

При цьому інформація представляється векторами N_{ijk} . Недоліком такого методу кодування є необхідність доповнення значення N_k коефіцієнтами i та j .

Задача підвищення ефективності кодування таких ДІ може бути вирішена двома шляхами:

а) розширенням системи взаємопростих модулів для внутрішнього кодування i та j ;

б) кодуванням цих координат у базисі Галуа.

У першому випадку пряме перетворення залишкових класів у розширеній системі модулів має вигляд:

$$N_{kij} = \text{res} \left(\sum_{i=1}^{k+2} b_i \cdot \beta_i \pmod{\wp^*} \right),$$

де \wp^* – розширений модуль, що включає систему модулів

$$\begin{aligned} &P_1, P_2, \dots, P_i, \dots, P_k, P_{k+1}, P_{k+2}, \\ &0 < P_{k+1} \leq \alpha - 1 \\ &0 < P_{k+2} \leq \beta - 1, \quad i = 1, 2, \dots, \alpha; \quad j = 1, 2, \dots, \beta. \end{aligned}$$

Коефіцієнт стиснення даного методу розраховується за формулою:

$$K_{c1} = \frac{\sum_{i=1}^k \hat{E}[\log_2 b_{i\max}] + \hat{E}[\log_2 \alpha] + \hat{E}[\log_2 \beta]}{\hat{E}[\log_2 (\wp^* - 1)]}.$$

При кодуванні в базисі Галуа значення N_k доповнюється одним бітом Галуа і ДІ представляється як масив даних:

$$\{G_i, N_k\},$$

де G_i – рекурентна послідовність бітів Галуа, $0 \leq G_i \leq \alpha \cdot \beta$.

Опитування ДІ при кодуванні їх номерів бітами Галуа відбувається послідовно.

Коефіцієнт стиснення дорівнює:

$$K_{c2} = \frac{\sum_{i=1}^k \hat{E}[\log_2 b_{i\max}] + \hat{E}[\log_2 \alpha] + \hat{E}[\log_2 \beta]}{1 + \hat{E} \left[\log_2 \left(\prod_{i=1}^n P_i - 1 \right) \right]}.$$

Запропонований метод дає змогу перейти від багатовимірного представлення параметрів контрольованого об'єкта до одновимірного, що, в свою чергу, суттєво спрощує протоколи передавання даних, зменшує обсяг службових даних, захищає дані від несанкціонованого доступу (рис. 13).

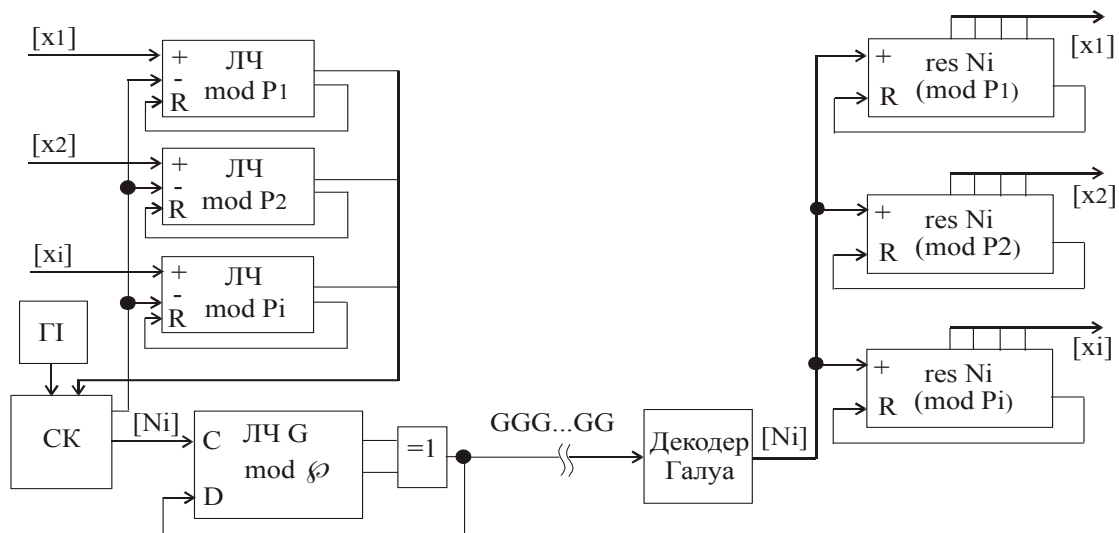


Рис.13. Структурна схема процесора обробки даних в базисі Крестенсона-Галуа:

$[X_i]$ – виміряні значення; ЛЧ $\text{mod } P_i$ – лічильник по модулю P_i ; ЛЧ G – лічильник Галуа; ГІ – генератор тактових імпульсів; СК – схема керування.

Використання рекурентних послідовностей Галуа дає змогу підвищити завадостійкість передавання вектору N_k .

У четвертому розділі розроблено апаратні та програмні засоби стиснення даних на основі запропонованих методів.

Наведено опис і функціональні схеми спецпроцесора обробки даних на основі перетворення Крестенсона – Галуа та процесора стиснення даних на основі базисних функцій Галуа.

Розроблено інтегрально-імпульсний перетворювач (ІП) з розширеними функціональними можливостями (рис. 14).

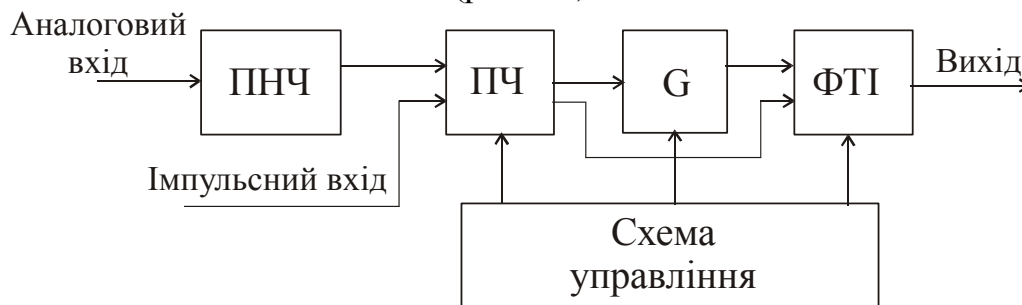


Рис. 14. Структурна схема ІП: ПНЧ – перетворювач напруга-частота; ПЧ – подільник частоти; G – генератор коду Галуа; ФТІ – формувач тривалості імпульсів.

Основні блоки ІП реалізовані на програмованій логічній інтегральній схемі фірми ALTERA EPM3064ALC44, що дало змогу підвищити надійність пристрою, суттєво прискорити час проектування і виготовлення, а отже, знизити вартість пристрою.

У додатках наведено інформаційні характеристики об'єктів низових комп'ютерних мереж, блок-схеми алгоритмів кодування-декодування

гармонічних сигналів, текстовий опис макрофункції G_GENERAT ІІП, блок-схема алгоритму стиснення даних на основі адаптивного кодування в базисі Галуа. Подано документи, що підтверджують впровадження результатів наукових досліджень за темою дисертації.

ВИСНОВКИ

У дисертаційній роботі розв'язано наукову задачу розробки та дослідження методів зменшення надлишковості цифрових даних у базисі Галуа, що забезпечують стиснення даних в процесі формування і передавання на низовому рівні РКС. При цьому отримано наступні результати.

1. Проаналізовано існуючі методи та інформаційні технології стиснення даних у комп'ютерних системах на основі представлення потоків повідомлень в різних теоретико-числових базисах і доведена перспектива використання базису Галуа для стиснення технологічних даних.

2. Розроблена архітектура системи ідентифікації станів ОК забезпечує на низовому рівні можливість зменшення обсягів даних на 1–2 порядки за рахунок кодування відповідного стану об'єкта та системи формальних параметрів, що описують стан об'єкта, замість інформації, що формується усіма сенсорами.

3. Проведений теоретичний аналіз методик оцінки коефіцієнта стиснення даних, на основі чого запропоновано визначення комплексного коефіцієнта ефективності методів стиснення.

4. Розроблений метод стиснення даних на основі запропонованих базисних функцій Галуа нульового і першого порядку, який забезпечує стабільний ефект зменшення обсягів даних процесів з різкими змінами амплітуди, для яких відомі методи неефективні, призводять до втрат інформації, збільшення обсягів даних стосовно первинних. Причому застосування базисних функцій Галуа першого порядку забезпечує рівномірність слідування бітів, що знижує навантаження і швидкісні вимоги до використання каналів зв'язку.

5. Запропонований і досліджений метод кодування гармонічних сигналів, що дає змогу закодувати період гармонічного сигналу по $1/8$ частині і однозначно його відновити, а отже, зменшити надлишковість представлення гармонічних сигналів у 7,6 раза.

6. Досліджено метод кодування багатовимірних ДД в базисі Крестенсона, що дає змогу перейти від багатовимірного до одновимірного представлення значень технологічних параметрів на передавальній стороні та зворотного перетворення на приймальній стороні. Запропоновано використання взаємопростих модулів розмірності $P = 2^n - 1$, що дало змогу передавати залишки послідовностями Галуа і реалізувати перетворення Крестенсона – Галуа. Проведені дослідження ефективності кодування номера ДД в базисах Радемахера, Крестенсона, Галуа

показали, що найменшу кількість біт для представлення номеру ДІ необхідно в базисі Галуа (1 біт) при послідовному опитуванні ДІ. Використання системи залишкових класів дає можливість захистити дані від несанкціонованого доступу. Одержані аналітичні залежності коефіцієнта стиснення розробленого методу.

7. Розроблено інтегрально-імпульсний перетворювач з розширеними функціональними можливостями, що дозволило розширити сферу його використання в РКС нафтогазової, енергетичної та інших галузей промисловості. Основні модулі ІІП реалізовано на ПЛІС, завдяки чому підвищена надійність і зменшено габарити пристрою.

8. На основі запропонованого методу стиснення гармонічних сигналів розроблено програмне забезпечення для використання у комп'ютерній системі реєстрації миттєвих значень струмів та напруг на первинному електричному обладнанні.

СПИСОК ПУБЛІКАЦІЙ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. 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. – Івано-Франківськ: ІФДТУНГ. – 2001. – Т. 6, № 37 – С. 183–186.

2. Яцків Н. Г., Николайчук Я. М. Нафтогазова автоматика та нові інформаційні технології – важливий важіль підвищення техногенної екологічної безпеки в паливно–енергетичному комплексі // Вимірювальна та обчислювальна техніка в технологічних процесах. – 2001. – № 4. – С. 173–176.

3. Николайчук Я. М., Яцків Н. Г. Методи стиснення даних в багатоканальних системах на основі кодів Галуа // Вісник Національного університету “Львівська політехніка”. Радіоелектроніка та телекомунікації. – Львів. – 2002. – № 443.– С. 135–138.

4. Яцків Н. Г., Николайчук Я. М. Системні характеристики джерел інформації та оцінка їх ентропії // Вісник Технологічного університету Поділля. Технічні науки. – Хмельницький. – 2002. – Т. 1, № 3. – С. 197–200.

5. Яцків Н. Г. Дослідження системних характеристик методів формування даних в різних теоретико-числових базисах // Вимірювальна та обчислювальна техніка в технологічних процесах: Збірник наукових праць. – Хмельницький: ТУП. – 2002.– Т. 2, № 9. – С. 132–136.

6. Яцків Н. Г., Король Р. І., Яцків В. В., Федчишин Т. Г. Спецпроцесор обробки даних на основі перетворення Крестенсона–Галуа // Вісник Технологічного університету Поділля. - Хмельницький. – 2003. – Т. 1, № 3. – С. 105 – 108.

7. Яцків Н. Г., Яцків В. В. Методи кодування та фізичного представлення

сигналів в інформаційно-керуючих системах // Розвідка і розробка нафтових і газових родовищ. Серія: Методи і засоби технічної діагностики. – Івано-Франківськ: ІФДТУНГ. – 2001. – Т.8, № 38. – С. 140–144.

8. Yatskiv N, Nykolaiychuk Y. Method of data compression in multichannel systems on the basis of Galois codes. //Proc. of the International Conf. TCSET 2002. – Lviv-Slavsko (Ukraine). – 2002. – P. 135.

9. Nykolaychuk Y., Yatskiv N., The coding of multichannel sources of information //Proc. of the International Conf. The Experience of Designing and Application of CAD Systems in Microelectronics. CADSM 2003. – Lviv-Slavsko (Ukraine). 2003. – P. 249 – 250.

АНОТАЦІЯ

Яцків Н. Г. Методи та засоби стиснення даних в розподілених комп'ютерних системах на основі кодів поля Галуа. – Рукопис.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.13 – обчислювальні машини, системи та мережі. – Національний університет “Львівська політехніка”, Львів, 2003.

Дисертація присвячена питанням розробки та дослідження методів зменшення надлишковості цифрових даних, створенню апаратних і програмних засобів стиснення даних у базисі Галуа. У дисертації запропоновано метод стиснення даних на основі базисних функцій Галуа нульового та першого порядку, що формує потік бітів Галуа з рівномірною швидкістю передавання.

Розроблено метод кодування багатомірних ДІ на основі системи залишкових класів, що дає змогу перейти від багатовимірного до одновимірного представлення значень технологічних параметрів на передавальному боці та зворотного перетворення на приймальному боці. Використання взаємопростих модулів розмірності $P = 2^n - 1$ дало можливість передавати залишки послідовностями Галуа.

На основі запропонованого методу стиснення гармонічних сигналів розроблено програмне забезпечення для використання у системі реєстрації миттєвих значень електричних параметрів на первинному електричному обладнанні.

Ключові слова: розподілені комп'ютерні системи, стиснення даних, кодування, базис Галуа, базисні функції, низові комп'ютерні мережі, теоретико-числові перетворення.

АННОТАЦИЯ

Яцкив Н. Г. Методы и средства сжатия данных в распределенных компьютерных системах на основе кодов поля Галуа. – Рукопись.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.13 – вычислительные машины, системы и сети. – Национальный университет “Львовская политехника”, Львов, 2003.

Диссертация посвящена вопросам разработки и исследования методов уменьшения избыточности цифровых данных, созданию аппаратных и программных средств сжатия данных в базисе Галуа.

Проведенный анализ и исследование технологических установок, как источников информации (ИИ), показали, что большинство их принадлежат к классу квазистационарных объектов управления (ОУ). Разработана архитектура системы идентификации семантических, технологических и информационных состояний ОУ.

В диссертации проведена обобщенная классификация известных и предложенных методов уменьшения избыточности данных по предметной области: универсальные и специальные. Исследование теоретико – числовых базисов Галуа, Уолша, Хаара, Крестенсона с целью определения минимальных затрат машинного времени и объемов памяти для обработки и сохранения данных показало преимущества использования дискретного базиса Галуа.

Исследован метод кодирования и сжатия данных на основе системы остаточных классов, коэффициент сжатия которого зависит от выбранного модуля и разрядности данных.

Теоретически обоснованы потенциальные возможности кодирования цифровых данных в базисе Галуа на основе вертикальной информационной технологии. Разработан и исследован метод кодирования и сжатия данных на основе базисных функций Галуа нулевого и первого порядка.

Предложен и исследован метод кодирования гармонических сигналов, который разрешает закодировать период гармонического сигнала по $1/8$ части и однозначно его восстановить, и, таким образом, уменьшить избыточность представления гармонических сигналов в 7,6 раза.

Разработаны адаптивные методы сжатия данных одномерных и многомерных ИИ на основе кодов поля Галуа. Коэффициент сжатия равен $k_c = 1,5 - 4$ и зависит от количества активных отсчетов. За счет рекуррентных свойств битов Галуа формируется поток данных с блочной синхронизацией, дающей возможность повысить помехоустойчивость закодированных данных.

Предложен метод кодирования многомерных ИИ на основе системы остаточных классов, который позволяет перейти от многомерного к одномерному представлению значений технологических параметров на передающей стороне и к обратному преобразованию на приемной стороне. Проведенные исследования эффективности кодирования номера ИИ в базисах Радемахера, Крестенсона, Галуа показали, что наименьшее количество бит для представления номера ИИ

необходимо в базисе Галуа (1 бит) при последовательном опросе ИИ. Использование системы остаточных классов разрешает защитить данные от несанкционированного доступа.

Разработан интегрально-импульсный преобразователь (ИИП) с расширенными функциональными возможностями, основной модуль ИИП реализован на ПЛИС, что повысило надежность разработанного устройства. Изготовлен экспериментальный образец ИИП с расширенными функциональными возможностями.

На основе предложенного метода сжатия гармонических сигналов разработано программное обеспечение для использования в компьютерной системе регистрации мгновенных значений электрических параметров на первичном электрическом оборудовании.

Ключевые слова: распределенные компьютерные системы, сжатие данных, кодирование, базис Галуа, базисные функции, низовые компьютерные сети, теоретико-числовые преобразования.

ABSTRACT

Yatskiv N. G. Methods and means of data compression on the basis of Galois field codes in the distributed computer systems. – Manuscript.

The dissertation on competition of the scientific degree of the candidate of engineering science on speciality 05.13.13 – computers, systems and networks. National university “Lvivska politehnika”, Lviv, 2003.

The dissertation is dedicated to questions of development of coding methods and reduction of redundancy of data, creation hardware and software of data compression in Galois basis. In the dissertation the compression method on the basis of Galois basic functions of the zeroth and first order, which forms a flow Galois bits with uniform data rate and provides compression ratio 4 – 10 is offered.

The coding method of multidimensional sources of information is developed on the basis of system of residual classes, which permits to transform from multidimensional to one-dimensional performance of values of technological parameters on the transmitting side and return transformation on the receiving side, use of mutually simple modules of dimension has permitted to transfer the rests by Galois sequences. On the basis of the offered compression method of harmonious signals the software, for use in system of registration of instant values of electrical parameters on primary electrical equipment is developed.

Keywords: distributed computer systems, data compression, coding, Galois basis, basic functions, low-level computer systems, theoretical-number transformations.

Підписано до друку 8.08.2003 р.
Формат 84x108/32. Папір офсетний. Друк на дублікаторі.
Умов. – друк. арк. 0,95. Обл. – вид. арк.0,90. Зам. № 1 – 8/08/03.
Тираж 100 прим.

Віддруковано у видавництві “Економічна думка”
Тернопільської академії народного господарства
46000, Тернопіль, вул. Львівська, 3, тел. 43-22-18