

МЕТОДИ ТА ЗАСОБИ ЗАБЕЗПЕЧЕННЯ ОПЕРАТИВНОГО ВІДНОВЛЕННЯ ТА ДОСТУПНОСТІ ДАНИХ В АВТОМАТИЗОВАНИХ ІНФОРМАЦІЙНИХ СИСТЕМАХ

Гончар Л.І.¹⁾, Шурдак Р.В.²⁾, Ваврух Р.В.³⁾, Пастернак Ю.І.⁴⁾, Олійник О.О.⁵⁾

*Західноукраїнський національний університет
1)к.е.н., доцент; 2,5)магістрант; 3)аспірант; 4)студент*

І. Постановка проблеми

В даний час забезпечення оперативного відновлення і доступності даних в автоматизованих інформаційних системах різного роду базується на вимогах міжнародних стандартів, за процедурами забезпечення доступності комп'ютерного сервісу і методам [1,2], в основі яких лежить створення плану забезпечення безперервності діяльності організації і плану відновлення в разі надзвичайних ситуацій, а також технологій забезпечення доступності та відновлення інформації.

Аналіз показує, що відомі комп'ютерні технології [2] допускають втрати інформації в таких обсягах, в результаті яких прості відновлення обчислювальної та мережевої інфраструктури втрачає сенс. Особливою складністю для зазначених вище процедур є зростаючий обсяг інформації, яка підлягає резервуванню. Відмінною особливістю сучасних комп'ютерних технологій, застосовуваних в компаніях, є також, орієнтація на створення єдиного інформаційного простору, в результаті чого виникли проблеми несумісності форматів представлення інформації.

II. Мета роботи

Метою дослідження є розробка методів та програмних засобів для забезпечення оперативного відновлення та доступності даних в автоматизованих інформаційних системах.

III. Метод стиснення кодів на основі двійково-десятькового перетворення

Основою запропонованого методу є наступні основні положення:

- не використовує модель надмірності інформації;
- не залежить від обсягу і формату представлення інформації, так як використовується альтернативний математичний апарат двійковий-десятькового перетворення.
- дозволяє стискати як об'єкти файлової системи, так і дампи пам'яті комп'ютера або їх будь-яких груп в мережі, центрів обробки даних в цілому, завдяки використанню прямої адресації початку і кінця бінарного коду, що стискається.
- дозволяє передачу великих обсягів інформації, використовуючи низькошвидкісні канали зв'язку.

У даній роботі запропонована альтернативна формула двійково-десятькового перетворення, заснована на заміні степеневої функції на лінійну:

$$N = \sum_{i=1}^m a_n \times n, \quad (1)$$

де N – невідоме число для заданого двійкового коду, m - довжина заданого коду в бітах, n - довжина коду в бітах, $a \in \{0,1\}$.

При використанні цієї формули для першого біта підсумовується одиниця, для другого - двійка і так далі, тобто номер біта позначає цифру, яка повинна бути приплюсована. Формулою розрахунку десяткового числа в даному випадку буде сума чисел від 1 до n . Для всього обсягу пам'яті, заповненого одиницями:

$$Z = \frac{n(n+1)}{2}, \quad (2)$$

де n – довжина коду в бітах.

При цьому для довгих двійкових кодів десяткове число буде на порядки коротше, ніж при класичному двійковий-десятьковому перетворенні. Строга математична постановка задачі стиснення для даного методу виглядає наступним чином: обчислити невідоме N для заданого двійкового коду, знаючи тільки M (суму на основі лінійної функції) і m (де m – довжина заданого коду в бітах). При цьому M в ході компресії обчислено для заданого коду по нижній формулі в системі рівнянь (3). Для

декомпресії потрібно розв'язати систему рівнянь при заданих вище значеннях. Причому, на кожному кроці підсумовування значення a_n буде однаковим в обох рівняннях системи.

$$\begin{cases} N = \sum_{n=1}^m a_n \times 2^n \\ M = \sum_{n=1}^m a_n \times n \end{cases} \quad (3)$$

Розглянемо приклад: задано двійковий код 10011011 (8 біт), знаходимо $M = 1+4+5+7+8 = 25$. Тоді для декомпресії потрібно розв'язати систему рівнянь (4).

$$\begin{cases} N = \sum_{n=1}^8 a_n \times 2^n \\ 25 = \sum_{n=1}^8 a_n \times n \end{cases} \quad (4)$$

В результаті розрахунку N при заданих значеннях має дорівнювати 217. Після чого, можна було б записати результат у вигляді двійкового коду. Але в зв'язку з тим, що можливі різні множини з a_n призводять до $M = 25$ (виникнення колізій), отримуємо множину розв'язків N , замість єдиного.

Висновок: потрібні додаткові члени системи рівнянь для реалізації єдиного розв'язку, або зниження кількості можливих рішень до прийнятної величини (для вказівки номера розв'язку).

IV. Програмна реалізація методу стиснення кодів на основі двійково-десятькового перетворення з використанням лінійних функцій

В даній роботі пропонується програмне забезпечення для стиснення інформації будь-яких форматів представлення без втрат, яке засноване на новому математичному методі двійково-десятькового перетворення інформації і псевдорегулярних числах, які задають постійну внутрішню структуру двійкових полів пам'яті комп'ютера і окремих інформаційних об'єктів.

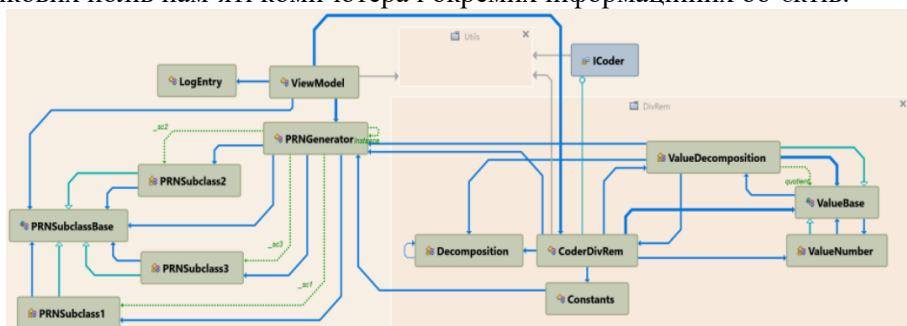


Рисунок 1 – Діаграма типів розробленої системи

На рисунку 1 представлено діаграму типів, а на рисунку 2 представлено діаграму класів, які використовуються для формування структур зберігання даних. Було проведено комп'ютерне моделювання за алгоритмом, який реалізований на запропонованих підходах.

Висновок

На основі розроблених методів реалізовано програмне забезпечення для стиснення та оперативного відновлення даних. Ефективність розроблених методів підтверджено експериментально.

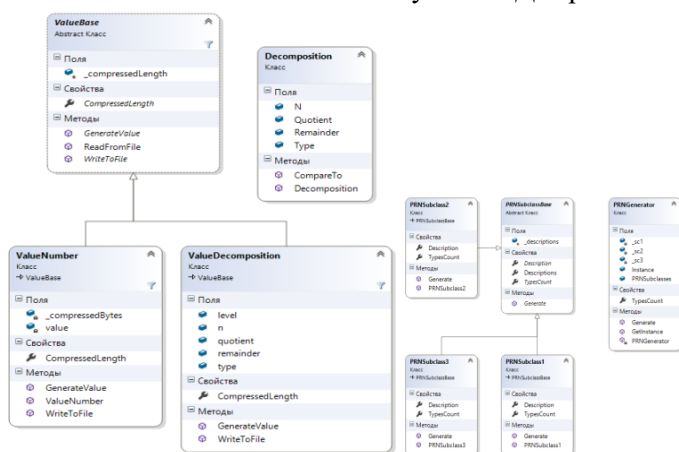


Рисунок 2 – Діаграма класів, що представляють структури зберігання даних

Список використаних джерел

1. Bentley, J. L. et al. (1986) "A Locally Adaptive Data Compression Algorithm," Communications of the ACM, 29(4):320–330.
2. Ryzhikov Y.I. Calculation of the Closed Multichannel Queueing Systems // Cybernetics and Algorithms in Intelligent Systems. Proceedings of 7th Computer Science Online Conference. 2018. Vol. 3. Pp.125–132.