

ВЕКТОРНИЙ МЕТОД ЛОКАЛІЗАЦІЇ ПОМИЛОК У ПОТОКАХ ДАНИХ

Вінницький національний технічний університет

Анотація

Розглянуто проблеми, що виникають під час передавання і зберігання інформації, а саме пошкодження даних під впливом зовнішніх завад. Обґрунтовано актуальність даної теми. Запропоновано метод побудови дерева згортки при векторному кодуванні, що дозволяє локалізувати подвійні помилки та відновити пошкоджені ділянки даних, за допомогою одного контрольного розряду на слово даних. Запропоноване дерево згортки надає можливість простого розширення розрядності даних. Таким чином, описаний метод достовірного передавання і зберігання інформації має мінімальну надлишковість даних та за рахунок регулярної будови дерева – підвищену ефективність.

Ключові слова: контроль, векторні коди, метод фруктового саду, надлишковість, передача інформації.

Abstract

The problems that arise during the transmission and storage of information is considered, namely data corruption under the influence of external interference. The relevance of this topic is substantiated. The method of constructing a convolution tree with vector coding is proposed, which allows to localize double errors and recover damaged parts of the data, using one control bit per data word. The proposed convolution tree allows you to easily extend the bit size of the data. Thus, the described method of reliable transmission and storage of information has a minimal redundancy of data and due to the regular structure of the tree - increased efficiency.

Key words: control, vector codes, orchard method, redundancy, information transfer.

Розширення інформаційної галузі потребує підвищення достовірності у потоках даних. Всі існуючі підходи до вирішення вказаних задач базуються на введенні у дані надлишковості для зберігання інформації з метою подальшого аналізу появи помилок. Відомо ряд публікацій, що стосуються використання надлишковості для підвищення ефективності обробки інформації [1-11]. До важливих критеріїв вирішення проблем достовірного передавання і зберігання великих обсягів даних, відносять високу ефективність виявлення помилок при незначній надлишковості. Розробка методу, що задовільнить ці критерії є актуальною задачею.

Одним з ефективних підходів до підвищення достовірності є методи локалізації помилок в двійкових даних та їх подальшого виправлення за допомогою векторних кодів [12-14]. Такі методи передбачають додавання одного контрольного біта до кожного двійкового слова і використання певних розрядів декількох сусідніх слів даних для його обчислення. Відомі публікації, де вказано, що використання векторного методу дозволяє виявляти та виправляти помилки в двійкових даних. Проте, даний підхід потребує подальшого дослідження та обґрунтування. У даній статті авторами запропоновано метод векторного кодування, який дозволяє локалізувати одинарні і подвійні помилки при передаванні і зберіганні великих обсягів інформації за допомогою лише одного контрольного розряду. Запропонований метод використовує вектори з регулярною будовою, що надає можливість простого нарощування розрядності кодових слів.

Найбільш простими способами побудови кодів, локалізуючих і коригуючих помилки передачі або зберігання двійкової інформації, є коди, що базуються на виявленні кратності конкретних бітів даних. Скоттом Е. та Гетшелем Д. [14] запропоновано векторний код, що дає можливість локалізувати і виправити помилки за допомогою лише одного контрольного розряду на слово. Метод дозволяє ефективно виявляти і виправляти помилку теоретично будь-якої кратності в межах блоку певної кількості слів. Як стверджують автори, така модель помилок найбільш вірогідна при передаванні, записуванні і зчитуванні потокової інформації. Для контролю вони запропонували так званий, метод фруктового саду, концепція якого базується на візуальному враженні, що складається в людини, при проходженні яблуневого саду з регулярно розміщеними рядами дерев. На рисунку 1 надана схема локалізації одинарної помилки за допомогою даного методу при використанні двох векторів згортки за модулем 2.

Всі вектори згортки в один контрольний розряд назвемо деревом згортки. З рисунку видно, що одинарна помилка призведе до порушення парності у двох контрольних розрядах K1 і K2, які назвемо контрольним слідом.

	1			1		2			2	
	1			1	2				2	
	1			■				2		
	1		2	1		2				
	1	2		1	2					
	K1			K2						

Рисунок 1 – Схема локалізації помилки дво-векторним кодом

Відстань між бітами контрольного сліду дозволяє однозначно встановити місце знаходження помилки. Отже, дво-векторний код дозволяє локалізувати і виправити одинарну помилку в межах блоку слів, кількість яких визначається довжиною дерева згортки.

У роботі [13] запропоновано використання три-векторного дерева згортки для локалізації одинарних та виявлення багатократних помилок, як це представлено на рисунку 2.

	2				1					3	
		2			1					3	
			2		1					3	
				2	1				3		
					2	1	3				
						K					

Рисунок 2 – Дерево згортки для симетричного три-векторного коду

На відміну від дерева згортки на рисунку 1, дана версія дозволяє локалізувати більше одинарних помилок. Таке дерево дозволяє легко визначити співвідношення для обчислення контрольних розрядів. Помилковий біт локалізується за допомогою двох розташованих підряд контрольних розрядів, для яких при декодуванні виявлено порушення парності.

Однак, у випадку виникнення подвійної помилки у даному блоці можна навести приклади, коли один і той самий контрольний біт є елементом двох контрольних слідів. Подвійне порушення парності такого розряду не буде поміченим.

Метод векторного кодування для локалізації подвійних помилок з можливістю розширення розрядності

Для виправлення подвійних помилок в словах з розрядністю п'ять, (включаючи контрольний) у [14] запропонована форма дерева згортки, зображена на рисунку 3.

№ розряду	№ слова																
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17
4	1					2							3				
3	1				2									3			
2	1			2											3		
1	1		2									3					
0	K	2															3

Рисунок 3 – Дерево згортки три-векторного коду

Така можливість базується на тому, що на кожному розрядному зрізі у згортці беруть участь три розряди, по одному з кожного вектора. Причому, відстані першим та другим, другим та третім, першим та третім розрядами є унікальними для даного зрізу і на других розрядних зрізах в дереві більше не повторюються. Тому навіть якщо відбудеться накладання біта контрольного сліду одного дерева та іншого, що приведе до неможливості визначення двох відстаней, то все одно залишиться третя унікальна відстань між бітами контрольного сліду, яка дозволить однозначно визначити помилковий розряд. Таким чином, при використанні

векторного коду для локалізації n-кратних помилок потрібно будувати дерево згортки, яке має (n+1) векторів, причому повинні бути унікальними всі можливі відстані між векторами на кожному розрядному зрізі.

Недоліком відомого дерева згортки [14], представленого на рисунку 3, є відсутність простої і зрозумілої закономірності побудови третього вектора для різних розрядностей. Це ускладнює нарощування розрядності слів при використанні векторного кодування для трьох і більшої кількості векторів.

Розглянемо умови яким повинно відповідати дерево згортки для локалізації і виправлення двох помилок. При наявності трьох векторів в формуванні одного контрольного розряду беруть участь три інформаційних розряди однієї ваги з різних слів. Контрольні розряди, в формуванні яких бере участь інформаційний розряд a_i^j назовемо контрольним слідом інформаційного розряду.

Нехай $\{a_f^k, a_g^k, a_h^k\} \subset D_i$, тоді $\{a_{f+b}^k, a_{g+b}^k, a_{h+b}^k\} \subset D_{i+b}$. Знайдемо контрольний слід розряду a_b^k . Нехай $b = f + x = g + y = h + z$, тоді правдивим є вираз (10)

$$\begin{aligned} \{a_b^k, a_{g+x}^k, a_{h+x}^k\} &\subset D_{i+x}; \\ \{a_{f+y}^k, a_b^k, a_{h+y}^k\} &\subset D_{i+y}; \\ \{a_{f+z}^k, a_{g+z}^k, a_b^k\} &\subset D_{i+z}. \end{aligned} \quad (10)$$

Визначимо умову накладання контрольних слідів від двох розрядів в двох точках.

Нехай $\{a_k^i, a_p^i, a_e^j, a_m^j\} \subset D_r$, нехай далі $k > p > b > m$ і $k - p = b - m = h$, тобто в дереві на рівнях і та j існують дві пари рівновіддалених вершин. Тоді для довільного a_d^i ($d = k + x = p + y$) отримаємо контрольний слід (11)

$$\begin{aligned} \{a_d^i, a_{p+x}^i\} &\subset D_{r+x}; \\ \{a_{k+y}^i, a_d^i\} &\subset D_{r+y}, \end{aligned} \quad (11)$$

при чому $k = d - x$, $p = d - y$ і, отже, $k - p = y - x = h$, тобто контрольний слід a_d^i включає два контрольних розряди, що знаходяться на відстані h один від одного.

Аналогічно для довільного a_c^j ($c = b + z = m + \tau$) отримаємо контрольний слід (12)

$$\begin{aligned} \{a_c^j, a_{m+z}^j\} &\subset D_{r+z}; \\ \{a_{b+\tau}^j, a_c^j\} &\subset D_{r+\tau}; \end{aligned} \quad (12)$$

при чому $b = c - z$, $m = c - \tau$ і, отже, $b - m = \tau - z = h$, тобто контрольний слід a_d^i включає два контрольних розряди, які знаходяться на відстані h один від одного. В тому випадку, коли $\tau = y$, то $z = \tau - h = x$, тобто два контрольних розряди контрольних слідів a_d^i і a_c^j співпадають. Отже, умова $k - p = b - m$ є достатньою, для того щоб виникало співпадіння двох різних контрольних слідів. Необхідність даної умови доводиться від зворотного. Дійсно, нехай контрольні сліди розрядів a_b^k, a_c^p мають спільні контрольні розряди d_j і d_{j+t} . Це означає, що як a_b^k так і a_c^p є елементами дерев згортки з вершинами в слові j в слові j+t (13).

$$\begin{aligned} \{a_b^k, a_c^p\} &\subset D_j; \\ \{a_b^k, a_c^p\} &\subset D_{j+t}. \end{aligned} \quad (13)$$

Оскільки формула згортки для всіх розрядів одна, то із (13) слідує (14)

$$\{a_{b+t}^k, a_{c+t}^p, a_b^k, a_c^p\} \subset D_{j+t}; \quad (14)$$

тобто дерево містить два розряди на рівні k з відстанню між ними t і два розряди на рівні p з відстанню між ними t.

Таким чином, наявність двох пар рівновіддалених вершин в дереві є необхідною і достатньою умовою співпадіння контрольних слідів різних розрядів в двох точках. Отже, для того щоб не було накладання контрольних слідів в двох точках, дерева згортки не повинні включати в себе рівновіддалені розряди однієї ваги. Оскільки, при три-векторному дереві контрольний слід складається з трьох контрольних розрядів, то при подвійній помилці накладання контрольних слідів більше ніж в двох точках неможливе.

У три-векторному дереві на кожному рівні знаходиться по три розряди a_i, a_l, a_k , де i, l, k – номери слів такі, що $i < l < k$. Нехай $l - i = x$, $k - l = y$, $k - i = z$. Для j-го рівня повинні виконуватись умови

$$\begin{aligned} x_j &\neq y_j; \\ x_j &\neq z_j; \end{aligned} \quad (16)$$

$$y_i \neq z_j.$$

Кількість слів, що беруть участь в одному дереві згортки позначимо як R . При чому, $z_{\max}=R$. Для зменшення R потрібно щоб множини $\{x\}$, $\{y\}$, $\{z\}$ містили натуральні числа, що заповнюють без проміжків відрізок числової вісі від одиниці до R . Надамо більш строгі формулювання умови для дерева згортки з мінімальним R .

Для будь-якого n визначити $\{x_1, \dots, x_n\}$, $\{y_1, \dots, y_n\}$, $\{z_1, \dots, z_n\}$ такі, що x, y, z – натуральні; $\{x\} \cap \{y\} = \emptyset$, $\{y\} \cap \{z\} = \emptyset$, $\{x\} \cap \{z\} = \emptyset$, причому, $\{\{x\} \cap \{y\} \cap \{z\}\} = \{1, 2, 3, \dots, 3n+1\}$.

У даній роботі авторами пропонується конфігурація дерева згортки, яка при всіх позитивних властивостях відомого рішення для будь-якої розрядності надає просту і зрозумілу закономірність побудови дерева з трьох векторів, представлену виразом для визначення контрольного розряду $K_i = a_i^0$ (17).

$$V1_i = \left(\sum_{j=1}^{n-1} (a_i^j) \right)_{\text{mod}2} ;$$

$$V2_i = \left(\sum_{j=0}^{n-1} (a_{i+j+1}^j) \right)_{\text{mod}2} ;$$

$$V3_i = \left(\sum_{j=0}^{(n-1)/2} (a_{i+2n+2-j}^{n-1-2j}) \right)_{\text{mod}2} \oplus \left(\sum_{j=0}^{(n-3)/2} (a_{i+3n+2-j}^{n-2-2j}) \right)_{\text{mod}2} ;$$

$$a_i^0 = V1_i \oplus V2_i \oplus V3_i.$$

На рисунку 4 зображено приклад запропонованого авторами дерева згортки для виправлення подвійних помилок, де K – контрольний розряд.

В запропонованому варіанті дерева згортки прослідковується проста закономірність його будови, що дозволить застосовувати даний метод для контролю слів даних будь-якої розрядності, а також при необхідності нарощувати розрядність пристрою контролю.

№ розряду	№ слова																
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17
4	1					2						3					
3	1				2												3
2	1			2							3						
1	1		2														3
0	K	2								3							

Рисунок 4 – Запропонований варіант дерева згортки для три-векторного коду,

Висновки

У доповіді подано теоретичні аспекти модифікованого алгоритму векторного коду. Описано існуючі методи векторного кодування даних. Наведено аналітичні вирази для опису побудови дерева згортки три-векторного коду, яке дозволяє виправляти подвійні помилки за допомогою одного контрольного розряду на слово. На основі проведеного аналізу запропоновано власний метод, який дозволяє будувати векторні коди з можливістю розширення розрядності даних за рахунок регулярної будови кодового дерева згортки, що і є підвищенням ефективності. З рисунків 2 і 4 видно, що довжина блоку кодів даних, в яких розміщено дерево згортки у запропонованому варіанті така ж, як у відомому. Тобто, у даному випадку авторам вдалося покращити одні параметри рішення не погіршуючи інших.

Список літератури

- [1] А. Д. Азаров, А. И. Черняк, "Полнофункциональная побитовая обработка результатов аналого-цифрового преобразования," на III міжнародній наук.-практ. конф. Методи та засоби кодування, захисту й ущільнення інформації, Вінниця, с. 208-209. 2011.

- [2] Olexiy D. Azarov, Olexander G. Murashchenko, Olexander I. Chernyak, Andrzej Smolarz, Gulzhan Kashaganova, "Method of glitch reduction in DAC with weight redundancy," in *16th Conference on Optical Fibers and Their Applications*, Proc. SPIE 9816, 98161T, Lublin and Naleczow, Poland, 2015; doi: 10.1117/12.2229045; <http://dx.doi.org/10.1117/12.2229045>.
- [3] О. Д. Азаров, О. І. Черняк, О. Г. Муращенко, "Метод побудови швидкодіючих фібоначчєвих лічильників," *Проблеми інформатизації та управління*, №2(46), с. 5-8. 2014.
- [4] О. Д. Азаров, О. І. Черняк, "Визначення довжини перенесення при додаванні в системах числення з адитивними та мультиплікативними співвідношеннями між вагами розрядів," *Наукові праці Донецького національного технічного університету. Серія: Обчислювальна техніка та автоматизація*, № 74, с. 401–407. 2004. ISSN 1996-1588.
- [5] О. Д. Азаров, О. І. Черняк, "Структурна організація побітового множення і ділення кодів золотої пропорції," *Проблеми інформатизації та управління*, №3(21). 2007. с. 5–13. ISSN 2073-4751.
- [6] О. Д. Азаров, О. І. Черняк, "Розрядність пристроїв порозрядного додавання в АМ-системах числення," *Наукові праці Вінницького національного технічного університету* № 4, с. 1-9, 2010. [Електронний ресурс]. Режим доступу : <http://praci.vntu.edu.ua/index.php/praci/article/view/233>. Дата звертання: Лис. 2020.
- [7] О. Д. Азаров, О. І. Черняк, "Структурна організація побітового додавання і віднімання кодів золотої 1-пропорції з урахуванням знаків," *Інформаційні технології та комп'ютерна інженерія*, № 3(22), с. 13–16. 2011. ISSN 1999-9941.
- [8] О. Д. Азаров, О. І. Черняк, "Аналіз витрат обладнання пристроїв побітової арифметики у системі числення золотої 1-пропорції," *Проблеми інформатизації та управління*, № 2(38), с. 5-9. 2012. ISSN 2073-4751.
- [9] О. Д. Азаров, О. І. Черняк, *Повнофункціональна побітова потокова арифметика зі зменшеними витратами обладнання: монографія*. Вінниця, Україна : ВНТУ, 2013.
- [10] О. Д. Азаров, О. І. Черняк, "Обмеження адитивних співвідношень при порозрядній потоковій обробці в АМ- системах числення," *Інформаційні технології та комп'ютерна інженерія*, № 3(31), с. 67-71. 2014. ISSN 1999-9941.
- [11] О. Д. Азаров, О. І. Черняк, О. Г. Муращенко, "Порозрядне додавання в АМ-системах числення на основі адитивних перетворень," *Проблеми інформатизації та управління*, № 1(45), с. 14-21. 2014. ISSN 2073-4751.
- [12] А.И. Черняк, А.П. Стахов, В.П. Марценюк, В.И. Пилипчак, О.А. Пленсак, "Устройство кодирования по векторному методу," *МКИ4 G 06 F 11/10 №1451700 A1*, 15.01.89.
- [13] О.П. Шафеева "Векторные коды для локализации ошибок в двоичных данных," *Омский научный вестник*, №3 (32). 2005.
- [14] Скотт Э., Гетшель Д. "Исправление многобитовых ошибок при помощи одного контрольного бита на слово," *Электроника*, № 9, с. 40-47. 1981.