

## **АНАЛІЗ МЕТОДУ РОЗПОДІЛЕНОГО ВИХІДНОГО КОДУВАННЯ СТИСНЕННЯ ДАНИХ В БЕЗПРОВОДОВИХ СЕНСОРНИХ МЕРЕЖАХ**

**<sup>1</sup>Третяк А.В., <sup>1</sup>Лисенко О.І., <sup>2</sup>Алексеева І.В., <sup>1</sup>Новіков В.І.**

*<sup>1</sup>Навчально-науковий Інститут телекомунікаційних  
систем КПІ ім. Ігоря Сікорського, Україна*

*<sup>2</sup>Фізико-математичний факультет КПІ ім. Ігоря Сікорського, Україна  
E-mail: novikov1967@ukr.net*

### **ANALYSIS OF THE DISTRIBUTED SOURCE CODING USING SYNDROMES METHOD FOR DATA COMPRESSION IN WIRELESS SENSOR NETWORKS**

The features of the application of the distributed source coding method for data compression in wireless sensor networks are analyzed.

Для зменшення витрати енергії, пов'язані з передачею даних та підвищення ефективності використання каналів зв'язку з обмеженою смугою пропускання в безпроводових сенсорних мережах (БСМ) застосовуються методи стиснення даних [1-2]. Розподілене кодування — це метод обробки даних, який дозволяє стискати інформацію з кількох джерел (наприклад, сенсорів), які генерують корельовані дані, без потреби прямого обміну інформацією між джерелами.

Кожне джерело стискає свої дані незалежно, передаючи лише мінімально необхідну інформацію. Декодування виконується в центральному вузлі, який використовує як отримані стисливі дані, так і "sideinformation" (допоміжну інформацію, яка може бути корельованими даними від інших джерел). Розподілене кодування вирішує проблеми, які виникають у системах із розподіленими джерелами даних, котрі зустрічаються у БСМ. У безпроводних сенсорних мережах кожен сенсор має обмежений запас енергії. Найбільші витрати енергії пов'язані з передачею даних. Замість передавання повного набору даних, розподілене кодування дозволяє сенсорам передавати тільки стислі дані (синдроми), що значно знижує витрати енергії. У системах із високою кількістю джерел (наприклад, сотні чи тисячі сенсорів) передача великих обсягів даних створює перевантаження каналу зв'язку. Розподілене кодування скорочує кількість переданих даних. Дані, які генеруються джерелами, що знаходяться у близькості (наприклад, сенсори температури), часто корельовані. Замість передачі всіх даних, розподілене кодування використовує цю кореляцію для стиснення.

Метод розподіленого вихідного кодування (англ. Distributed source coding using syndromes, DISCUS) [3] базується на теоремі Слєпіана-Вулфа, котра теоретично обґрунтувала можливість стиснення корельованих джерел без втрати ефективності, навіть якщо кодування виконується незалежно.

Розподілене кодування підходить для великих систем, де централізовані підходи можуть стати вузьким місцем через обмеження обчислювальної потужності чи пропускну здатності.

Висока кореляція між даними означає, що значення, зібрані різними сенсорами, мають схожі або залежні одне від одного значення. Це часто спостерігається у фізичних системах, де сусідні сенсори вимірюють параметри в одній і тій самій області (наприклад, температуру, вологість, тиск). Тобто якщо сенсор S1 вимірює температуру 25°C, а сенсор S2, розташований поруч, вимірює 24°C або 26°C, то дані S1 та S2 корельовані. У таких випадках ми можемо використати знання про значення одного сенсора (або групи сенсорів) для передбачення або точного відновлення значення іншого. Кореляція дозволяє зменшити надлишковість у даних. Замість передачі повного набору даних, можна передавати тільки відмінності або мінімальну кількість інформації (синдроми), достатню для відновлення даних у центральному вузлі.

Центральний вузол мережі використовує кореляцію наступним чином:

1. Центральний вузол знає дані одного або кількох сенсорів (Y) як допоміжну інформацію. Це може бути або повні дані сусіднього сенсора, або вже відновлені дані.

2. Синдром, отриманий від джерела, визначає набір можливих значень (косет).

3. Кореляція з допоміжною інформацією дозволяє звужити цей набір до одного правильного значення.

4. Вузол виконує операцію декодування, використовуючи матрицю H, яка визначає, як саме стискати й декодувати синдроми.

5. Знаходиться значення, яке відповідає отриманому синдрому і найближче до допоміжної інформації.

Параметри котрі повинні бути встановлені заздалегідь на вузлі наступні:

1. Матриця H, яка визначає, як стискаються дані і як декодуються синдроми. Центральний вузол повинен мати ту саму матрицю H, яка використовується у сенсорах.

2. Центральний вузол повинен знати, яким чином дані корельовані тобто знати модель кореляції (наприклад, максимальна різниця значень між сенсорами).

3. Центральний вузол повинен отримувати або вже мати дані від сусідніх сенсорів, щоб використовувати їх як допоміжну інформацію для відновлення стислих даних.

4. Вузол повинен мати алгоритм для обчислення можливих значень із синдрому (кандидатів) і вибору правильного значення на основі кореляції.

Розглянемо наступний приклад стискання бітів інформації в DISCUS

Початковий обсяг даних: X має розмір 1024 байти (тобто  $1024 \times 8 = 8192$  біт).



Кожен сенсор кодує свої дані у вигляді синдрому.

Центральний вузол вже знає значення сусідніх сенсорів (вони передаються або вже декодовані).

Сенсор S2 використовує S1 як допоміжну інформацію. Відмінності між S1 і S2 [1,1,1,1,1]

У двійковому вигляді: [1,1,1,1,1] кодується в компактний синдром.

Передається лише синдром (2–3 біти замість 8 біт).

Сенсор S3 використовує S2 як допоміжна інформація. Відмінності між S2 і S3: [1,1,1,1,1]

Знову передається лише синдром. Центральний вузол отримує дані S1 у повному вигляді {25,26,25,27,26}. Синдроми від S2 та S3. За допомогою S1 і синдрому від S2, вузол відновлює дані S2 {26,27,26,28,27}. За допомогою S2 і синдрому від S3, вузол відновлює дані S3 {27,28,27,29,28}.

Аналіз використання методу DISCUS наведено в таблиці 1.

Таблиця 1. Обсяг переданої інформації з методом DISCUS.

| Параметр                                     | Розмір (байт) |
|----------------------------------------------|---------------|
| Загальна кількість переданих даних(S1,S2,S3) | 15            |
| Повні дані для S1                            | 5             |
| Синдроми для S2 і S3                         | 1.25          |
| Загальний обсяг                              | 7.5           |
| Коефіцієнт стиснення                         | 2.00          |

Стиснення в методі DISCUS напряму залежить від кореляції між даними сусідніх вузлів або джерел. Чим вища кореляція, тим менший обсяг інформації потрібно передавати, оскільки центральний вузол може ефективніше відновити дані з використанням допоміжної інформації.

Таким чином, у доповіді проведено аналіз сучасного методу стиснення даних BCM DISCUS. Напрямом подальших досліджень є експериментальне дослідження ефективності методів стиснення, таких як MRCQ, DIMENSIONS, S-LZW-MS, оцінка їх продуктивності за такими критеріями, як коефіцієнт стиснення, використання пам'яті, кількість операцій та середньоквадратичної помилки, якщо присутня, з використанням ПЗ Matlab2024b.

### Література

1. Akyildiz I. F. et al. A survey on sensor networks //IEEE Communications magazine. – 2002. – Т. 40. – №. 8. – С. 102-114.
2. Nelson M., Gailly J. L. The data compression book 2nd edition //M & T Books, New York, NY. – 1995.
3. Z. Xiong, A.D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," IEEE Signal Processing Magazine, vol. 21, no. 5, pp. 80–94, 2004.