

ПОРІВНЯЛЬНИЙ АНАЛІЗ МЕТОДІВ СТИСКУ ДАНИХ В БЕЗПРОВОДОВИХ СЕНСОРНИХ МЕРЕЖАХ

Третяк А.В., Лисенко О.І.

*Навчально-науковий інститут телекомунікаційних
систем КПІ ім. Ігоря Сікорського, Україна
E-mail: petbka.tasher@gmail.com*

COMPARATIVE ANALYSIS OF DATA COMPRESSION METHODS IN WIRELESS SENSOR NETWORKS

The article analyzes various methods of data compression used in wireless sensor networks. Four main categories are considered: string-based compression, image-based compression, distributed source coding, and compressed sensing. Comparison of these methods for data recovery, lossless or lossy. Algorithms in each category are given, and their principle of action is explained.

Бездротова сенсорна мережа (WSN) складається з одного або декількох віддалених приймачів і великої кількості сенсорних вузлів. Кожен сенсорний вузол є компактним бездротовим пристроєм, який безперервно збирає навколишню інформацію та передає дані датчиків приймачам за допомогою схеми множинної маршрутизації з кількома переходами [1]. WSN відкривають нові можливості для повноцінного обчислення та контекстуального моніторингу фізичного середовища. Зазвичай їх використовують для спостереження за конкретними явищами або відстеження об'єктів. Практичні застосування WSN включають, наприклад, моніторинг тварин, трансформацію сільського господарства, охорону здоров'я, спостереження всередині приміщень та розумні будівлі.

Для зменшення обсягу даних від датчиків необхідно їх стиснути у мережі. В залежності від можливості відновлення даних, можна виділити стиснення: без втрат, із втратою. Стиснення без втрат позначає, що після проведення операції декомпресії ми отримаємо точно такі ж дані, як і до стиснення. Один із стандартних прикладів - це кодування Хаффмана [2]. Стиснення з втратою означає втрату деяких деталей даних через операцію стиснення. До цієї категорії відноситься більшість схем стиснення зображень і відео, таких як JPEG2000 [3].

Методи стиснення даних у WSN, можна класифікувати на категорії: стиснення на основі рядків, стиснення на основі зображень, стиснення за кодуванням розподіленого джерела, стиснутого зондування.

1) Методи стиснення на основі рядків розглядають дані датчиків як послідовність символів, а потім застосовують схеми стиснення даних, які використовуються для обробки текстових даних, щоб стиснути ці дані датчиків.

Успадковані від цих схем стиснення текстових даних методи стиснення на основі рядків також можуть забезпечити стиснення без втрат.

Прикладом може бути алгоритм Lempel-Ziv-Welch (LZW) [4] який динамічно створює словник для кодування нових рядків на основі рядків, які зустрічалися раніше. Відправник, і одержувач мають однакові початкові словникові статті, а всі нові словникові статті можуть бути отримані з існуючих словникових статей і потоку вхідних даних. Спочатку він сканує вхідний потік символів, доки не знайде підрядок, відсутній у словнику. Кодує найдовшу знайдену підрядку індексом відповідності у словнику. Додає цю нову підрядку до словника з черговим доступним кодом. Процес повторюється під час сканування наступних символів вхідного потоку

2) Методи стиснення на основі зображень організовують WSN в ієрархічну архітектуру, а потім застосовують деякі схеми стиснення зображень, такі як вейвлет-перетворення, щоб забезпечити багаторазову роздільну здатність даних зондування всередині мережі. Деякі другорядні функції даних зондування можуть бути втрачені через операції стиснення, тому метод стиснення на основі зображення підтримує стиснення з втратою.

Зображення зазвичай складається з багатьох малих пікселів, і його можна представити у вигляді матриці, де кожен елемент відображає значення одного пікселя. Застосовуючи вейвлет-перетворення до цієї матриці, можна виділити ключові характеристики зображення в частотному діапазоні. В результаті розмір зображення може бути значно зменшений, зберігаючи лише важливі характеристики. Підходи до стиснення даних на основі зображень використовують подібний підхід. Вони організовують бездротову сенсорну мережу в ієрархічну структуру і розглядають дані зондування, надіслані всіма вузлами датчиків, як зображення з кількома пікселями. Після цього застосовується вейвлет-перетворення для отримання просторового та часового узагальнення цих даних зондування. У випадку, коли дані зондування мають високу просторову або часову кореляцію, методи стиснення на основі зображень можуть додатково зменшити обсяг передаваних даних. Нижче описані методи стиснення на основі зображень: структура DIMENSIONS [5] і структура стиснення та запитів із різною роздільною здатністю MRCQ [6].

Метод стиснення даних DIMENSIONS:

1. Використовує ієрархічну архітектуру бездротової сенсорної мережі.
2. Розбиває мережу на кластери, у кожному вибирається головний вузол.
3. Головні вузли збирають дані з датчиків та виконують стиск.
4. Застосовується 3D дискретне вейвлет-перетворення для просторово-часового підсумовування даних.
5. Дані на нижніх рівнях ієрархії мають кращу роздільну здатність.

6. Дані на верхніх рівнях - грубішу роздільну здатність, але менший обсяг.

7. Дозволяє запитувати дані з потрібним дозволом у різних головних вузлів.

Основна перевага полягає в ефективному стисненні великих обсягів корельованих даних датчиків, тому DIMENSIONS добре підходить для стиснення просторово-часових даних датчиків у великих сенсорних мережах. Недолік порівняно з MRCQ - висока обчислювальна складність.

Метод стиснення даних MRCQ (multi-resolution compression and query):

1. Використовує ієрархічну архітектуру бездротової сенсорної мережі.

2. Розбиває мережу на кластери та рівні, у кожному кластері вибирається вузол обробки даних.

3. Вузли нижнього рівня застосовують 2D дискретне косинусне перетворення.

4. Вузли верхніх рівнів передають лише частину найбільш значимих перетворених даних.

5. Дозволяє отримувати дані з різною роздільною здатністю на різних ієрархічних рівнях.

3) Методи кодування розподіленого джерела стискають дані зондування всередині мережі згідно з теоремою Слепяна-Вольфа, яка доводить, що два або більше корельованих потоків даних можуть бути закодовані незалежно, а потім спільно декодовані в приймачі зі швидкістю, що дорівнює їхній спільній ентропії. Таким чином, розподілені методи кодування джерела можуть підтримувати стиснення без втрат [7].

Метод розподіленого вихідного кодування - базується на теоремі Слепяна-Вольфа, яка показує, що при кодуванні двох чи більше корельованих джерел даних, кожне з яких кодується незалежно, а потім спільно декодується, можливо досягти стиснення даних без втрат зі швидкістю бітів, що дорівнює їх спільній ентропії. Він дозволяє сенсорним вузлам стискати дані зондування без кооперації та узгодження між собою, але вимагає попереднього знання кореляції в даних. Застосовується для стиснення даних зондування в бездротових сенсорних мережах у випадках, коли є корельовані потоки даних від різних сенсорних вузлів.

4) Методи стиснутого зондування вказують на те, що будь-які досить стислі дані можна точно відновити з невеликої кількості неадаптивних рандомізованих вибірок лінійної проекції. Таким чином, вони можуть використовувати стиснення, не покладаючись на будь-які попередні знання або припущення щодо даних зондування. З огляду на наведене вище спостереження, методи стиснення можуть забезпечити стиснення без втрат [8].

Метод стиснутого зондування - використовує стисливість даних без покладання на конкретні припущення щодо даних чи їх кореляцій. Баується на

теорії стисненого сприйняття, згідно з якою будь-які достатньо стислі дані можна точно відновити з невеликої кількості випадкових лінійних проєкцій цих даних. Для розріджених даних з невеликою кількістю ненульових елементів цей метод дозволяє отримати їх стиснений набір зі значно меншою розмірністю. Даний метод дозволяє здійснити стиснення даних без втрат без попередньої обробки та координації між сенсорними вузлами.

Коротка порівняльна характеристика цих методів наведена в таблиці 1

Таблиця 1. Коротка порівняльна характеристика методів.

<i>Метод</i>	<i>Категорія стиснення</i>	<i>Відновлення даних</i>	<i>Теорія стиснення</i>
LZW	На основі рядків	Без втрат	-
DIMENSIONS	На основі зображення	З втратами	Вейвлет-перетворення
MRCQ	На основі зображення	З втратами	Вейвлет-перетворення
Розподілене кодування	Розподілене вихідне кодування	Без втрат	Теорема Слепяна-Вольфа
Зондування	Зондування	Без втрат	Зондування

Література

1. I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," IEEE Comm. Magazine, vol. 40, no. 8, pp. 102–114, 2002.
2. M. Nelson and J. L. Gailly, The data compression book.
3. D.S. Taubman and M.W. Marcellin, JPEG2000: Standard for Interactive Imaging.
4. T.A. Welch, "A technique for high-performance data compression," Computer, vol. 17, no. 6, pp. 8–19, 1984.
5. D. Ganesan, B. Greenstein, D. Estrin, J. Heidemann, and R. Govindan, "Multiresolution storage and search in sensor networks," ACM Trans. Storage, vol. 1, no. 3, pp. 277–315, 2005.
6. Y.C. Wang, Y.Y. Hsieh, and Y.C. Tseng, "Multiresolution spatial and temporal coding in a wireless sensor network for long-term monitoring applications," IEEE Trans. Computers, vol. 58, no. 6, pp. 827–838, 2009.
7. 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.
8. J. Haupt, W.U. Bajwa, M. Rabbat, and R. Nowak, "Compressed sensing for networked data," IEEE Signal Processing Magazine, vol. 25, no. 2, pp. 92–101, 2008.