

АЛГОРИТМИ СТИСНЕННЯ ДАНИХ В БЕЗПРОВІДНИХ СЕНСОРНИХ МЕРЕЖАХ

Коченко А.О.

Тернопільський національний економічний університет, магістрант

I. Постановка задачі

Алгоритми стиснення даних почали свою історію розвитку ще задовго до появи першого комп'ютера. В основі всіх алгоритмів стиснення даних закладена проста ідея: якщо подавати елементи, які часто використовуються короткими кодами, а ті, що використовуються рідше – довгими, то для збереження блоку даних буде потрібно менший обсяг пам'яті, порівняно з тим, якщо би всі елементи були представлені кодами однакової довжини [1]. В статичних методах компресії використовують статичну модель даних, і якість стиснення інформації залежить від того, наскільки добре сформована дана модель. Методи, які базуються на словниковому підході не використовують коди змінної довжини, а деякі послідовності символів, які зберігаються в словнику [2]. Тому актуальною задачею є проведення дослідження алгоритмів стиснення даних в безпроводних сенсорних мережах для здійснення розробки асиметричного за складністю алгоритму стиснення даних.

II. Мета роботи

Метою роботи є аналіз існуючих алгоритмів стиснення даних в безпроводних сенсорних мережах.

III. Аналіз алгоритмів стиснення даних у безпроводних сенсорних мережах

Безпроводні сенсорні мережі (БСМ) невинним потоком впливають в повсякденні комп'ютерні пристрої, засоби автоматики та робототехніки, системи електроспоживання та обліку енергії, тощо [3].

Безпроводна сенсорна мережа – це розподілена мережа, яка складається з безлічі датчиків та виконавчих пристроїв, які об'єднані між собою за допомогою радіоканалу. Область покриття подібної мережі може становити від декількох метрів до декількох кілометрів за рахунок здатності ретрансляції повідомлень від одного вузла до іншого.

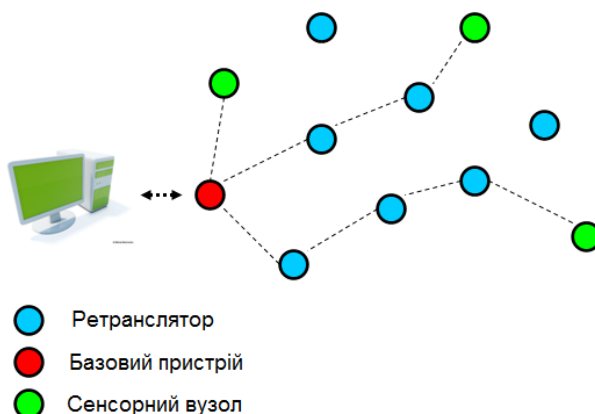


Рисунок 1 – Архітектура типової безпроводної сенсорної мережі

Характерною особливістю більшості типів даних є їх надлишковість. Ступінь надлишковості даних залежить від типу даних. Іншим фактором, що впливає на ступінь надлишковості є прийнята система кодування.

До технічних характеристик процесів стиснення даних та результату їх роботи відносять:

- ступінь стиснення (compress rating) або відношення (ratio) обсягів вихідного і результуючого потоків;
- швидкість стиснення – це час, який використовується на стиснення деякого обсягу інформації вхідного потоку, до одержання з нього еквівалентного вихідного потоку;
- якість стиснення – величина, що показує наскільки сильно запакований вихідний потік, за допомогою застосування до нього повторного стиснення з цього ж або іншого алгоритму.

Для проблеми стиснення інформації використовують декілька різних підходів. Одні мають досить складну теоретичну математичну базу, інші засновані на властивостях інформаційного потоку і алгоритмічно досить прості. Будь-який спосіб, підхід і алгоритм, який реалізує стиснення даних, призначений для зниження обсягу вихідного потоку інформації в бітах за допомогою її зворотного або незворотного перетворення. Тому, перш за все, за критерієм, пов'язаним з характером або форматом даних, всі способи стиснення можна розділити на дві категорії: зворотне і незворотне стиснення.

Під незворотним стисненням розуміється така зміна вхідного потоку даних, при якій вихідний потік, який заснований на певному форматі інформації схожий за своїми зовнішніми характеристиками на вхідний потік, який відрізняється від нього лише обсягом [4].

Зворотне стиснення завжди призводить до зниження обсягу вихідного потоку інформації без зміни його інформативності, тобто – без втрати інформаційної структури.

Існує багато різних практичних методів стиснення без втрати інформації, які, як правило, мають різну ефективність для різних типів даних та різних обсягів. Однак, в основі цих методів лежать три теоретичних алгоритми:

- Алгоритм RLE (Run Length Encoding);
- Алгоритми групи KWE (KeyWord Encoding);
- Алгоритм Хаффмана.

Після проведених досліджень алгоритмів стиснення, визначено степінь стиснення, тобто який відсоток від початкового розміру файлу складає розмір стиснутого файлу, та розархівування – це час, який використано на розархівування в мілісекундах, представлено в таблиці 1.

Таблиця 1

Порівняння алгоритмів стиснення для даних різного обсягу

Програма	Малі (30-45 байт)		Середні (2 кбайта)		Великі (200кбайт)	
	Степінь стиснення	Час, мс	Степінь стиснення	Час, мс	Степінь стиснення	Час, мс
Cmix	116.80	22287	41.09	23600	24.97	112930
Nanozip	297.20	13	53.04	17	32.06	50
Gzip	180.39	1	51.53	1	37.51	3
7zip	420.52	4	57.30	4	33.19	11
Huffman	75.20	1	64.29	1	65.18	27
Huffman with V	47.15	1	27.91	1	28.91	16

Результати, що наведені в таблиці показують, що при малому розмірі даних всі алгоритми збільшували розмір початкового файлу, окрім алгоритму Хаффмана, але кращим за нього є алгоритм Хаффмана зі словником. При середньому розмірі даних переваги має алгоритм Хаффмана з словником, окрім цього всі протестовані алгоритми дозволили зменшити розмір початкового файлу. Для файлів великого розміру найкращим є алгоритм Cmix, але його швидкість роботи дуже низька, тому цей алгоритм не є доцільним для використання. Отже, після проведення дослідження найкращим алгоритмом для стиснення даних є алгоритм Хаффмана зі словником для даних різного розміру.

Висновок

У роботі проведено аналіз алгоритмів стиснення даних в безпроводних сенсорних мережах. Використання алгоритмів стиснення даних дозволить підвищити продуктивність роботи безпроводних сенсорних мереж, зокрема зменшить затрати енергії на передачу даних і тим самим збільшить тривалість роботи мережі від автономного джерела живлення.

Список використаних джерел

1. Cleary J.G., Witten I.H. Data compression using adaptive coding and partial string matching // IEEE Transactions on Communications, Vol. 32(4), pp.396-402, April 1984.
2. Howard P.G. The design and analysis of efficient lossless data compression systems // PhD thesis, Brown University, Providence, Rhode Island. 1993
3. Бойко Ю.М. Концептуальні особливості реалізації безпроводних сенсорних мереж / О.Ю.М. Бойко, В.М. Локазюк, В.В. Мішан // Вісник Хмельницького національного університету. - г 2010. - № 2. - С. 94-97.
4. Селомон Д. Стиснення даних, зображень та звуку/ Москва: Техносфера, 2004. – С. 81.