

# Адаптивный алгоритм сжатия данных в беспроводных сенсорных сетях

Б.Я. Лихтциндер<sup>1</sup>

<sup>1</sup>Поволжский государственный университет телекоммуникаций и информатики, Льва Толстого 23, Самара, Россия, 443010

**Аннотация.** В статье рассматриваются методы сжатия данных в беспроводных сенсорных сетях (БСС). Обсуждаются пять категорий методов, включая сжатие на основе цепочек, сжатие на основе изображений, кодирование с распределенным источником, сжатие на основе агрегации данных. Наиболее подробно рассмотрен предлагаемый автором адаптивный метод сжатия. Показана эффективность применения алгоритма "дырявое ведро" для адаптации сжатия. Рассмотрена структура, реализующая указанный алгоритм. Показаны преимущества введения адаптации при сжатии данных в БСС.

## 1. Введение

Беспроводная сенсорная сеть состоит из одного или нескольких удаленных приемников и большого количества сенсорных узлов. Каждый сенсорный узел представляет собой небольшое беспроводное устройство, которое может непрерывно собирать информацию и сообщать данные в приемник через схему множественной маршрутизации с несколькими переходами [1]. БСС предоставляют новую возможность для повсеместного компьютерного и контекстно-ориентированного мониторинга физических сред. Они обычно развертываются в областях, представляющих интерес, для наблюдения специфических явлений или объектов. Практические применения БСС включают: мониторинг животных, предприятий сельского хозяйства, здравоохранение, надзор и интеллектуальные здания [2-6]. Поскольку узлы датчиков обычно питаются от батарей, и многие приложения БСС нацелены на долгосрочный мониторинг окружающей среды, то важной проблемой является сохранение энергии сенсорных узлов, с целью продления срока их службы. Существует два общих решения экономии энергии сенсорных узлов. Одним из решений является использование избыточности узлов, выбирая для активизации те подмножества сенсорных узлов, которые имеют наибольший запас энергии, по сравнению с другими [7-9]. Выбранный поднабор активных сенсорных узлов должен охватывать всю область мониторинга и поддерживать сетевое подключение. Другими словами, эти активные узлы датчиков должны следить за тем, чтобы сеть по-прежнему функционировала нормально. Путем выбора различных подмножеств сенсорных узлов, которые будут активны по очереди, мы можем помешать некоторым узлам датчиков потреблять слишком много энергии и, таким образом, продлить срок службы сети. Однако, часто избыточность узла использована быть не может (например, из-за занятости сети [10], [11] или из-за прерывания, например, механизмом сна). Альтернативным решением является уменьшение количества данных, отправляемых датчиками, поскольку передача является одной из наиболее энергозатратных операций сенсорных узлов. Такое решение особенно полезно, когда узлы

датчиков должны регулярно сообщать данные своих измерений приемнику в течение длительного промежутка времени. Чтобы уменьшить количество передаваемых данных, необходимо сжать их внутри сети. В зависимости от возможности восстановления данных можно классифицировать схемы сжатия данных по трем категориям: без потерь, с потерями и необратимое сжатие. Сжатие без потерь означает, что после выполнения операции декомпрессии мы можем получить точно такие же данные, что и перед выполнением операции сжатия. Кодирование Хаффмана [12] является одним из типичных примеров. Сжатие с потерями означает, что некоторые детализированные (и обычно незначительные) данные могут быть потеряны из-за операции сжатия. Большинство схем сжатия изображений и видео, таких как JPEG2000 [13], относятся к этой категории. Наконец, сжатие с потерями означает, что операция сжатия необратима. Другими словами, нет операции декомпрессии. Например, можно сжать набор значений, принимая их среднее значение, но каждое из исходных чисел не может быть получено из этого среднего значения.

Методы сжатия данных в БСС, могут быть классифицированы по пяти категориям.

1) Методы сжатия на основе построения префиксных кодов рассматривают воспринимаемые данные как последовательность символов, а затем применяют к ним схемы сжатия, используемые для обработки текстовых данных. Унаследованные от этих методов способы сжатия на основе строк также могут обеспечивать сжатие без потерь.

2) Методы сжатия на основе изображений организуют БСС в иерархическую архитектуру, а затем применяют некоторые схемы сжатия изображений, такие как вейвлет-преобразование, чтобы обеспечить множественное разрешение воспринимаемых данных внутри сети. Некоторые незначительные характеристики таких данных могут быть потеряны из-за операций сжатия, и, таким образом, технология сжатия на основе изображения поддерживает сжатие с потерями.

3) Методы кодирования распределенных источников сжимают данные считывания внутри сети в соответствии с теоремой Слепьяна -Вольфа и работой [14], которая доказывает, что два или более коррелированных потока данных могут быть закодированы независимо друг от друга и затем декодироваться совместно в приемнике со скоростью, соответствующей их совместной энтропии. Поэтому методы кодирования с распределенным источником могут поддерживать сжатие без потерь.

4) Методы сжатого зондирования показывают, что любые достоверно сжимаемые данные могут быть точно восстановлены из небольшого числа неадаптивных рандомизированных образцов линейной проекции. Таким образом, они могут использовать сжимаемость, не полагаясь на какие-либо предварительные знания или предположения об измеряемых данных. При вышеуказанном наблюдении методы сжатого зондирования могут обеспечить сжатие без потерь.

5) Методы агрегирования данных передают сжатые агрегированные данные в приемник [17]. Поскольку из этих агрегированных данных не могут быть получены данные исходных измерений, сжатие методами агрегирования данных является необратимым. К таким методам относится рассматриваемый в этой статье метод адаптивного сжатия.

## **2. Адаптивное сжатие**

Наиболее часто сжатие передаваемых данных осуществляется с помощью вейвлет-преобразований. Некоторые незначительные характеристики измеряемых данных могут быть потеряны из-за операций сжатия, и, таким образом, подобная технология поддерживает сжатие с потерями. Сущность предлагаемого нами метода заключается в том, что указанные потери (мгновенные отклонения реального процесса от восстановленного из сжатого процесса) анализируются, и при выходе анализируемого параметра отклонения за определенные установленные пределы, начинает передаваться полная информация о процессе в не сжатом виде. В качестве анализирующего алгоритма предлагается использовать известный алгоритм «дырявое ведро», широко применяющийся при анализе трафика сетей связи с коммутацией пакетов. Именно в указанном случае анализа параметра отклонения возможно применение

результатов обобщенных формул разработанного нами интервального метода анализа очередей в системах массового обслуживания [15].

Рассмотрим работу указанного алгоритма на рисунке 1. Измерительная информация, поступающая от сенсора, кодируется с помощью АЦП, и в виде цифровых отсчетов  $m_{ik}$  поступает на вход буферной памяти, где отсчеты хранятся последовательно. Одновременно, в памяти может находиться  $N$  отсчетов. При поступлении очередного отсчета, наиболее ранний отсчет удаляется.

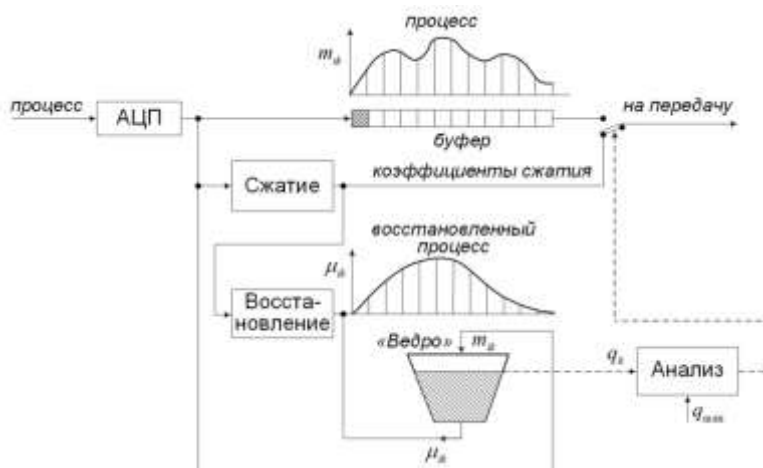


Рисунок 1. Структурная схема адаптивного сжатия.

Отсчеты также поступают на устройство сжатия. Коэффициенты, получаемые в результате сжатия процесса на данном цикле (любым из известных способов, например, с помощью вейвлет), поступают циклически на передачу. Одновременно отсчеты поступают на суммирующий вход обработки алгоритмом "дырявое ведро".

алгоритмом "дырявое ведро". По коэффициентам, получаемым в результате "сжатия", производится восстановление процесса, результаты которого в виде отсчетов  $\mu_{ik}$  поступают на вычитающий вход обработки алгоритмом "дырявое ведро". Уровень "наполнения ведра"  $q_{ik}$  сравнивается с заданным максимальным значением  $q_{max}$ , и в случае превышения, происходит переключение на передачу всех не сжатых данных о процессе из буферной памяти. Если в течение цикла превышение отсутствует, то на передачу поступают только коэффициенты, полученные в результате сжатия исходного процесса.

Способ сжатия с помощью вейвлет требует задержки передачи информации на интервал времени, равный длительности одного цикла, поскольку значения всех коэффициентов разложения могут быть получены лишь по завершению цикла. Если в сети производится циклический доступ методом TDMA [16], то рассмотренный способ сжатия можно считать оправданным. Если же используется метод случайного доступа и измерительная информация поступает в сеть непосредственно, то может возникнуть задержка между моментом существенного отклонения процесса и моментом передачи информации об этом отклонении. В указанных случаях целесообразней использовать способ сжатия на основе прогнозирования.

### 3. Алгоритм сжатия на основе прогнозирования

Рассмотрим такой алгоритм подробнее:

Аналого-цифровой преобразователь (АЦП) сенсора периодически с интервалом времени  $\tau$  определяет значения  $m_{ik}$ , которое записывается в последнюю ячейку буферной памяти, при этом значение из первой ячейки памяти удаляется (в памяти всегда должно находиться  $N$  значений).

$m_{i0}$  – начальный отсчет на  $i$ -м интервале времени (цикле). Указанные отсчеты считаются главными и передаются в сеть.

$m_{ik}$  – результат измерения на  $k$ -м такте  $i$ -го интервала времени. Указанные отсчеты являются промежуточными и обрабатываются в передатчике.

$N$  – число тактовых измерений, одинаковое для всех циклов.

На передающей стороне, по результатам обработки  $n$  главных отсчетов, предшествовавших отсчету  $m_{i0}$  ( $m_{i-j,0}$   $j=1,2,\dots,n$ ), определяются разности до  $(n-1)$ -го порядка, а по ним – значения коэффициентов разложения в ряд Тейлора. По указанным коэффициентам производится восстановление значений на всех тактах  $i$ -го цикла.

$\mu_{ik}$  – результат восстановления на  $k$ -м такте измерений  $i$ -го цикла  $u_{ik} = m_{ik} - \mu_{ik}$  – разность между действительным и спрогнозированным значениями процесса на  $k$ -м такте  $i$ -го цикла. Примем  $u_{i0} = 0$ .  $q_{ik}$  – очередь (наполнение "ведра" на  $k$ -м такте  $i$ -го цикла). Значения  $q_{ik}$  определяются уравнением баланса [15].

$$q_{i(k+1)} = q_{ik} + u_{ik} \text{ если } q_{ik} + u_{ik} \geq 0 \\ q_{i(k+1)} = 0 \text{ если } q_{ik} + u_{ik} < 0$$

Именно величина  $u_{ik}$  показывает отличие спрогнозированного процесса от реального. Оценкой степени отклонения является интегральная величина  $q_{ik}$ . Если ее значение не превышает допустимого, то на данном цикле передается только главный отсчет, Если значение превышает допустимое, то передается вся информация о процессе из буферной памяти в несжатом виде.

Получив значения главных отсчетов, приемная сторона производит определение коэффициентов и восстановление промежуточных значений аналогично тому, как это происходило на передающей стороне. При этом, имеется гарантия, что восстановленный процесс незначительно отличается от реального. В противном случае, в приемник поступит полная информация о процессе из буферной памяти.

#### 4. Заключение

Сенсорные узлы, как правило, питаются от батарей, и, следовательно, сохранение их энергии является основной проблемой БСС. Сжатие данных в сети помогает уменьшить количество измеряемых данных, которые сенсорные узлы должны регулярно сообщать приемнику и, следовательно, значительно снизить потребление энергии. Адаптация алгоритма сжатия к постоянно изменяющимся во времени данным, значительно повышает его эффективность. Предлагаемый алгоритм позволяет не только отслеживать текущие изменения, но также, в случае значительных отклонений, предоставляет информацию о предыстории контролируемого процесса.

#### 5. Литература

- [1] Akyildiz, I.F. A survey on sensor networks / I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci // IEEE Comm. Magazine. – 2002. – Vol. 40(8). – P. 102-114.
- [2] Zhang, P. Hardware design experiences in ZebraNet / P. Zhang, C.M. Sadler, S.A. Lyon, M. Martonosi // Proc. ACM Int'l Conf. Embedded Networked Sensor Systems, 2004. – P. 227-238.
- [3] Tseng, Y.C. iMouse: an integrated mobile surveillance and wireless sensor system / Y.C. Tseng, Y.C. Wang, K.Y. Cheng, Y.Y. Hsieh // Computer. – 2007. – Vol. 40(6). – P. 60-66.
- [4] Wark, T. Transforming agriculture through pervasive wireless sensor networks / T. Wark, P. Corke, P. Sikka, L. Klingbeil, Y. Guo, C. Crossman, P. Valencia, D. Swain, G. Bishop-Hurley // IEEE Pervasive Computing. – 2007. – Vol. 6(2). – P. 50-57.
- [5] Huang, Y.M. Pervasive, secure access to a hierarchical sensor-based healthcare monitoring architecture in wireless heterogeneous networks / Y.M. Huang, M.Y. Hsieh, H.C. Chao, S.H. Hung, J.H. Park // IEEE J. Selected Areas in Comm. – 2009. – Vol. 27(4). – P. 400-411.
- [6] Yeh, L.W. iPower: an energy conservation system for intelligent buildings by wireless sensor networks / L.W. Yeh, Y.C. Wang, Y.C. Tseng // Int'l J. Sensor Networks. – 2009. – Vol. 5(1). – P. 1-10.

- [7] Improving wireless sensor network lifetime through power aware organization / M. Cardei, D.Z. Du // *ACM Wireless Networks*. – 2005. – Vol. 11(3). – P. 333-340.
- [8] Zou, Y. A distributed coverage- and connectivity-centric technique for selecting active nodes in wireless sensor networks / Y. Zou, K. Chakrabarty // *IEEE Trans. Computers*. – 2005. – Vol. 54(8). – P. 978-991.
- [9] Zhao, Q. Lifetime maximization for connected target coverage in wireless sensor networks / Q. Zhao, M. Gurusamy // *IEEE/ACM Trans. Networking*. – 2008. – Vol. 16(6). – P. 1378-1391.
- [10] Heo, N. Energy-efficient deployment of intelligent mobile sensor networks / N. Heo, P.K. Varshney // *IEEE Trans. Systems, Man and Cybernetics – Part A: Systems and Humans*. – 2005. – Vol. 35(1). – P. 78-92.
- [11] Wang, Y.C. Efficient placement and dispatch of sensors in a wireless sensor network / Y.C. Wang, C.C. Hu, Y.C. Tseng // *IEEE Trans. Mobile Computing*. – 2008. – Vol. 7(2). – P. 262-274.
- [12] Nelson, M. The data compression book / M. Nelson, J. L. Gailly – MIS Press, 1996.
- [13] Taubman, D.S. JPEG2000: fundamentals, standards and practice / D.S. Taubman, M.W. Marcellin // Kluwer Academic Publishers, 2002.
- [14] Haupt, J. Compressed sensing for networked data / J. Haupt, W.U. Bajwa, M. Rabbat, R. Nowak // *IEEE Signal Processing Magazine*. – 2008. – Vol. 25(2). – P. 92-101.
- [15] Лихтциндер, Б.Я. Интервальный метод анализа трафика мультисервисных сетей / Б.Я. Лихтциндер // *Модели инфокоммуникационных систем: разработка и применение. Приложение к журналу ИКТ. Вып. 8.* – 2011. – С. 101–152.
- [16] Голубничая, Е.Ю. Применение детерминированного расписания TDMA в беспроводных сенсорных сетях // *Проблемы техники и технологии телекоммуникаций: XX Междунар. науч.-техн. конф. Оптические технологии в телекоммуникациях: XVI Междунар. науч.-техн. конф.* – Уфа: РИК УГАТУ. – 2018. – Т. 1. – С. 135-137.
- [17] Голубничая, Е.Ю. Агрегирование данных в беспроводных сенсорных сетях мониторинга // *Проблемы передачи информации в инфокоммуникационных системах: сборник докладов и тезисов – Волгоград: Изд-во ВолГУ, 2017.* – С. 37-42.

## Adaptive data compression algorithm in wireless sensor networks

**B.Ya. Likhttcinder<sup>1</sup>**

<sup>1</sup>Povolzhskiy State University of Telecommunications and Informatics, L. Tolstoy street 23, Samara, Russia, 443010

**Abstract.** The article provides an overview of the latest methods of data compression in wireless sensor networks (WSN). Five categories of techniques are discussed, including chain-based compression, image-based compression, distributed-source encoding, and data-aggregation compression. The adaptive compression method proposed by the author is considered in most detail. The efficiency of the holey bucket algorithm for compressing adaptation is shown. The structure that implements the specified algorithm is considered. The advantages of introducing adaptation during data compression in the WSN are shown.