Chapter

Data Compression Strategies for Use in Advanced Metering Infrastructure Networks

Chih-Wei Hsu and Sun-Yuan Hsieh

Abstract

Internet of Things technology has advanced rapidly. For example, numerous sensors can be deployed in a city to collect a variety of data, and such data can be used to monitor the city's situation. A possible application of such data is smart metering implemented by power suppliers for their consumers; smart metering involves installing a multiplicity of smart meters that, in conjunction with data centers, form a smart grid. Because a smart gird must collect and send data automatically, the establishment of advanced metering infrastructure (AMI) constitutes the primary step to establishing a smart grid. However, problems remain in smart metering: data traffic from smart meters flows rapidly at a huge volume, resulting in bandwidth bottlenecks. Thus, this chapter proposes some data compression technologies as well as a novel scheme for reducing the communication data load in AMI architectures.

Keywords: smart grid, smart meter, concentrator, compression, advanced metering infrastructure, data traffic

1. Introduction

The Internet of Things (IoT) connects a multiplicity of devices to make life convenient. Smart grids constitute one implementation of IoT, and many countries have launched smart grids to develop integrated energy supply systems. A smart grid incorporates automation, bidirectional communication, and advanced sensor measurement systems to streamline interactions between the client and power supplier [1–4]. In contrast to traditional grids, smart grids enable power suppliers to distribute power efficiently and to control the use of power during a given period [5]. Furthermore, the data collected from a smart grid enable automatic billing.

The analysis of energy consumption data has many advantages for both consumers and suppliers: Consumers can track their energy usage, particularly as it varies with the seasons. Power suppliers can monitor how power is utilized across the distribution grid, which can help them formulate power management and energy-saving measures [6–8]. The establishment of advanced metering infrastructure (AMI) constitutes the first step to constructing such an intelligent power gird.

AMI is central to a smart grid system and enables the system to automatically monitor usage. **Figure 1** illustrates the architecture of a typical AMI, indicating that it comprises several concentrators and smart meters that are connected to a meter data management system (MDMS) [9]. The smart meters send data to the

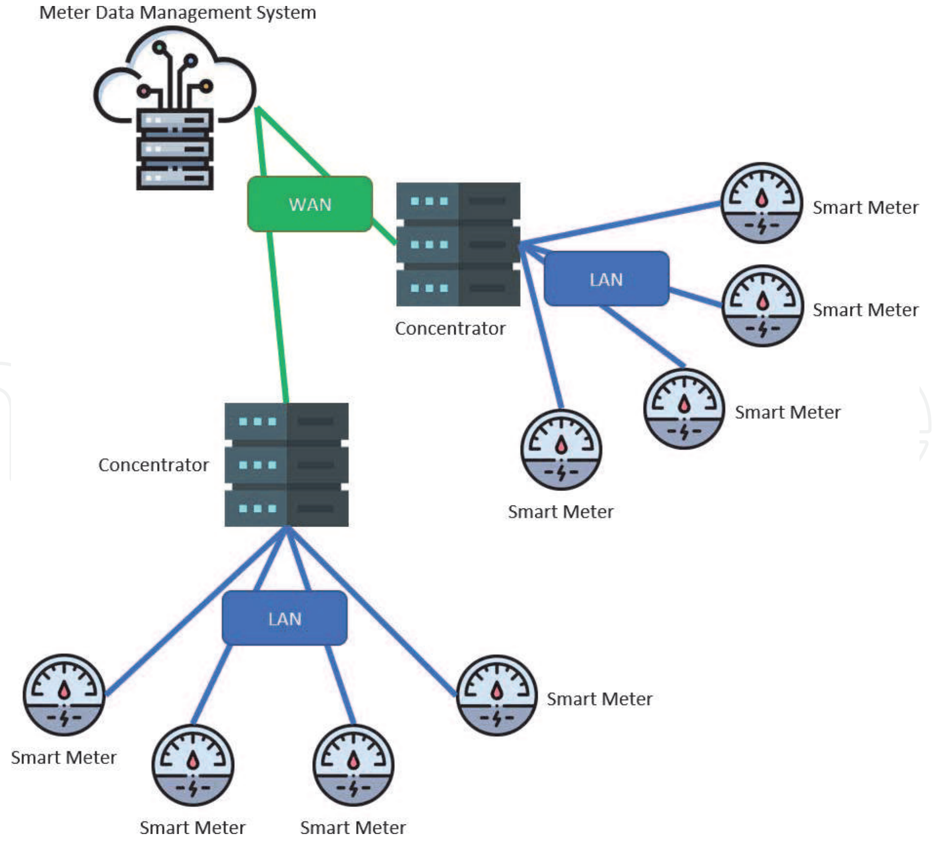

Figure 1.
AMI structure.

concentrators, and the concentrators send the data to the MDMS. The AMI architecture requires high-quality and high-speed networks for providing stable service and efficient monitoring, and it comprises two such networks, namely a local area network (LAN) and wide area network (WAN). The LAN connects the concentrators and smart meters and leverages cutting-edge communication technologies such as power-line communication and the Zigbee specification. The WAN connects the concentrators and MDMS and constitutes the most important part of the AMI system; this network supports a host of high-speed technologies such as broadband-over-power-line technology, 3G, and long-term evolution [10].

Nonetheless, smart meters have a very large data frequency—with a small packet being generated every 15–60 min—that will exceed the capacity of existing communication technologies [11, 12]; this makes data difficult to transmit. Furthermore, smart grids face challenges in storing these large volumes of data. To solve these two problems, data size in both communication and storage should be reduced. This chapter introduces some novel approaches to doing so.

2. Preliminaries

Smart grids have expanded considerably, and many organizations have noticed the advantages engendered by smart meters. Thus, research on smart grids has similarly expanded. Specifically, smart grids enable the automatic distribution of power on the part of suppliers and provide usage data, which can be used to develop various applications.

Smart meters constantly upload data to the MDMS. Such data pertain to, for example, interval data readings, meter remote disconnections, meter remote reconnections, and meter firmware patches. The volume of such data is considerably large, potentially causing network collapse. The reduction of packet size in a smart grid addresses this problem, and such a reduction has become a major topic of research. Thus, this section introduces some basic strategies that are currently adopted to solve this problem.

2.1 Basic method for data reduction in an AMI system

Among existing methods, the conventional approach is combining the messages sent by multiple meters to reduce either protocol overhead or the frequency of transmission. In this approach of message concatenation, energy consumption data, control signal messages, and firmware update packages are combined. This scheme entails a lower transmission volume because the packet header is reduced, where multiple messages can be sent in a header. However, one issue in this scheme is where messages ought to be concatenated.

2.1.1 Meter side

Message concatenation at the meter side results in a greater reduction in data size than does that at the concentrator side. Data size reduction is greater at the meter side primarily because a meter generates data within a certain period; thus, a longer interval between instances of meter-to-concentrator data transmission yields disproportionately large savings in data size. However, most meters contain low-performance hardware, hampering this method. Consequently, meters are either overloaded or unable to compress the data.

2.1.2 Concentrator side

Messages can also be concatenated at the concentrator side prior to transmission. This is the appropriate site for message concatenation [13, 14] because concentrators have more powerful hardware and process data more efficiently than meters do. In concentrator-side concatenation, meters send packages to the concentrator, the concentrator subsequently aggregates the data into one package, and the compressed message is finally sent to the MDMS. Concentrator-side concatenation not only reduces the volume of messages but also stabilizes the system.

However, despite the advantages of concentrator-side concatenation, existing devices on the market are incapable of supporting such concatenation. Currently, concentrators on the market are capable of executing only simple integrations, and these products all follow the PRIME standard of WAN communication [15].

3. Novel approaches

A message generated by a smart meter contains a 40–60-B header within the packet. Concentrators can collect messages sent from the meter side according to a Poisson process [16] before combining these messages to minimize possible traffic. The aforementioned procedure comprises two steps, namely combination and compaction, which are detailed as follows [17].

3.1 Combination

Combination entails combining messages into a larger packet to solve the problem of protocol overhead during communication. **Figure 2** illustrates how such combination reduces traffic. Combining meter messages on smart grids is efficient because each piece of data generated by a meter have a size of only a few bytes.

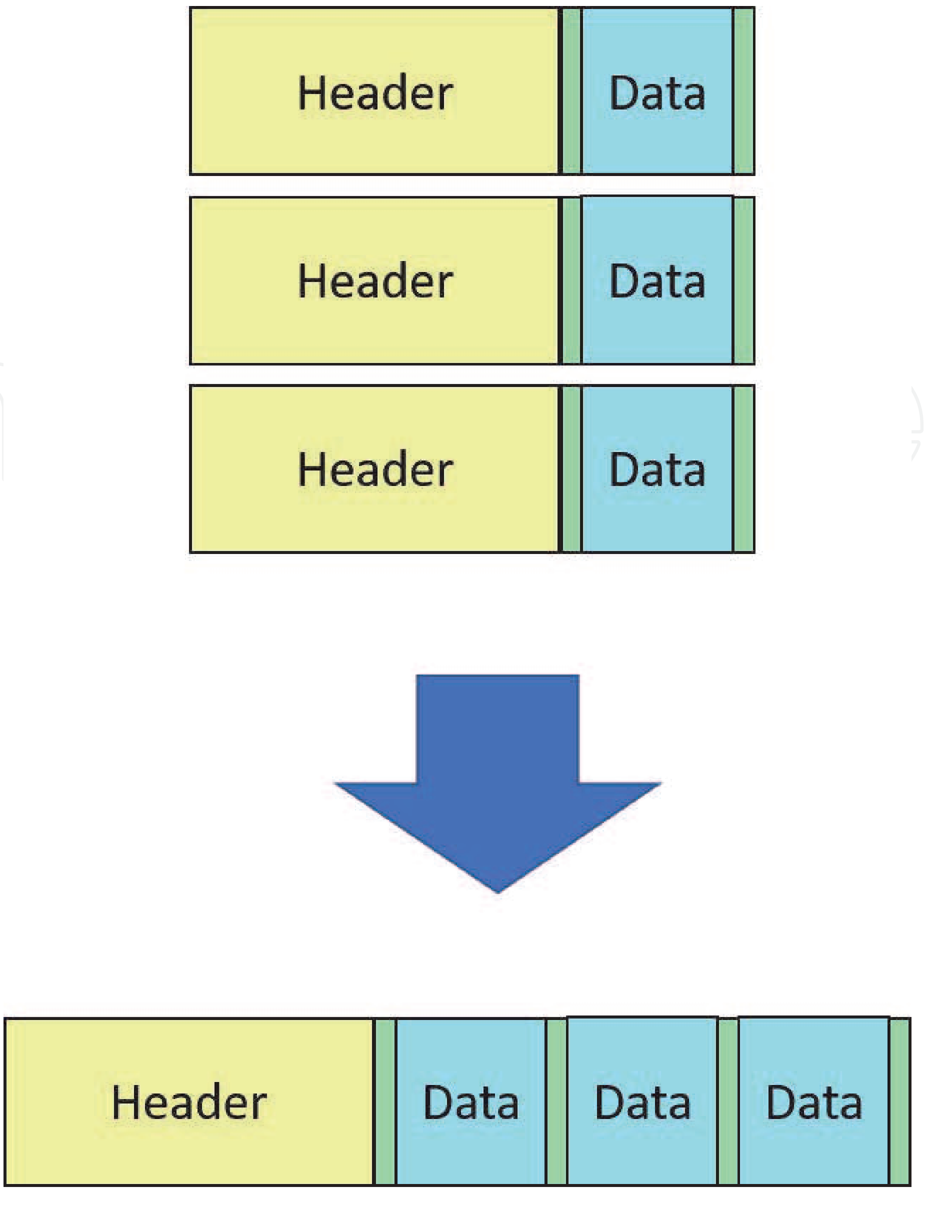

Figure 2.
Concept underlying the combination procedure.

Concentrators combine messages that have been sent by meters in a specified period, after which they compact the messages.

Figure 3 illustrates the operating procedures of a typical combination algorithm. Specifically, the concentrators receive meter messages, place the messages into a combination queue, combine the messages in the queue once some predetermined quantity of messages have been accumulated, and finally compact the messages into a package [17]. This method enables solving the protocol overhead problem. The data compaction scheme is detailed as follows.

3.2 Compaction

Compaction reduces the size of headers in large-scale AMI architecture, which is very helpful for reducing data volume. Messages can be compacted through either full compaction (FC) or loose compaction (LC) [17].

3.2.1 Full compaction

Figure 4 illustrates the operating procedures of an FC algorithm. Specifically, any piece of data received from the meter is stored in Q_c, and the usage in the

DOI: http://dx.doi.org/10.5772/intechopen.94974

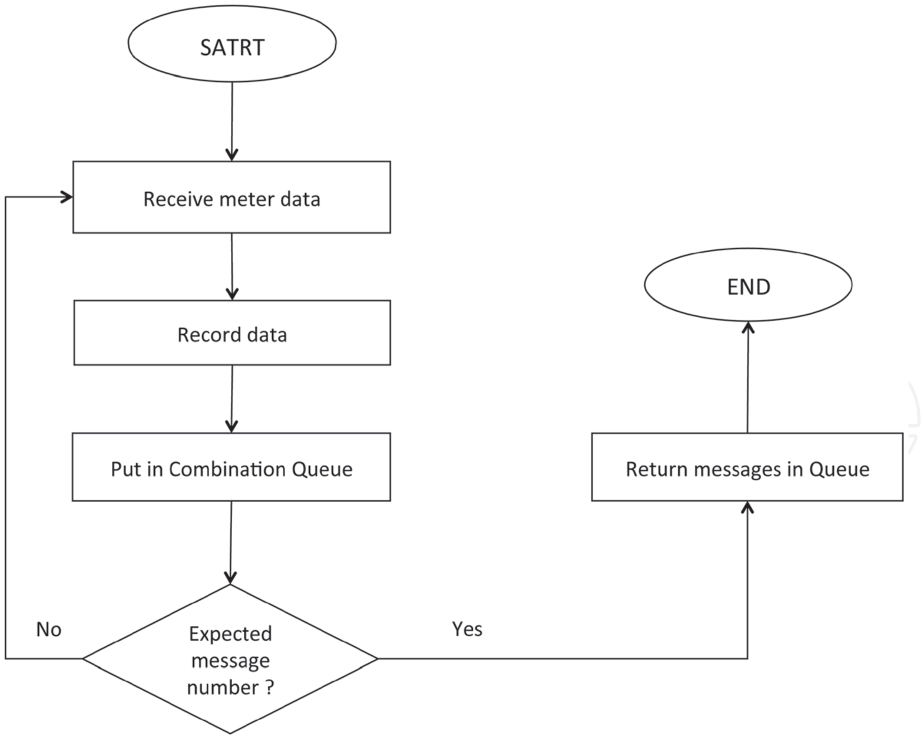

Figure 3.
Combination procedure.

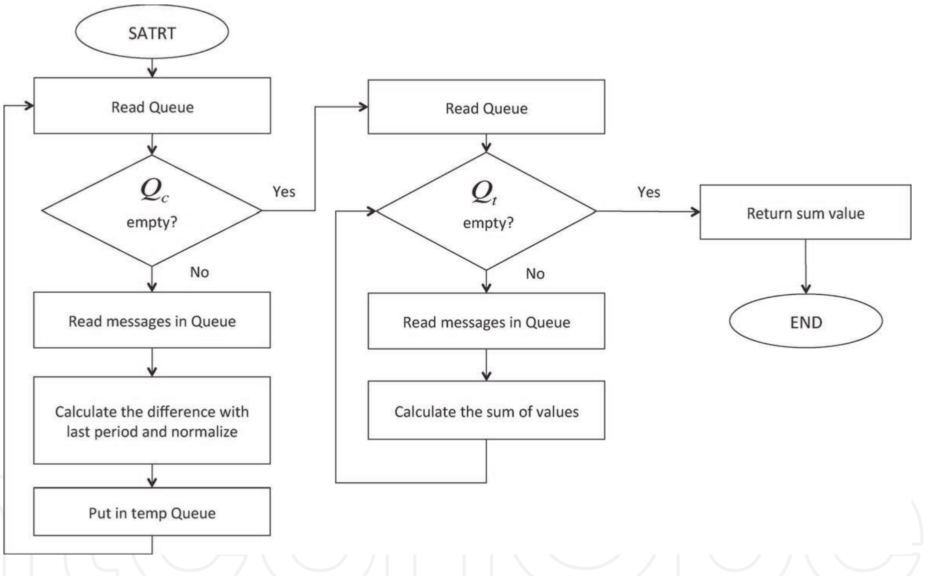

Figure 4.
FC procedure.

current period is calculated by determining the difference between the message in the current period and the message in the last period; this process is performed until Q_c is empty. Subsequently, before the usage data are stored in Q_t, the calculated value should be normalized; this can specifically be achieved by multiplying the value by 10^3 to convert any floating-point number into an integer. Once all messages are in Q_t, the concentrators can start calculating the sum of all values in Q_t. Once the FC procedure is complete, the concentrators obtain the total usage from the meter data, which helps power suppliers monitor the usage of an area to optimize power delivery.

To illustrate the FC procedure, **Table 1** lists some energy usage values in an area. First, all values are input into Q_c. Second, the difference between the previous usage and current usage is calculated for each period. Third, the values obtained from these calculations (212.880, 323.769, 388.112, 46.115, and 610.762 in this example)

Meter Number	Previous Usage	Current Usage	Difference
1	2153.232	2366.112	212.880
2	5698.564	6022.333	323.769
3	23154.003	23542.115	388.112
4	1542.121	1588.236	46.115
5	56213.225	56823.987	610.762

An Example of meter data.

Table 1.
Example meter data from an area.

are input into Q_t. Finally, these values in Q_t are summed to indicate the net difference in energy usage (1581.638 in this example).

3.2.2 Loose compaction

LC differs from FC in that LC is recoverable, whereas FC is not. **Figure 5** illustrates operating procedures of an LC algorithm. First, the concentrators obtain meter data by dequeuing Q_c and subtracting the meter value for the current period from the meter value for the previous period. Second, similar to second step of the FC algorithm, the value obtained in the first step is multiplied by 10^3 for normalization and subsequently input into Q_t. Third, LC sorts values in Q_t in descending order. Fourth, the positive difference between a value and its lower-valued neighbor is computed and then input into Q_l; the initial, lowest value in Q_t thus remains unchanged. For example, if $[a, b, c, d]$ is input into Q_t, then $[a, (b-a), (c-b), (d-c)]$ is input into Q_l. Finally, the values returned by the LC algorithm constitute the compacted data.

To illustrate the LC procedure, consider the values in **Table 1**. First, the LC algorithm takes the difference between past and present usages and inputs these differences into Q_c. In this example, these differences can be listed as follows: [212.880, 323.769, 388.112, 46.115, 610.762]. This list is then sorted in descending order as [46.112, 212.880, 323.769, 388.112, 610.762] and input into Q_t. The LC algorithm then dequeues each piece of data from Q_t and calculates the positive difference between a value and its lower-valued neighbor. The list of such

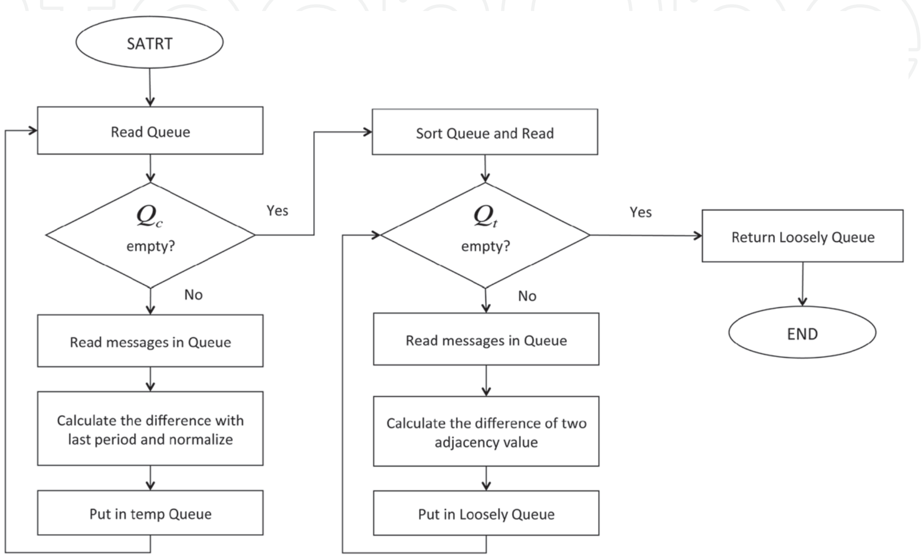

Figure 5.
LC flow.

differences is as follows: [46.112, 166.768, 110.889, 64.343, 222.650]. This list is input into Q_l and returned.

After compacting the data, the concentrators send the data to the MDMS. When the MDMS receives the data, it executes the LC recovery (LCR) algorithm to extract the raw data. **Figure 6** illustrates the LCR algorithm, which is the reverse of the LC algorithm. First, the LCR algorithm inputs the received data into Q_l; the algorithm then recovers those values prior to subtraction from a lower-valued neighbor. These recovered values are then input into Q_t. Second, the LCR algorithm denormalizes each value in Q_t and adds to it the value of the usage in the previous period. Finally, the MDMS outputs all values into Q_c and returns them.

Consider the values for the example presented in **Table 1**. First, the LCR algorithm recovers the compacted data; specifically, the MDMS inputs [46.112, 166.768, 110.889, 64.343, 222.650] into Q_l. Subsequently, the LCR algorithm recovers the

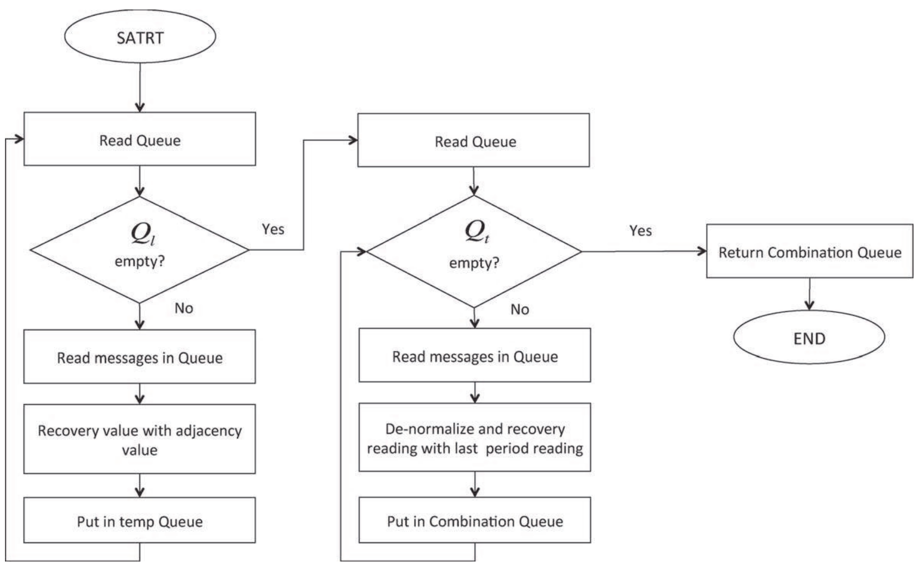

Figure 6.
LC recovery procedure.

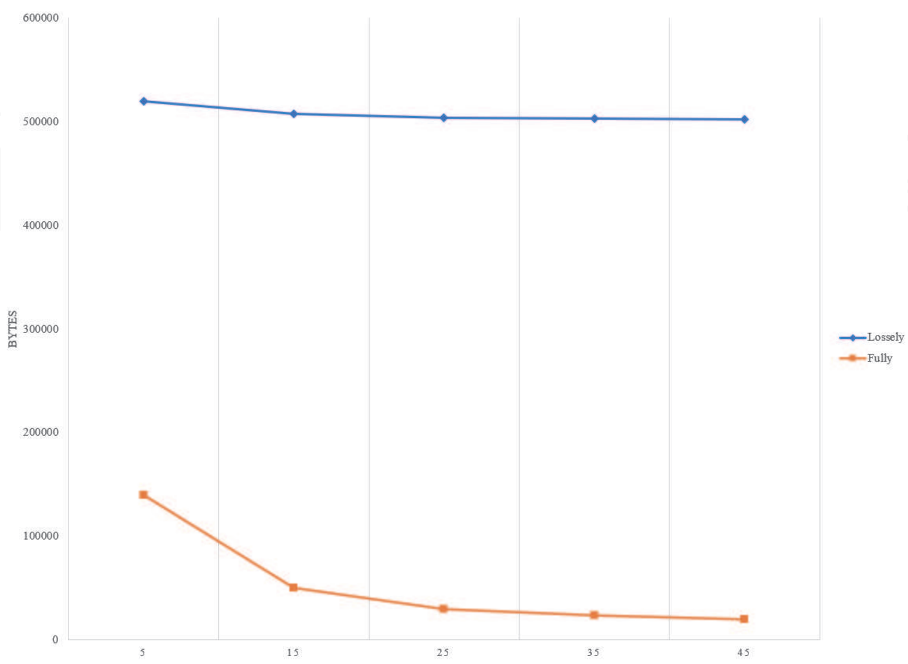

Figure 7.
FC and LC analysis.

values prior to subtraction from a lower-valued neighbor; specifically, the MDMS obtains [46.112, 212.880, 323.769, 388.112, 610.762] and inputs this list into Q_t. Finally, the LCR algorithm denormalizes the data by multiplying each value by 10^{-3}; specifically, the MDMS obtains the list of raw data [46.112, 212.880, 323.769, 388.112, 610.762].

3.2.3 Result analysis

First of all, all of messages sent from meters were encoded in ASCII. FC and LC compaction performance are provided in this section. The **Figure 7** shows the compression ratio of the FC and LC method. The vertical scale label present the total bytes used before compaction and the horizontal scale label is the number of meter messages compacted when compaction. As the **Figure 7** shows, FC has better performance than LC. However, FC does not support recovery, so it can achieve the highly compaction ratio. On the contrary, LC is the trade-off algorithm when recovery function is required and have demand of compression. If the system only need the sum of the energy usage, then FC would be the proper way for compressing the data.

4. Compression technologies

The choice of compression algorithm has become a critical issue because most IoT devices have hardware limitations, particularly in their low-power-consumption microprocessors. The best compression algorithm provides the best compression ratio given the hardware specifications of an IoT device [18, 19]. This section discusses two well-known approaches to data compression [20].

4.1 Lempel–Ziv–Markov chain algorithm

The Lempel–Ziv–Markov chain algorithm (LZMA) is a member of the LZ-family, and it is based on the LZ77 algorithm, which uses a dictionary-based scheme. The LZMA yields excellent compression ratios without being demanding on hardware, making it suited to IoT environments by virtue of its exclusive dictionary structure [19].

4.2 Prediction by partial matching

Prediction by partial matching (PPM) can predict a subsequent pattern using present or previous symbols. Moreover, PPM can be used with a Markov model to construct a compression algorithm. For example, the RAR algorithm, developed by Eugene Roshal and Alexander Roshal, uses PPM and the Lempel–Ziv–Storer–Szymanski algorithm to achieve impressive compression [20].

5. Conclusions

Smart grids will be key infrastructure considering the rapid developments in IoT technology. This chapter presents data compression techniques, such as combination and compaction, developed for reducing the communication data load in a smart grid. These techniques not only reduce the frequency between instances of transmission but also considerably reduce data volume. Moreover, FC and LC can be used in different situation. FC provides extremely compact ratio for the user who have less bandwidth for transport the meter data. LC can be used when the system requires the raw data of the energy usage and having sufficient network bandwidth. In addition, some well known data compression techniques are also introduce in the chapter. Proposed algorithm can be implement with the compression technique provided above to decrease the volume of the meter data. Also, this chapter explains recent advances in smart grid technology. Readers can build on the aforementioned algorithms to formulate novel contributions of their own.

Author details

Chih-Wei Hsu and Sun-Yuan Hsieh*
National Cheng Kung University, Taiwan

*Address all correspondence to: hsiehsy@mail.ncku.edu.tw

References

[1] Xi Fang, Satyajayant Misra, Guoliang Xue, and Dejun Yangand, "Smart grid-the new and improved power grid: a survey," IEEE Communications Surveys & Tutorials, vol. 14, no. 4, pp.944–980, October 2012.

[2] Nanpeng Yu, Sunil Shah, Raymond Johnson, Robert Sherick, Mingguo Hong and Kenneth Loparo, "Big data analytics in power distribution systems, " 2015 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT), June 2015.

[3] Ruofei Ma, Hsiao-Hwa Chen, Yu-Ren Huang, and Weixiao Meng, "Smart grid communication: its challenges and opportunities," IEEE Transactions on Smart Grid, vol. 4, no. 1, pp. 36–46, March 2013.

[4] D. Alahakoon, X. Yu, "Smart Electricity Meter Data Intelligence for Future Energy Systems: A Survey, "IEEE Transactions on Industrial Informations, vol. 12, no. 1, February 2016.

[5] Melike Erol-Kantarci, and Hussein T. Mouftah, "Energy-efficient information and communication infrastructures in the smart grid: a survey on interactions and open issues, "IEEE Communications Surveys & Tutorials, vol. 17, no. 1, pp. 179–197, March 2015.

[6] Z. Fan, P. Kulkarni, S. Gormus, C. Efthymiou, G. Kalogridis, M. Sooriyabandara, Z. Zhu, S. Lambotharan, and W. H. Chin, "Smart grid communications: overview of research challenges, solutions, and standardization activities, "IEEE Communications Surveys & Tutorials, vol. 15, no. 1, pp. 21–38, February 2013.

[7] M. Ghofrani, M. Hassanzadeh, M. Etezadi-Amoli, and M. Fadali, "Smart meter based short-term load forecasting for residential customeers, "In proceedings of North American Power Symposium (NAPS), pp. 1–5, August 2011.

[8] S. Wang, L. Cui, J. Que, D. Choi, X. Jiang, S. Cheng, and L. Xie, "A randomized response model for privacy preserving smart metering," IEEE Transactions on Smart Grid, voll.3, no. 3, pp. 1317–1324, September 2012.

[9] V. Cagri Gungor, Dilan Sahin, Taskin Kocak, Salih Ergut, Concettina Buccella, Carlo Cecati, and Gerhand P. Hancke, "Smart grid technologies: communication technologies and standards," IEEE Transactions on Industrial Informatics, vol. 7, no. 4, pp. 529–542, November 2011.

[10] Ye Yan, Yi Qian, Hamid Sharif, and David Tipper, "A survey on smart grid communication infrastructures: motivations, requirements, and challenges," IEEE Communications Surveys & Tutorials, vol. 15, no. 1, pp. 5–20, February 2013.

[11] David Bernaudo, Larry Barto, Shannon Spizzirri, and Terry Penn, "SmartGrid/AEIC AMI interoperability standard guidelines for ANSI C12.19 / IEEE 1377 / MC 12.19 end device communications and supporting enterprise devices, networks and related accessories," The Association of Edison Illuminating Companies, Meter and Service Technical Committee report, ver. 2, May 2010.

[12] E. E. Queen, "Adiscussion of smart meters and RF exposure issues," Edison Electric Institute, Washington, C.C, A Joint Project of the EEI and AEIC MEter Committee, March 2011.

[13] B. Karimi, V. Namboodiri, and M. Jadliwala, "On the scalable collection of metering data in smart grid through message concatenation," In preceedings of IEEE International Conference on

Smart Grid Communications, pp. 318–323, October 2013.

[14] B. Karimi, V. Namboodiri, and M. Jadliwala, "Scalable meter data collection in smart grids through message concatenation," IEEE Transactions on Smart Grid Communications, vol. 6, no. 4, pp. 1697–1706, July 2015.

[15] PRIME Alliance, "PRIME: powerline intelligent metering evolution," Available from: http://www.prime-alliance.org [Accessed: 2016-5-2].

[16] H. Li, Z. Han, L. Lai, R. Qiu, and D. Yang, "Efficient and reliable multiple access for advanced metering in future smart grid," In proceedings of IEEE International Conference on Smart Grid Communications, pp. 440–444, October 2011.

[17] Sun-Yuan Hsieh, Dr. Hemant Kumar Gianey, Chih-Wei Hsu, Cheng-Han Yeh, and Geng-Hua Zhang, "Novel Scheme for Reducing Communication Data Traffic in Advanced Metering Infrastructure Networks," Submission Committing.

[18] K. Barr and K. Asanovic, "Energy Aware Lossless Data Compression," in Proceedings of 1st International Conference on Mobile systems, Applications, and Services (MobiSys), pp. 231–242, 2003.

[19] M. Ringwelski, C. Renner, A. Reinhardt, A. Weigel, V. Turau,"The Hitchhiker's Guide to Choosing the Compression Algorithm for Your Smart Meter Data," 2nd IEEE ENERGYCON conference & Exhibition, 2012.

[20] Sun-Yuan Hsieh, and Geng-Hua Zhang, "A Strategy for Real-Time Energy Data Compression to Reduce Data Traffic based on Smart Grid AMI Networks," Submission Committing.