



Article Cost-Effective Data Aggregation Method for Smart Grid

Hsi-Chou Hsu¹, Shi-Ren Zhuang² and Yung-Fa Huang^{3,*}

- ¹ Department of Computer and Communication, National Pingtung University, Pingtung 90004, Taiwan; chrishsu@mail.nptu.edu.tw
- ² Department of Computer Science and Information Engineering, National Pingtung University, Pingtung 90004, Taiwan; renrenjuang@gmail.com
- ³ Department of Information and Communication Engineering, Chaoyang University of Technology, Taichung 313310, Taiwan
- * Correspondence: yfahuang@cyut.edu.tw; Tel.: +886-4-23323000

Abstract: Finding a more efficient use of energy is an important problem that needs attention. Compared with the traditional power grid, a smart grid can monitor users' electricity situation and electricity consumption instantly. However, it involves many problems of deploying network equipment. Consequently, it is vital to promote smart grids by collecting data from smart meters efficiently and keeping costs low. In this article, we propose a two-stage method of data collection for smart grids. The main contribution of this paper is to lower the number of data aggregation points (DAPs) so that the cost can be reduced. By using the K-means method, an entire smart grid can be divided into many smaller parts. In addition, the needs of transmitting and receiving data in the entire smart grid can be met by installing the least number of DAPs. Finally, the simulations show that the proposed two-stage method of data collection can use fewer DAPs to collect data than other methods which use one-stage methods, so the proposed scheme is more cost-effective.





check for

Citation: Hsu, H.-C.; Zhuang, S.-R.; Huang, Y.-F. Cost-Effective Data Aggregation Method for Smart Grid. *Electronics* 2021, 10, 2911. https:// doi.org/10.3390/electronics10232911

Academic Editor: Tadeusz A. Wysocki

Received: 31 October 2021 Accepted: 23 November 2021 Published: 24 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). 1. Introduction

Because of the effects of climate change, more people have started focusing on the issues of energy conservation, carbon reduction, efficient energy use, and so on. In recent years, with the exuberant development of network, the concept of the smart grid is promoted vigorously. By promoting this, people hope to use and control energy efficiently; moreover, the traditional power grid can be replaced with smart grids. Compared with the traditional power grid, a smart grid can monitor users' electricity situation and electricity consumption at a faster rate. Moreover, thanks to the availability of bi-directional communications, a smart grid can raise the efficiency of electricity and the reliability of the power grid. The infrastructure which supports bi-directional communication, called advanced metering infrastructure (AMI), was regarded as the fundamental structure of the smart grid [1]. To construct AMI, it is necessary to install smart meters (SMs) at home and all the electrical appliances at home need to be installed with sensors so that they can connect with SMs by a network. For example, in recent years, people often use wireless sensor networks which use Bluetooth, Wi-Fi and ZigBee to solve communication problems. Furthermore, users can monitor the electricity situations and transmit data using SMs to the control center instantly. As a result, the control center can immediately obtain the information on the electricity situations from each family, and users can also learn about the situations of every electrical appliance so that they can begin with allocating the utilization of electricity.

According to Greentech Media (GTM) Research, a smart grid can be divided into a power layer, a communications layer, and a smart grid application layer [2]. Besides, according to the communication range, AMI can be divided into three categories which are home area network (HAN), neighborhood area network (NAN), and wide area network (WAN). HAN contains SMs and other smart devices and often uses lower-cost communication techniques, such as Wi-Fi, ZigBee, PLC, Z Wave, and so on. NAN is between HAN and WAN. In NAN, the data from HAN will be collected and then be transmitted to WAN. Because of the need to handle bigger data in NAN, it uses wired networks, such as PLC, optical networks, or wireless networks, which have high data rates such as WiMAX. As for WAN, it connects several NANs and transmits data to the control center. In WAN, long-range communication techniques are often used, such as 3G, LTE, and LoRa.

In the smart grid system, when a great number of SMs transmit data at the same time, network congestion or collision will happen and even delay the entire system. As a result, communication quality is quite important for the smart grid. Researchers have proposed schemes to reduce delay in WAN [3-5]. A method was proposed to reduce delay in NAN by using the concept of data aggregation and defining a role, called data aggregation point (DAP), to receive all data from SMs [6]. However, choosing proper positions to install DAPs and be able to reduce delay at the same time is an important issue. We usually use a high-speed wired network to transmit data from DAPs to the control center, so it needs extra cost to install DAPs. Therefore, it is important to use a smaller number of DAPs [7,8]. In [9], applications of different optimization techniques are summarized. In [10], the methods which can reduce the cost of DAP in wired networks and wireless networks were proposed, while in [11], a cost-effective method to install DAP by using a utility pole was proposed. In [12], a heuristic algorithm using K-means was proposed to split the original set covering problem (SCP) into smaller ones that are optimally solved. Authors in [13] have formulated a constrained optimization problem, called cost minimization DAP placement (CMDP) to minimize DAP installation cost while satisfying communication QoS requirements. Because CMDP is proven NP-hard, a heuristic algorithm based on K-means was proposed to produce sub-optimal solution in reasonable time. In [14], a heuristic threephase algorithm was presented for the optimal DAP placement in a multi-hop routing scenario, where SMs can serve as small relay devices so that communications among SMs can be leveraged to reduce the overall communication cost. K-means is a popular clustering algorithm that has advantages of high speed, but it suffers from the problem of local optimal. In [15], a clustering algorithm based on K-medoids was proposed. Although the mean square error for K-medoids is lesser than K-means, K-medoids is lacking in performance [16]. Owing to a bad choice of initial centroid locations and trapping into the local optimum easily, many articles have proposed to improve K-means by the particle swarm optimization (PSO) algorithm [17–21].

The aim of this paper is to reduce the number of DAPs. A role called cluster head (CH) is added to collect data from SMs and then transmit it to DAPs. We propose a new cluster algorithm to solve the problems of installing DAPs and choosing CHs. First, we divided SMs into several groups by using clustering and subsequently selected proper SMs from every cluster to be CH. In every cluster, all SMs could transmit data to DAPs through CHs.

The remainder of this paper is organized as follows. Section 2 presents the signal model. Section 3 describes the proposed algorithm. Section 4 illustrates the evaluation results. Finally, Section 5 draws the main conclusions.

2. Signal Model

Here, a smart grid is divided into four categories, including SM, CH, DAP, and control center. Their main functions, as shown in Figure 1, are:

- SM: SM is installed in every home and the main function is to monitor the electricity situations in every family.
- CH: CH is a SM that is chosen from every cluster. Its main function is to collect data from all SMs in every cluster and transmit it to DAP.
- DAP: the main function of DAP is to collect data from CHs and then transmit it to the control center.
- Control center: the control center is used to receive data from DAPs and monitor the entire smart grid.

As shown in Figure 1, every house represents a family and has an SM in every house. First, we chose proper positions to install DAPs and then divided SMs into several clusters. In any area, SMs were divided into three clusters. Then, a proper SM was chosen to be a CH in every cluster. In this paper, we focus on the problems such as the transmission of data from SMs to CHs and CHs to DAPs. Our goal is to use the least number of DAPs and an optimal number of CHs.



Figure 1. The smart grid system model.

The required signal-to-noise ratio (SNR) when CHs receive data from SMs, and DAPs receive data from CHs can be expressed as:

$$\gamma_{SM,CH,DAP} = \frac{P_t \times FSPL}{\sigma} \tag{1}$$

where P_t is the transmission power, *FSPL* is free-space path loss. The *FSPL* in dB can be expressed as:

$$FSPL = 10 \times \log_{10}\left(\left(\frac{4\pi \times d_{snr} \times f}{c}\right)^2\right)$$
(2)

where *f* is frequency, *c* is the speed of light, and d_{snr} is the threshold for distance. However, the supposed environment is not in free space, so the received signal power and the distance are *n*th power fading. Different values of *n* are presented in Table 1. In this paper, the value of *n* is considered as 3, so the function can be rewritten as:

$$FSPL = 20 \, \log_{10}(f) + 20 \, \log_{10}\left(\frac{4\pi}{c}\right) + 30 \, \log_{10}(d_{snr}) \tag{3}$$

Besides, σ is the thermal noise power in Watts which can be calculated by the formula as:

$$\sigma = K \cdot T \cdot B \tag{4}$$

where *K* is Boltzmann constant (at about 1.38×10^{-23}); and *T* is the temperature in degree K (°K). Because γ is the least SNR, $\gamma_{SM,CH,DAP}$ should be larger than or equal to γ . As a result, the SNR threshold is determined as follows:

$$\gamma_{SM,CH,DAP} \ge \gamma$$
 (5)

At last, the corresponding d_{snr} in different SNR can be determined by Equations (1), (3), and (5). According to [13] and (4), assuming that packet error probability (E) is 0.01, packet length (L) is 1800 bits, bit per second (R) is 2 Mbps and bandwidth (B) is 1 MHz. In this case, the γ was estimated to be about 46 dB. Because γ is the least SNR, the $\gamma_{SM,CH,DAP}$ must be large or equal to γ . For example, giving n = 3, $P_t = 30$ dBm, T = 298.15 °K (about 25 °C), f = 1 GHz, and c is 3×10^5 km, the d_{snr} was measured to be about 1.53 km while γ is 46 dB.

Environments	Path Attenuation Index (n)
Free space	2
Metropolitan area	3~5
Building	4~6
Factory	2~3

Table 1. Path loss exponent in various environment.

3. Two-Stage Clustering Algorithm

In this section, the proposed scheme is described in detail. We aimed to aggregate data efficiently and reduce network delay by using the least number of DAPs and the optimal number of CHs. Using K-means to cluster SMs, it just groups SMs into several clusters and searches for the centroid in each cluster. However, it does not consider whether some SMs is too far from the DAP or not. When wireless communication techniques are used to transmit data, there are different distance restrictions due to transmission power and channel environments. Thus, a distance threshold (d_{snr}) is added to constrict the farthest distance of SMs to CHs and CHs to DAPs. In this paper, we assumed that the d_{snr} of SMs to CHs and CHs to DAPs were the same.

In smart grids, how to choose the number of DAPs and the locations to install DAPs is a non-deterministic polynomial (NP) problem [13]. To reduce the complexity, in this work it is used the K-means algorithm to partition SMs into several smaller parts.

For the K-means algorithm, Euclidean sum of squares is defined as fitness function:

$$Fitfunction = \sum_{i=1}^{N} [(X_i - Cen(j))]^2$$
(6)

where X_i is the coordinate of *i*th SM and Cen(j) is the initial centroids when dividing into *j* clusters. We expected to find the centroid, which means that the sum of squared shortest distances of all SMs to the centroid was minimized.

A two-stage clustering algorithm to determine the least number of DAPs and an optimal number of CHs is proposed. At the first stage, the required amount of DAPs, denoted by *K*, is determined and a single DAP (i.e., K = 1) is tested at the beginning. At this stage, all SMs were divided into *K* clusters, and let *K* DAPs be placed at the *K* centroids. At the second stage, in each of the *K* clusters, SMs were divided into smaller sub-clusters, while a SM was chosen as the CH in each sub-cluster, until the distance from all SMs to the CH in each sub-cluster and all CHs to the DAP in each of the *K* clusters were shorter than the distance threshold (d_{snr}). The proposed clustering algorithm is described as below:

- (1) Set K = 1, regard all SMs as one cluster.
- (2) Use the K-means algorithm to find out the least fitness as well as the optimal corresponding *K* centroids and install a DAP in each centroid, as shown in Figure 2.
- (3) Clustering all SMs covered in each DAP.
 - a. When the clustering result met the distance threshold (d_{snr}) , the optimal number of clusters is determined. Therefore, the algorithm is stopped.
 - b. If no candidate CH for any SM can meet the distance threshold, SMs are continually grouped into smaller clusters until all SMs become CHs. As a result, one DAP (K = K + 1) is added and return to Step (2).



Figure 2. Coverage of all SMs in the area under study with one DAP (the triangle symbol), K = 1.

Finally, the number of DAPs needed to be installed in this area and the number of CHs needed in every DAP are determined. As shown in Figure 3, the optimal number of clusters was 6, so 6 CHs (blue dots in the figure) could be obtained. However, on adding the other one DAP, as in Figure 2 (i.e., K = 2), the obtained clustering result is shown in Figure 4.



Figure 3. Selection of the optimal number of clusters for all SMs with one DAP.



Figure 4. Addition of another DAP to meet the distance threshold, *K* = 2.

The K-means algorithm can also determine the centroid in every cluster, and therefore, we choose the SM closest to centroid as CH in every cluster because a centroid is a point

where the sum of squares of the shortest distance to every SM is minimum. Although SMs are not mobile devices, new SMs might be added to a cluster, and other SMs might be removed from a cluster. As a result, if that happens, the cluster should be reconfigured.

When data are transmitted from transmitter to receiver, they must be affected by the environment or by passing too many hops so that there may be a delay. Delay may be of two types, propagation delay, and media access delay, as described below:

- (1) Propagation delay: this delay time is the duration when a packet transmits from transmitter to receiver. This time is affected by the distance, and the farther the distance is, the higher the delay is. Assuming that the distance between transmitter and receiver is d and transmission rate is v, so, the time it costs is d/v. However, because the speed of transmission equals to speed of light, the delay caused by propagation delay is very low.
- (2) Media access delay: this delay time is the duration when a packet transmits successfully. If a packet transmits unsuccessfully, it is retransmitted until it is successful. Suppose that the media access delay D_m is a packet that transmits once successfully and if a packet transmits unsuccessfully for the first time, then the delay time when it transmits for the second time is $2D_m$. Hence, if this packet transmits successfully at the *n*th time, the delay time is $n \times D_m$. As a result, the greater number of times the packet is retransmitted, the higher the delay is. The media access delay is the primary delay when a packet transmits.

The proposed two-stage clustering algorithm utilizes CH to receive data from SMs and transmit it to DAP, so it can reduce the number of DAPs with the same distance threshold. Compared with the method presented in [13], the proposed method causes a greater delay because data are transmitted through more hops. Suppose *D* is the total delay when a packet transmits from transmitter to receiver, D_p is propagation delay, D_m is media access delay, assuming $D_p + D_m \cong D_m$, where D_p is significantly short and can be ignored. Here, *D* can be calculated by the packet error rate (PER). Thus, *D* can be expressed as:

$$D = (1 - P) \times D_m + P \times (1 - P) \times 2D_m + P^2 \times (1 - P) \times 3D_m + \dots$$
(7)

where *P* is the PER. Then, Equation (7) can be rewritten by multiplying *P*, it can be expressed as:

$$P \times D = P \times (1-P) \times D_m + P^2 \times (1-P) \times 2D_m + P^3 \times (1-P) \times 3D_m + \dots$$
(8)

Subtracting Equation (7) from (8) and we can obtain:

$$(1-P) \times D = (1-P) \times D_m + P \times (1-P) \times D_m + P^2 \times (1-P) \times D_m + \dots$$
 (9)

Finally, Equation (9) can be simplified as below:

$$D = D_m \sum_{i=0}^{\infty} (P)^i$$

= $\frac{D_m}{1-P}$ (10)

In the proposed method, SMs will transmit data to CHs, and then CHs will transmit these data to DAPs. Assuming that P_{ch} and P_{dap} are the PER which are when SMs transmit data to CH and the PER when CHs transmit data to DAP, respectively. Finally, we can calculate the delay by PER. Suppose $\frac{D_m}{1-P_{ch}}$ and $\frac{D_m}{1-P_{dap}}$ are expressed as the delay when SMs transmit data through two hops. So, the total delay can be expressed as:

$$D = \frac{D_m}{1 - P_{dap}} + \frac{D_m}{1 - P_{ch}}$$
(11)

4. Evaluation Results

To evaluate the performance in the delay time (*D*) and required amount of DAPs, we compared our two-stage data aggregation method with the method presented in [13]. In this paper, MATLAB is used for simulation. To verify that our proposed scheme can be applied to various distributions of SMs, the command 'randn' in MATLAB is used to generate locations for SMs, which are shown as a normal distribution. As a result, most SMs concentrated in the middle area. Figure 5 shows the simulation result of SMs' distribution. A total of 100 SMs are produced, which were randomly distributed in $40 \times 40 \text{ km}^2$ area. Most SMs were concentrated in the middle and were only seldom SMs scattered in faraway areas.



Figure 5. The result of SMs' distribution.

Assuming PER to be 0.01, packet size L of 1800 bits, transmission rate R of 2 Mbps, bandwidth B of 1 MHz, the SNR threshold γ was 46 dB, and distance threshold d_{snr} was 7.1 km. Figure 6 shows the result of clustering. The triangle icon is expressed as DAP, the bigger circle icon is expressed as CH in every cluster, and the SMs which have different colors are expressed the SMs belong to different clusters.

Because K-means suffers from the problem of local optimal, the combination of PSO and K-means is used to improve the K-means clustering problem [18]. Figure 7 shows the comparison of the sum of the squares of the distance between K-means and K-means improved by PSO, denoted as "PSO + K-means". It has a lower sum of the distance while K-means is improved by PSO. A lower sum of the distance means that the PSO + K-means scheme yields a better clustering result.

The comparison of the required amount of DAP for different numbers of SMs is shown in Figure 8. The dotted line is the compared, one-stage method presented earlier [13]. In the one-stage method, SMs transmit data to DAPs directly. When the distance between SM and DAP is longer than the distance threshold (d_{snr}), it will add one DAP and restart clustering until the distance between each SM and a DAP is shorter than the distance threshold. As shown in Figure 8, the horizontal axis is expressed as the number of SMs, from 50 to 350, and the vertical axis is expressed as the amount of DAP. The distance threshold is 7.1 km. With the increase in the number of SMs, the amount of DAP necessary to install also increase. Obviously, the amount of DAP necessary in this study was less than that in the earlier proposed method [13].



Figure 6. The result of clustering.



Figure 7. Comparison between K-means and PSO + K-means algorithms.

Figure 9 shows the relationship of distance threshold (d_{snr}) according to different SNR thresholds (γ). Here, γ is about 30 dB where the corresponding d_{snr} is about 24 km; γ is about 40 dB where the corresponding d_{snr} is about 11 km; and when γ is about 50 dB, the corresponding d_{snr} is about 5 km. This is because the received power is lost when the distance increases. As a result, to maintain higher SNR, the distance is not too far.



Figure 8. The comparison of the required number of DAPs with different numbers of SMs ($d_{snr} = 7.1$ km).



Figure 9. The relationship of distance threshold (d_{snr}) according to different SNR thresholds (γ) .

Figure 10 shows the required number of DAPs in different γ . In this simulation, 100 SMs are used, and the value of γ was from 45 dB to 60 dB. According to this figure, when the value of γ raises, the required number of DAPs also increases. According to Figure 9, the higher the γ is, the smaller is the d_{snr} . As a result, when d_{snr} is smaller, the distance between every DAP and CHs as well as the distance between every CH and SMs becomes shorter. While d_{snr} becomes smaller, the range every DAP can cover is also reduced and the required number of DAPs increases. The dotted line is the method proposed by the earlier study [13] and the solid line represents our method. The required number of DAPs with the proposed method was fewer than that in the earlier proposed method [13].



Figure 10. The relationship of the required number of DAPs with SNR threshold (γ).

The average delay is calculated by Equation (11). The packet error rate can be calculated by bit error rate (BER), as below:

$$P = 1 - (1 - b)^L \tag{12}$$

where *b* is BER; and *L* is the number of bits per packet. By using Formula (13) mentioned in [22], the BER can be calculated by a given SNR. In this paper, the SNR was calculated by Equations (1) and (3) with a known distance. Therefore, P_{ch} was derived by the distance between an SM and its CH, and P_{dap} was derived by the distance between a CH and its DAP.

Figure 11 shows the comparison of delay with different numbers of SMs. The horizontal axis is the number of SMs, and the vertical axis represents the average delay. The average delay for the compared method and our proposed method was calculated using Equations (10) and (11), respectively. According to this figure, it is shown that the delay in the proposed method was higher than the compared method [13]. However, in 50 SMs, the average number of DAPs in the proposed method was 6, while that of the compared method is 6.5. Furthermore, when the number of SMs increased to 150, the average amount of DAPs in the proposed method was 7, and that of the compared method was 11. Therefore, it is concluded that although the delay in our proposed method is higher, the average number of DAPs is fewer than that of the compared scheme. Moreover, while the number of DAPs increased, the average delay with the compared method was lowered because the distance between every DAP and SMs was shortened.

Figure 12 shows the relationship of average delay with our proposed method and the compared method [13] with the same number of DAPs, while 50 SMs are used. With a single DAP, the average delay with the compared method was much higher than that in our proposed method, because the compared method divides all SMs into one cluster. As a result, the average distance between the DAP and all SMs is longer than that in our proposed method. Significantly, the longer distance brings about a higher PER. As shown in Equation (10), the time delay depends on the PER. However, with the number of DAPs larger than three, the average delay was worse with our proposed method than the compared method. This may be because if the SMs in one area has good results of clustering, too many hops cause higher delay.



Figure 11. The comparison of delay in different numbers of SMs.



Figure 12. The relationships of average delay (D) in cases of different numbers of DAPs.

5. Conclusions

Owing to the high cost of installing a DAP, our aim was to use as few DAPs as possible. In this study, we used K-means clustering to divide SMs into several clusters and used a two-stage method to aggregate data. The proposed two-stage clustering algorithm utilizes CH to receive data from SMs and transmit it to DAP, so it can reduce the number of DAPs. However, because our method is with a two-stage mechanism, an SM needed two hops to transmit its data through a CH to a DAP. It may cause a greater delay because data are transmitted through two hops, but the per-hop delay is reduced because the transmission distance is shorter. To fairly compare average delay of our proposed scheme to the compared scheme, the delay time was formulated as Equations (10) and (11). As a result, the required number of DAPs with our proposed method was significantly

fewer although the average delay with our proposed method was a little higher than that with the compared method. Thus, our proposed method is more cost-effective than the compared method.

Author Contributions: Conceptualization, H.-C.H. and Y.-F.H.; data curation, H.-C.H. and S.-R.Z.; investigation, H.-C.H. and S.-R.Z.; methodology, H.-C.H. and Y.-F.H.; software, S.-R.Z.; validation, H.-C.H. and S.-R.Z.; writing—original draft, H.-C.H., S.-R.Z. and Y.-F.H.; writing—review and editing, H.-C.H. and Y.-F.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Ministry of Science and Technology (MOST), R.O.C. grant number MOST 110-2221-E-324-008.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Gungor, V.C.; Sahin, D.; Kocak, T.; Ergut, S.; Buccella, C.; Cecati, C.; Hancke, G.P. Smart Grid Technologies: Communication Technologies and Standards. *IEEE Trans. Ind. Inform.* **2011**, *7*, 529–539. [CrossRef]
- Leeds, D.J. The Smart Grid in 2010: Market Segments, Applications and Industry Players; GTM Research: Boston, MA, USA, 2009; pp. 1–145.
- 3. Briscoe, B.; Brunstrom, A.; Petlund, A.; Hayes, D.; Ross, D.; Tsang, I.-J.; Gjessing, S.; Fairhurst, G.; Griwodz, C.; Welzl, M. Reducing Internet Latency: A Survey of Techniques and Their Merits. *IEEE Commun. Surv. Tutor* **2014**, *18*, 2149–2196. [CrossRef]
- 4. Wang, G.; Ren, Y.; Li, J. An effective approach to alleviating the challenges of transmission control protocol. *IET Commun.* **2014**, *8*, 860–869. [CrossRef]
- Wang, G.; Wu, Y.; Dou, K.; Ren, Y.; Li, J. AppTCP: The design and evaluation of application-based TCP for e-VLBI in fast long distance networks. *Futur. Gener. Comput. Syst.* 2014, 39, 67–74. [CrossRef]
- Wang, G.; Zhao, Y.; Huang, J.; Winter, R.M. On the Data Aggregation Point Placement in Smart Meter Networks. In Proceedings of the 2017 26th International Conference on Computer Communication and Networks (ICCCN), Vancouver, BC, Canada, 31 July–3 August 2017; pp. 1–6.
- Rolim, G.; Passos, D.; Albuquerque, C.; Moraes, I.; Carrano, R.; Sousa, C.; Bettiol, A.; Passos, L.; Homma, R.; Andrade, R.; et al. Scalability evaluation of the data aggregator positioning problem in smart grids. In Proceedings of the 2016 IEEE PES Transmission & Distribution Conference and Exposition-Latin America (PES T&D-LA), Morelia, Mexico, 20–24 September 2016; pp. 1–6.
- Rolim, G.; Passos, D.; Moraes, I.; Albuquerque, C. Modelling the Data Aggregator Positioning Problem in Smart Grids. In Proceedings of the 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, Liverpool, UK, 26–28 October 2015; pp. 632–639.
- 9. Islam, M.; Nagrial, M.; Rizk, J.; Hellany, A. Review of Application of Optimization Techniques in Smart Grids. In Proceedings of the 2018 2nd International Conference on Electrical Engineering (EECon), Colombo, Sri Lanka, 28 September 2018; pp. 99–104.
- Huang, X.; Wang, S.; Wang, C. Aggregation Points Planning for Smart Grid Communications: Wired and Wireless Cases. In Proceedings of the IEEE Global Communications Conference (GLOBECOM), San Diego, CA, USA, 6–10 December 2015; pp. 1–6.
- 11. Aalamifar, F.; Shirazi, G.N.; Noori, M.; Lampe, L. Cost-efficient data aggregation point placement for advanced metering infrastructure. In Proceedings of the 2014 IEEE International Conference on Smart Grid Communications (SmartGridComm), Venice, Italy, 3–6 November 2014; pp. 344–349.
- 12. Rolim, G.; Passos, D.; Albuquerque, C.; Moraes, I. MOSKOU: A Heuristic for Data Aggregator Positioning in Smart Grids. *IEEE Trans. Smart Grid* 2017, 9, 6206–6213. [CrossRef]
- 13. Kong, P.Y. Cost Efficient Data Aggregation Point Placement With Interdependent Communication and Power Networks in Smart Grid. *IEEE Trans. Smart Grid* 2017, 10, 74–83. [CrossRef]
- 14. Lang, A.; Wang, Y.; Feng, C.; Stai, E.; Hug, G. Data Aggregation Point Placement for Smart Meters in the Smart Grid (Early Access). *IEEE Trans. Smart Grid* 2021, 1. [CrossRef]
- 15. Gallardo, J.L.; Ahmed, M.A.; Jara, N. Clustering Algorithm-Based Network Planning for Advanced Metering Infrastructure in Smart Grid. *IEEE Access* 2021, *9*, 48992–49006. [CrossRef]
- Arbin, N.; Suhaimi, N.S.; Mokhtar, N.Z.; Othman, Z. Comparative Analysis between K-Means and K-Medoids for Statistical Clustering. In Proceedings of the 2015 3rd International Conference on Artificial Intelligence, Modelling and Simulation (AIMS), Kota Kinabalu, Malaysia, 2–4 December 2015; pp. 117–121.
- 17. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the IEEE International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.
- Tan, L. A Clustering K-means Algorithm Based on Improved PSO Algorithm. In Proceedings of the 2015 Fifth International Conference on Communication Systems and Network Technologies, Gwalior, India, 4–6 April 2015; pp. 940–944.

- Sheta, A.F.; Solaiman, B. Evolving clustering algorithms for wireless sensor networks with various radiation patterns to reduce energy consumption. In Proceedings of the 2015 Science and Information Conference (SAI), London, UK, 28–30 July 2015; pp. 1037–1045.
- Rani, A.; Parthiban, L. Improved particle swarm optimization and K-means clustering algorithm for news article. In Proceedings
 of the IET Chennai Fourth International Conference on Sustainable Energy and Intelligent Systems (SEISCON 2013); Institution
 of Engineering and Technology (IET), Chennai, India, 12–14 December 2013; pp. 412–420.
- 21. Salma, M.U. PSO based fast K-means algorithm for feature selection from high dimensional medical data set. In Proceedings of the 2016 10th International Conference on Intelligent Systems and Control (ISCO), Coimbatore, India, 7–8 June 2016; pp. 1–6.
- 22. Elshabrawy, T.; Robert, J. Closed-Form Approximation of LoRa Modulation BER Performance. *IEEE Commun. Lett.* **2018**, *22*, 1778–1781. [CrossRef]