03

A solution method for data aggregator allocation in smart grids through a modelling

Sami Nasser Lauar

Coordenadoria de Engenharia Elétrica – Campus Vitória
Instituto Federal do Espírito Santo (Ifes), Vitória, ES

Tainã Ribeiro de Oliveira

Coordenadoria de Engenharia Elétrica – Campus Vitória
Instituto Federal do Espírito Santo (Ifes), Vitória, ES

Mário Mestria

Coordenadoria de Engenharia Elétrica – Campus Vitória
Programa de Pós-Graduação em Tecnologias Sustentáveis
Instituto Federal do Espírito Santo (Ifes), Vitória, ES

DOI: 10.47573/aya.5379.2.66.3

ABSTRACT

A model to solve data aggregator allocation problem in smart grids is presented in this paper. We use a model based on a set covering from the literature plus dispersion constraints among data aggregators. These constraints are incorporated in the proposed model. We applied an algorithm in a way that uses two steps to solve this model. Several instances were used in a simulation way with different smart grid scenarios to test proposed model. The results were conducted with different real-world parameters, such as data aggregator costs, a number of smart meters, and number of the data aggregators. The computational results reached the best performance using medium and small size instances. The ratio between the number of smart meters and the number of allocated aggregators was in the range of 28-36.25 according to the literature. The test of all instances demonstrated that the algorithm yielded optimal solutions in short computational times, given in seconds. Considering that the data aggregator allocation problem is NP-hard, in future work, we propose the use of metaheuristics to reach high-quality solutions in large size instances.

Keywords: data aggregator. smart meter. smart grid. modelling. combinatorial optimization problem. set covering problem.

INTRODUCTION

There are challenges in reducing technical losses from the Joule effect in traditional electric distribution systems. The electric power distribution utilities must offer customers quality and reliability in the use of energy. To overcome these problems, Cardenas *et al.* [1] describe a new concept called smart grids (more details next section). The concept of a smart grid covers a range of research challenges, such as distributed control, fault detection, and data communication using the smart meters and data aggregators.

In this new concept of smart grids, data are sent in real time at short and constant periodic intervals for one or more data aggregators over a wireless network infrastructure. The aggregators will transmit data that have been collected from a region containing the smart meters for the electric power distribution utilities. Therefore, a medium-distance data communication network infrastructure emerges.

The smart meters are installed in the final consumers and the data aggregators, in general, are on top of the distribution utility poles. According to [2] some communication protocols can be used to communicate between smart meters and data aggregators within a limited area, from short to medium range communication as the IEEE 802.11 and IEEE 802.15.4 (ZigBee). Moreover, the authors in [2] claimed that in the smart grids are regularly used for long range communication, the protocols as GPRS, 3G, 4G, 5G or IEEE 802.16 (WiMax), between data aggregators and the electric power distribution utilities.

The current technologies to communicate aggregators and the meters limit the positing of data aggregators along electric power distribution network due decreasing signal propagation [3]. Thus, allocation of the data aggregators in smart grids brings in a hard problem. We proposed a formulation based on the set covering problem with the constraints that the aggregators may not be extended over a large area.

In the literature, there is a lack in the modeling that solves data aggregator allocation problem in smart grids. In the models is not incorporating the constraints in distributions among the aggregators, called here dispersion. Another shortcoming in the literature is referred about testing smart grid electrical parameters as: the acquisition cost of aggregators for the utilities that are settled by simple way as single cost and constant number of smart meters in testing scenarios.

In this paper, a new mathematical optimization model is presented based on electrical parameters of smart grids, incorporating the constraints of dispersion. With other words, a new model, with the above described restrictions, makes it possible to reduce redundancy and avoid delays in the communication of data packages. The model also addresses the quality of service by reducing delay and jitter in the transfer of data packets from the smart meters to the aggregators. Thus, it is introduced a limiting the distance between smart meter and aggregator. Consequently, this model avoids packet loss and collisions, reduces network congestion, and minimizes signal interference.

Beyond, in the model the electrical parameters can be tested including different aggregator costs, dispersion among the aggregators, and flexible numbers of smart meters. However, the model does not address ensuring data security that concerns customer privacy of confidential information about their energy consumption habits.

In the smart grids there are different types of smart meters that they can communicate with the aggregators impacting significantly the total cost in the smart grids. In our mathematical optimization model, the bandwidth requirements of each smart meter, it is assumed to be identical. We call each smart meter of a standard smart meter. Further, a location problem of data aggregators in smart grids can incorporate hybrid network topology (e.g., urban, suburban, and rural) into the optimization model. But, in our mathematical optimization model is designed in an urban scenario without interference source with several standard smart meters.

The proposed mathematical model in this work will solve the data aggregator allocation problem in the smart grids. This model is partially based on the set covering problem. The standard smart meters can send data to the nearest aggregators. Thus, the constraints of dispersion among the aggregators are incorporated into the proposed mathematical model. So, the constraints minimize the number of standard smart meters to be assigned to the aggregators and set up a minimum distance among the aggregators.

To solve the model, in this paper was proposed an algorithm with two steps for allocation problem in the smart grids: first, we allocate the smart meters to the nearest aggregators (that can be one or more aggregator); second, and then to reduce dispersion, a simple, but effective branch-and-bound (B&B) exact method that minimizes the cost acquisition of aggregators for the utilities is applied. A goal of this paper is also rationalizing the numbers of standard smart meters and allocated aggregators. We highlight that the limitations of our mathematical optimization model are: data transmitted security is not addressed; standard smart meter is assumed; urban scenario without interference source is taking on.

This model with dispersion proved to be robust in computational tests using an algorithm (to be involved in the two steps) with dissimilar parameters. Simulation results showed that the proposed exact method for new optimization model was capable to find several high-quality optimal solutions in a matter of seconds for medium-size instances.

In summary, the main contributions of this work, which are not considered in previously published works, can be listed as follows.

- A novel model is proposed based on smart grid electrical parameters.
- The elements needed in planning and implementing the model in smart grids are described.
- The minimum acquisition cost of aggregators for the utilities is optimized and allocated.
- The numbers of standard smart meters and allocated aggregators are rationalized.
- The performance of the proposed algorithm provides benefits for the electric utilities, consumers, and smart cities.

The paper is structured as follows: Section 2 is described about smart grids. Section 3 provides an optimization method for the data aggregator allocation problem in smart grids. In Section 4, the computational results are provided. In Section 5, discussion is presented. In the last section, conclusions are described together with future work.

SMART GRIDS

The smart power grid covers other issues as forecasting, grid stability, and demand response. Therefore, smart grids constitute a multidisciplinary area that presents many challenges [4]. Smart grids are based on efficient energy distribution using state-of-the-art electronics; use of renewable energy resources to feed the electric distribution systems; active participation of consumers in all chains of generation and distribution of electricity; presentation for customers of consumption in real time through smart meters; and, in the future, using the energy from electric vehicle batteries to store and distribute energy for the power electric systems.

Boccardo *et al.* [5] discussed sustainability in energy production that could entail a complete transformation in consumer habits: The consumers could choose to schedule their consumption of energy when the sustainability index is most favorable; customers would also have incentive to use alternative energy sources. That enables electric power distribution utilities to promote the adoption of renewable energy sources (e.g., wind power or solar power). In such cases, for example, energy regulators such as smart meters could impose the best pricing strategy aligned with cleaner energy usage.

Smart grids are in a rapid development phase in the digital transformation of electric power distribution utilities. In fact, in future, everything will be connected through so-called smart zones: smart water grids, smart traffic systems, smart manufacturing, smart buildings, and smart cities. Smart cities use technology to improve the transport network, reduce traffic, increase mobility, improve energy efficiency, and promote urban sustainability.

According to Lopes *et al.* [6], smart cities aim to improve people's quality of life by providing shared information and allow innovations in several commercial and industrial areas. In this sense, smart cities can use data sharing through the Internet in the context of the Internet of Things [7].

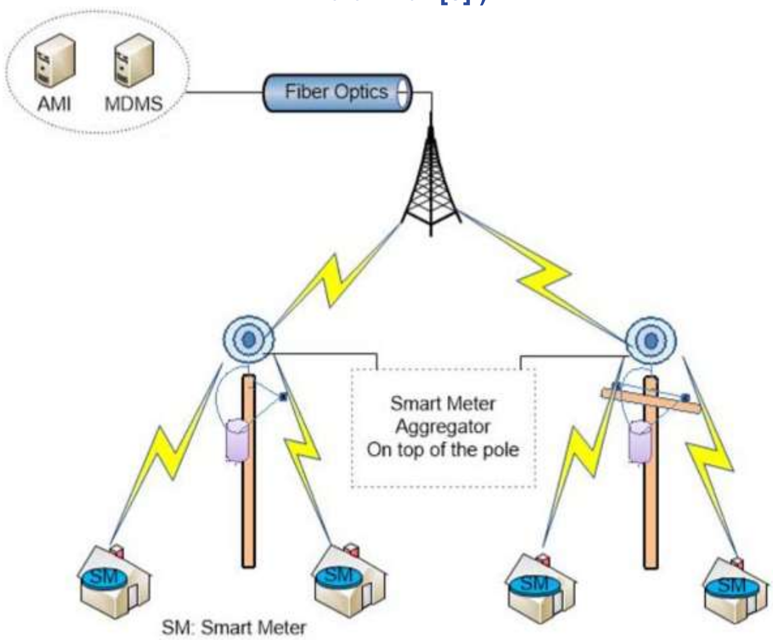
The wide popularity of smart meters allows a huge amount of electricity consumption data to be collected. Billing is no longer the only function of smart meters. Collection of data with higher effectiveness of the service, integrality, and universal access from smart meters is provided, producing valuable information about consumer electricity consumption, behaviors, and lifestyles [9].

Smart meters installed in customer environments can store electricity consumption data periodically for each individual consumer. These data are sent for one or more data aggregators over a smart electric grid.

Figure 1 shows an example of a smart grid. It consists of smart meters, data aggregators on top of utility poles communicating over wireless links, advanced metering infrastructure (AMI), and the meter data system management (MDSM), which can collect the data measured by the smart meters. The AMI is considered a key component of smart electricity grids, integrating software and hardware components, meter data management systems, monitoring systems, and information and control systems [10].

The aggregator is responsible collecting all the data from the several smart meters. It can become a bottleneck in communications because the amount of transmitted data is high. Therefore, the bandwidth for the data aggregator must be high. There are different bandwidth requirements of the smart meter from the aggregator point of view. Therefore, the aggregator is a vital element in the network [11].

Figure 1 –Example of a smart electric grid with aggregators and smart meters. (Adapted from Aalamifar [8].)



The efficient choice of aggregator locations in a smart grid is a difficult task. This task is aggravated in large cities, which contain thousands of customers and consequently thousands of smart meters [12]. According to Carniel and Mestria [13], the limitation in bringing into play smart meters and aggregators is their high cost, principally the aggregators. That cost varies with the number of aggregation channels and their functionalities, principally with introduction of data security features. The cost of an aggregator was estimated equal to a phasor measurement unit (PMU) [14]. A search of the literature reveals a minimum cost of a PMU in $40,000 (for two supported channels) and an additional cost of $4,000 per channel [14], in addition to the costs of fiber optics, wireless links, and switches (where here and in the following sections, all costs are quoted in U.S. dollars).

Santos [15] used a strategy to connect smart meters to aggregators, each of which had a capacity to connect with an average of 30 smart meters. The smart meters collected electrical data parameters at a rate ranging from once every five minutes to once an hour. This means that electric power distribution utilities need to handle a considerable amount of data. In this case, it was estimated that the amount of data collected in a month would be a factor of 3000 greater than that from a single smart meter and this only for only one type of data: energy consumption [16].

Sharma and Saini [17] developed a metering distributed system in which the smart meters have as a standard a single channel or three channels. They also studied the secure communication protocols needed in the smart grids. In addition, they described the following power measurement parameters: watt-hour accuracy, measurement range, supply voltage, analog-to-digital conversion, packet types in data transport, and the effect of harmonics on metrology that impacts the reliability of smart metering infrastructure.

Tavasoli, Yaghmaee, and Mohajerzadeh [18] presented a location problem for the data aggregators in a hybrid communication network, including fiber optics and WiMAX. The results demonstrated that the optimization model at data aggregation points enables minimized costs and density of data aggregators. They also proposed a location problem incorporating quality of service metrics and mixed topologies (e.g., urban, suburban, and rural) into the optimization model. It is worth noting that the total number of aggregators allocated in their hybrid topology is less than the sum of aggregators needed for each one individually. This is because the areas overlap in the mixed networks.

The allocation of smart meters and data aggregators is important in the smart grids because it will enable the management in the electric power quality and provides services to the customers. In future, the electricity markets will ensure that consumers have access to the energy services with flexible and low cost. In this sense, the power electric systems based on clean energy sources as wind and solar power can offer an economic alternative energy service. Thus, it contributes as an important way to sustainable development in the world.

MATERIAL AND METHODS

In this work, a research of the literature was conducted in the Scielo, Scopus, Science Direct, IEEE Xplore, and CAPES databases using the following terms: smart grid, aggregator, set covering, smart meter, optimization, distribution system, and heuristics. Our research methodology is applied and exploratory with an inductive and quantitative method. We propose a mathematical model for the data aggregator allocation problem in smart grids. The method applied to the proposed mathematical model was a B&B exact method.

Mathematical model

The proposed model adds in the Set Covering Problem (SCP) formulation described in Beasley [19] the introduction of the dispersion constraint among the aggregators called the dispersion set covering problem (DSCP).

The DSCP is NP-hard because it can be reduced to the SCP. The SCP is formally defined as follows: Given a set of m elements of M = {1, …, m} and a collection of n subsets N = {S_j

$\subseteq$ M, with $1 \leq j \leq n$}, with non-negative costs, a collection of subsets $X \subseteq N$ is a cover of M if the following holds: $\cup S_{j\in X}$ Sj = M. Then, X is an optimal cover of M without any redundant subset in X; i.e., X will not cover M if any subset is removed from X. The goal of the SCP is to find a minimal cost cover X of M.

Henceforward, we will call a standard smart meter as smart meter, in short. For the data aggregator allocation problem in a smart grid, the elements to be covered are the smart meters. A cluster formed by the smart meters should be covered by at least one candidate aggregator (each with a cost of cj), where X is composed of clusters of all allocated aggregators [20]. In this work, we have assumed one-hop communication between the smart meters and the data aggregators (i.e., a single signal path from the smart meter to the aggregator). The DSCP model is formulated as a binary integer nonlinear programming problem as follows:

$$\min z = \sum c_j . x_j,\ j = 1, \ldots, n, \qquad (1)$$

subject to

$$\sum aij.xj \geq 1,\ \ j = 1, \ldots, n;\ i = 1, \ldots, m, \qquad (2)$$

$$r x_j x_k \leq dist_{jk},\ r > 0,\ \ \forall j,k \in N, \qquad (3)$$

$$x_j \in \{0,1\},\ \ j = 1, \ldots, n, \qquad (4)$$

where $a_{ij} = 1$ if $i \in S_j$ and $a_{ij} = 0$ otherwise, r is the minimum distance between two allocated aggregators, distjk is the distance between candidate aggregator j and candidate aggregator k, and the decision variable $x_j = 1$ if the subset S_j belongs to cover X and $x_j = 0$ otherwise.

The objective function (1) minimizes the allocation of candidate data aggregators to the smart meters. Note that there are several types of aggregators in the real world, each with different costs. In the literature, these are simplified as a single cost cj = C, for all j = {1, …, n}, where C is a constant value. In this work, the different costs for candidate data aggregators will be introduced, where each has a number of supported channels ranging from a few simple ones ($40,000) to more sophisticated setups ($152,000). The aggregator costs do not account for the costs of data security (e.g., the use of encryption of the data sent by the smart meters).

Constraints (2) ensure that every smart meter must be covered by at least one aggregator. Constraints (3) ensure that the aggregators have a minimum distance between them. The purpose of constraints (3) is to ensure that the aggregators are not in close proximity. These constraints allow different smart meters to be served by different aggregators.

Constraints (3) decrease redundancies, preventing delays between data packets. The integrality constraints are outlined in (4) with the integer variables in the DSCP model equal to 0 or 1 to represent a decision.

For a radio communication between a smart meter and an aggregator to take place, a maximum distance (md) is required between these two devices. This means that the smart meter–aggregator distance has to be less than or equal to the transmission range of the smart meter. Therefore, to reduce the redundancies of the aggregators, each one has to be at a minimum radial distance (r) from the others ($r \geq md$). The value of md was set equal to 100 m, in accordance with the study of Aalamifar *et al.* [8]. Therefore, in constraints (3), $r \geq 100$ is assumed.

Two steps for solving allocation problem

The mathematical optimization model is solving by an algorithm in the two steps: first, the smart meters were allocated to the nearest aggregators (in this case, each smart meter can be allocated one or more aggregator). A smart meter is allocated to the aggregators, since that it reaches until a distance (md). This distance is required between these two devices (smart meter and aggregator) due to the radio communications.

After that, in second step, an effective B&B exact method that minimizes the cost acquisition of aggregators is applied to model showed in section 3.1. We call attention to that the model is formulated as an integer programming problem. The data input to proposed model are compose by cost of aggregators (coefficients in the objective function of DSCP model), coverage matrix (constraints number 2 in this model), distances between candidate data aggregators (right-hand side of the constraints number 3), and minimum radial distance.

An exact B&B method will be used to solve the DSCP with the LINDO solver [21]. The basic idea of the B&B algorithm is to solve in the first phase the linear programming problem. Then, the B&B algorithm performs several steps until the solution found is an integer solution. In the first phase, if the algorithm does not find an integer solution, i.e., for any variable x_i ($i = 1, 2, \ldots, n$), then the B&B algorithm chooses a fractional variable $x_m = K$, which is a branch variable, where K is a value belonging to the set of real numbers ($K \in R$) and m is an integer ($m \in N$), with $m = (i = 1, 2, \ldots,$ or $n)$.

Let L be the truncated integer value of K. New subproblems can then be created by alternately attaching one of two constraints: $x_m \leq L$ or $x_m \geq L+1$. This branching is continued as long as there are fractional variables. The B&B algorithm performs several feasibility tests to be satisfied. In the end, if the B&B algorithm has not exceeded the memory limit and it has found a feasible solution, the B&B algorithm stops and it presents the optimal solution. The Algorithm 1 shows the pseudocode to solving data aggregator allocation in the smart grids.

The Algorithm 1 reads data input: coefficients of the objective function, coverage matrix, distances between candidate data aggregators, minimum radial distance, and total number of candidate data aggregators and smart meters. In a first step the Algorithm 1, each smart meter is allocated to the nearest aggregators using nearaggregators procedure. After, in a second step, the Algorithm 1 obtains a file data using formulDSCP procedure according to the DSCP model described in section 3.1. This model is formed by data input and the temporary matrix that consists of the candidate aggregators and meters. Next, into the second step, the Algorithm 1 applies a branch-and-bound approach by exactmethod procedure. Finally, the Algorithm 1 prints the final solution.

Algorithm 1: Two Steps

Notation:

C[*j*] - vector of objective function *j* coefficients

M[*i,j*] - matrix of coverage of *i* meters and *j* aggregators

D[*j,k*] - matrix of distances between *j* candidate data aggregators and *k* candidate data aggregators

md - minimum radial distance

n - total number of candidate data aggregators

l - total number of smart meters

T[*i,j*] - temporary matrix that consists of *j* candidate aggregators and *i* meters

f - file data using the DSCP model

S - final solution

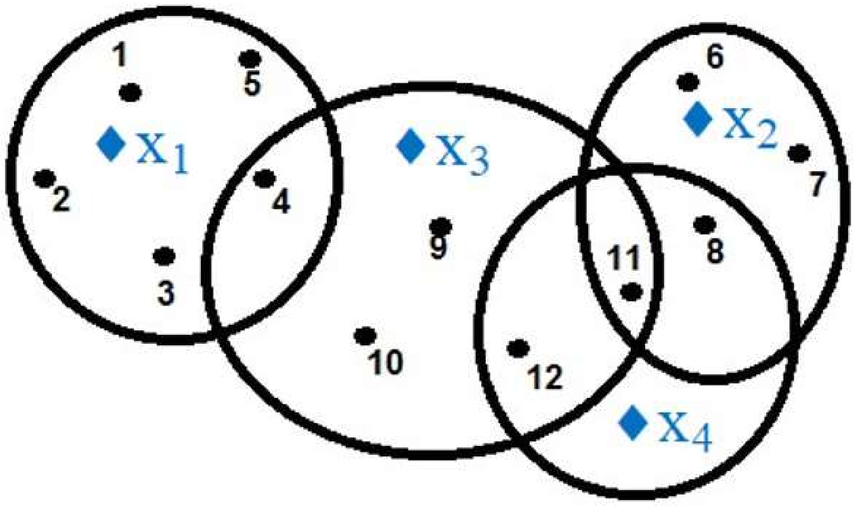
```
Input (C, M, D, md, n, l);
Output (S);
S={ };
Begin
    //first step
    for (i=1 to l) do
      for (j=1 to n) do
        if (meter i is nearest to aggregator j) then
          T=nearaggregators(i,j);
        endif
      endfor
    endfor
//second step
f=formulDSCP(C, M, D, md, n, l, T);
S=exactmethod(f);
Print S Solution;
End
```

Example of a smart grid

To illustrate the data aggregator allocation problem with the DSCP mathematical model, we present an example consisting of four candidate aggregators (x_1, x_2, x_3, and x_4) and 12 smart meters (numbered in sequence from 1 to 12), as shown in Figure 2. All smart meters primarily were allocated to the nearest aggregators, conform first step. For example, the smart meters of number 1, 2, 3, and 5 were allocated to the aggregator of number 1 (denoted by x_1); smart meter of number 4 was allocated to the aggregators (1 and 3); the smart meters of number 6 and 7 were allocated to the aggregator 2 (denoted x_2).

After that, the smart meter of number 8 was allocated to the aggregators (2 and 4); the smart meters of number 9 and 10 were allocated to the aggregator 3 (denoted x_3). Next, smart meter of number 11 was allocated to the aggregators (2, 3, and 4); smart meter of number 12 to the aggregators (3 and 4), with aggregator of number 4 represented by variable x_4.

Figure 2 – Example of a smart grid with candidate aggregators and smart meters.



It needs to call attention to that some smart meters were allocated only to an aggregator and other two or more aggregators, conform is shown in Figure 2. In this way, then we use an optimization mathematical formulation (Figure 3) according to the DSCP model proposed in section 3.1. The DSCP model is composed with the input data as: coverage matrix and cost of aggregators (both showed in Table 1), r minimum radial distance, and distances between candidate data aggregators (Table 2).

Table 1 – Example of coverage matrix composed of smart meters and aggregators.

Costs of candidate aggregators ($1000s)	$c_1 = 10$	$c_2 = 15$	$c_3 = 18$	$c_4 = 4$
Decision variables	x_1	x_2	x_3	x_4
Number of smart meters				
1	1	0	0	0
2	1	0	0	0
3	1	0	0	0
4	1	0	1	0
5	1	0	0	0
6	0	1	0	0
7	0	1	0	0
8	0	1	0	1
9	0	0	1	0
10	0	0	1	0
11	0	1	1	1
12	0	0	1	1

An exact method to the optimization mathematical formulation is applied to decrease redundancies and to ensure that the aggregators are not in close proximity. The solution reached by exact method is presented in Figure 4. Next, it is showed the formulation and solution optimization using input data from Figure 2, in more details.

Table 1 lists the costs of the candidate aggregators in thousands of dollars, the decision variables, and the coverage matrix of the smart meters distributed to the candidate aggregators, corresponding to the example of Figure 2. Table 2 lists the distances in linear meters between the candidate data aggregators (with distjk that is the distance between candidate aggregator j and candidate aggregator k) and r minimum radial distance between them.

Table 2 – Distances between the candidate data aggregators and a minimum radial distance (r = 100 m).

Decision variable (candidate aggregators)	x_1	x_2	x_3	x_4
x_1	0	140	110	200
x_2	140	0	110	120
x_3	110	110	0	140
x_4	200	120	140	0

Figure 3 – Optimization mathematical formulation as data input to the solver.

```
min 10x1 + 15x2 + 18x3 + 4x4
st
x1 >=1
x2 >=1
x3 >=1
x1 + x3 >=1
x2 + x4 >=1
x2 + x3 + x4 >=1
x3 + x4>=1
100x1*x2<=140
100x1*x3<=110
100x1*x4<=200
100x2*x3<=110
100x2*x4<=120
100x3*x4<=140
end
GIN x1
GIN x2
GIN x3
GIN x4
```

Figure 3 presents an optimization mathematical formulation using the proposed model applied to the DSCP as input data (via a text file) to the solver LINDO. Figure 3 uses the data from Tables 1 and 2 and the smart grid information described in Figure 2. Figure 4 presents a solution showing the objective function value ($43,000) and the allocated data aggregators (x1, x2, and x3).

Table 3 presents more details about the obtained solution with the smart grid information described in Figure 2. The first, second, and third columns of Table 3 list the instance (scp4-12-4-18), the number of candidate aggregators, and the number of smart meters, respectively. The fourth and fifth columns list the lower and higher costs of aggregators, respectively. The next section will describe the details of how was generated the instances.

Figure 4 – Solution of the optimization mathematical formulation using input data from Figure 3

```
OBJECTIVE FUNCTION VALUE
      1)     43.00000
 VARIABLE        VALUE          REDUCED COST
      X1         1.000000        10.000000
      X2         1.000000        15.000000
      X3         1.000000        18.000000
      X4         0.000000         4.000000
```

Table 3 – Obtained values for the instance of Figure 2

Instance	Candidate aggregators	sm	Lower cost	Higher cost	OF ($1000s)	t (s)	Allocated aggregators
4-12-4-18	4	12	4	18	43	< 1	1, 2, 3

The sixth and seventh columns of Table 3 present the objective function (OF) value in thousands of U.S. dollars and the computational time (t) reached by using the B&B exact method (<1 s), respectively. The last column gives the allocated data aggregators (1, 2, and 3) that are represented by (x_1, x_2, and x_3) variables, all set equal to 1, to communicate to the smart meters.

Instances and scenarios applied in smart grids

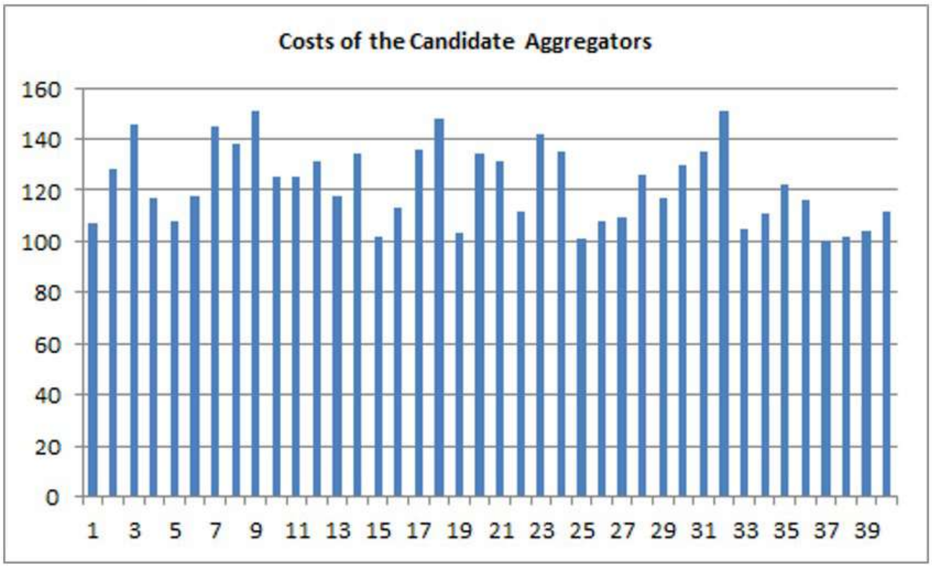
Tests using the proposed mathematical model applied to the DSCP were performed for several scenarios composed of several instances. In each instance was generated with smart meters and candidate data aggregators distributed in a planar region for a generated smart grid in a simulation way. It is assumed that the bandwidth of each smart meter to be identical, so-called standard smart meter. In this instance, the topology of the smart grids, in planar regions, consists in the installation of the smart meters in an urban scenario without interference source. In this scenario, there are the data aggregators that will be selected among candidate data aggregators.

The following parameters used in the instances were varied: costs of the aggregator in thousands of dollars with a uniform distribution on the range [40–152], the number of candidate aggregators randomly distributed in a planar region in the range [21–290], and the number of smart meters in the range [120–725].

It is guaranteed that the data aggregators are distributed in the planar region in a way that each smart meter will be covered by one or more candidate data aggregators. For example, instance scp40-140-100-152 has in its scenario 40 candidate aggregators distributed in a planar region, 140 smart meters assigned in the same region, and the costs of each candidate aggregator in a uniform distribution between 100 and 152, as shown in Figure 5.

The distance from the smart meter to the candidate aggregator must be equal to md, where md is the maximum distance permitted. Recall that some smart meters may be assigned to more than one aggregator. Moreover, in the proposed mathematical model, the aggregators are at the minimum distance between them.

Figure 5 – Costs of the candidate data aggregators



Two scenarios (1 and 2) were created for small- and medium-size instances. In the first scenario (1), the numbers of smart meters were kept fixed, but their positions were different in each instance, and they were distributed in a single planar region. In this scenario, the number of aggregators and costs are different.

In scenario 1, other tests were also performed with a constant number of candidate aggregators and different costs. In this last case, the numbers of fixed smart meters distributed in a single planar region in different locations were also varied.

In scenario 2 (the case of the more robust smart grids), a new planar region was generated with five different planar regions, where each one is like that of scenario 1.

COMPUTATIONAL RESULTS

This section presents the computational results using the LINDO solver release 6.1 run on a computer with an i7 processer, operating at 2.2 GHz, with 8 GB of memory and 6 cores. The tests were conducted for the two scenarios described in the last section. In both scenarios, extremely short computational times were reached, with some being <1 s.

Scenario 1

Table 4 presents the values obtained for various instances, with the number of smart meters equal to 140 and the higher cost of the aggregator equal to $152,000. The first, second, third, and fourth columns of Table 4 list the instance type, the number of candidate aggregators, the lower cost of the aggregator, and the value of the objective function (OF) in thousands of dollars, respectively.

Table 4 – Obtained values for the instances in scenario 1

Instance	Aggregators	Lower cost	OF ($1000s)	t (s)	Allocated aggregators
scp21-140-40-152	21	40	336	<1	1, 6, 15, 19, 21
scp22-140-40-152	22	40	329	<1	1, 2, 13, 15, 19
scp23-140-40-152	23	40	320	<1	1, 4, 10, 15, 19
scp24-140-40-152	24	40	335	<1	1, 6, 12, 15, 19
scp25-140-40-152	25	40	273	<1	5, 6, 15, 19, 25
scp26-140-60-152	26	60	377	1	4, 15, 16, 25, 26
scp27-140-60-152	27	60	361	1	1, 5, 15, 16, 19
scp28-140-60-152	28	60	298	1	5, 13, 15, 19
scp29-140-60-152	29	60	385	1	1, 4, 16, 25, 27
scp30-140-60-152	30	60	344	1	8, 15, 26, 27
scp31-140-80-152	31	80	403	2	11, 16, 21, 31
scp32-140-80-152	32	80	369	1	17, 18, 21, 30
scp33-140-80-152	33	80	366	1	14, 21, 22, 33
scp34-140-80-152	34	80	434	2	6, 14, 17, 21, 31
scp35-140-80-152	35	80	388	1	15, 17, 22, 34

scp36-140-100-152	36	100	448	2	6, 26, 29, 33
scp37-140-100-152	37	100	460	2	12, 15, 25, 28
scp38-140-100-152	38	100	463	4	5, 14, 16, 26
scp39-140-100-152	39	100	502	12	11, 18, 22,29
scp40-140-100-152	40	100	438	2	19, 28, 33, 39

The fifth column of Table 4 gives the computational time demanded by the exact method in seconds, and the last column gives the number of allocated aggregators needed to serve the smart meters. Table 4 list four sets of instances distributed by the lower costs of the aggregators (40, 60, 80, and 100), but the higher cost of the aggregators is a constant value of 152.

We observe that, as more candidate aggregators are introduced, the cost of the objective function decreases in each individual set of instances. In the first set of instances, it was observed that aggregators 15 and 19 ($x_{15} = x_{19} = 1$ in the solver) were chosen in every case in this set. These aggregators have lower costs (fewer number of supported channels), but they are still able to serve a portion of the smart meters.

Something similar was observed with aggregator 21 ($x_{21} = 1$ in the solver) in the third set of instances, except in one instance. In Table 4, the average number of allocated aggregators was 4.45. As the number of aggregators is an integer, and the ratio between the numbers of smart meters to the number of allocated aggregators is obtained as 5, we obtain the average number of 28 smart meters for each allocated aggregator.

Table 5 presents the values obtained for various instances, with a constant number of the candidate aggregators equal to 50, 140 smart meters, and the lowest and highest costs of the aggregators being 100 and 152 in thousands of dollars, respectively. The first, second, and third columns of Table 5 list the instance type, the objective function value in thousands of dollars, and the computational time in seconds, respectively. The fourth column gives the number of allocated aggregators needed to serve the meters, and the last column gives the type of distribution (TD) of smart meters in the planar region. We identified the different distributions of smart meters over the planar regions by the numbers 1, 2, 3, 4, and 5. For example, instance scp50-1-140-100-152 was generated in scenario 1 with 50 candidate aggregators, 140 smart meters, and the costs of each candidate aggregator in a range from 100 to 152.

Table 5 – Obtained values for instances in scenario 1 with a constant number of aggregators.

Instance	Objective function ($1000s)	t (s)	Allocated aggregators	TD
scp50-1-140-100-152	439	4	2, 17, 30, 49	1
scp50-2-140-100-152	427	7	8, 19, 29, 43	2
scp50-3-140-100-152	439	5	3, 19, 30, 35	3
scp50-4-140-100-152	467	6	25, 26, 30, 38	4
scp50-5-140-100-152	472	9	8, 42, 46, 47	5

In Table 5, we observe that the average cost of the objective function was 448.8, with a standard deviation 19.6. Aggregator 30 ($x_{30} = 1$ in solver) was chosen in three instances because it has a lower cost. The aggregator has fewer supported channels, but it serves several smart meters. The number of allocated aggregators was four in all instances, with the ratio between the numbers of smart meters to the number of allocated aggregators equal to 35.

Table 6 presents the values obtained by instances with the number of candidate aggregators ranging from 56 to 60, a fixed number of smart meters equal to 120, the lowest aggregator cost equal to $110,000, and the highest cost equal to $152,000. The first, second, third, and fourth columns of Table 6 present the type of instance, the value of the objective function in thousands of dollars, the computational time in seconds, and the number of allocated aggregators, respectively.

Table 6 – Obtained values with different parameters on smart meters, aggregators, and costs.

Instance	OF ($1000s)	t (s)	Allocated aggregators
scp56-120-110-152	453	6	1, 12, 19, 29
scp57-120-110-152	467	8	14, 29, 38, 42
scp58-120-110-152	478	7	27, 38, 41, 55
scp59-120-110-152	494	7	20, 38, 52, 59
scp60-120-110-152	471	9	15, 20, 27, 29

Table 6 presents aspects of scenario 1 different from those previously studied, with a variety of parameters of the smart grids. We observe in this scenario that the number of candidate aggregators was increasing, but costs of the candidate aggregator were decreasing. From Table 6, one can also observe that, for a diverse number of aggregators, the average cost of the objective function was 472.6, with a standard deviation of 15.0.

The number of allocated aggregators was four in all instances; the ratio between the numbers of smart meters and the aggregators was equal to 30. Aggregator 29 (x_{29} = 1) was chosen in three instances because of its low cost.

Scenario 2

Table 7 lists the values obtained for instances with a powerful smart grid. The first, second, third, and fourth columns of Table 7 present the instance identifier (Id), the number of candidate aggregators, the number of smart meters (sm), and the lower cost of the aggregators, respectively.

Table 7 – Values obtained for the instances in scenario 2.

Id	Aggregators	sm	Lower cost	Higher cost	OF ($1000 US)	t (s)	taa	R
1	190	600	110	152	2349	10	20	30
2	215	700	120	152	2509	25	20	35
3	240	700	120	152	2516	50	20	35
4	265	725	120	152	2533	51	20	36.25
5	290	600	110	152	2363	32	20	30

The fifth, sixth, and seventh columns in Table 7 present the higher cost, the value of the objective function (OF) in thousands of dollars, and the computational time demanded in seconds. The eighth column gives the total number of allocated aggregators (taa) demanded by the smart meters, and the last column gives the ration R (=sm/taa) between the smart meters (sm) and the total number of allocated aggregators.

We can observe that, when the costs of the aggregators are high, the R ratio is increasing, as can be seen in Tables 5 and 7. This occurs because most expensive aggregators have a larger number of supported channels with capacity to allocate more smart meters.

DISCUSSION

The mathematical model was applied for several instances with a variety of parameters. The parameters tested were size instances, aggregator costs, numbers of aggregators, dispersion between aggregators, and numbers of standard smart meters. It proved to be robust and supported variations in the parameters.

We performed several tests using a new mathematical model applied to the DSCP. The computations revealed the minimum cost for allocation of aggregators in covering the standard smart meters in all instances. This model can be used as a decision-making tool for engineers in the deployment of smart grids and will be of great importance for the electric power distribution utilities. Beyond of that provides benefits to the consumers in a smart energy context in the smart cities. Smart energy aims to serve energy demands incorporating renewable energy sources to maintain sustainability while minimizing adverse effects on the environment [22].

In the proposed scenarios in this work, computational results yielded a ratio from 28 standard smart meters to 36.25 for each allocated aggregator. This ratio is close to the value of Santos [15], who found a value of 30 standard smart meters for each allocated aggregator.

Additional research tasks can be undertaken to overcome the limitations of the current study. These limitations are due to the B&B algorithm that solves the DSCP for a certain number of aggregators and standard smart meters. The main computer memory does not support processing the algorithm for large-size instances. We know that the solution to the DSCP is affected by an increase in the number of aggregators. When the numbers of aggregators are large, use of the main computer memory will be substantially increased.

According to Araújo and Mestria [23], finding an optimal solution in a reasonable computational time is possible only for small- or medium-size instances. In this sense, heuristic or metaheuristic methods can be used to solve the data aggregator allocation problem in large smart grids.

CONCLUSIONS AND FUTURE WORK

Given the importance of the deployment of smart grids and metering infrastructure, we addressed the modeling of and solution to the data aggregator allocation problem in smart grids. This deployment brings up electric power quality control and provides flexible energy services to the customers with low cost using sustainable energy sources.

For this purpose, we used an optimization mathematical model based on the set covering problem. We added dispersion constraints between the aggregators to the coverage model.

The solutions obtained by an exact method using optimization mathematical model proved to be robust when it applied different parameters. We conclude that new optimization model was capable to solve the problem of data aggregator allocation in the smart grids. Thus, the set of standard smart meters and aggregators were rationalized providing benefits for the electric utilities. The computational results showed that ratio between standard smart meters and an allocated aggregator is line with the literature.

The limitations concern customer privacy of confidential information about their energy consumption habits. Therefore, the data aggregators must receive and send information without revealing confidential information. In this case, data acquired from the aggregators must be encrypted to preserve the privacy of the customer static data [24]. Thus, the aggregators' acquisition costs will be higher. Questions about security in the AMI infrastructures can also be addressed in a future study [25].

Finally, the model explored in this work can be modified by considering a maximum budget (MAX) to purchase the data aggregators by electric power distribution utilities. Therefore, the following constraints should be introduced into the DSCP model [(1)–(4)] as a binary integer nonlinear programming problem:

$$\sum c_j x_j \leq MAX, \quad j = 1, \ldots, n. \quad (5)$$

In future work, we propose using the chemical reaction optimization metaheuristic of Yu, Lam, and Li [26] for the DSCP. Because the DSCP is NP-hard, this will enable quality solutions to be found in a reasonable computational time for large-size instances.

Acknowledgments

We would like to thank a Language Editing service of the Elsevier (Certificate Serial number: LE-208038-4E1D581D8576, Date: 17-Mar-2021) for English language editing.

REFERENCES

1. Cardena JA, Gemoets L, Rosas JHA and Sarfi R. A literature survey on smart grid distribution: an analytical approach. Journal of Cleaner Production, vol. 65, p. 202-216, 2014.

2. Rolim G, Passos D, Albuquerque C and Moraes I. MOSKOU: A Heuristic for Data Aggregator Positioning in Smart Grids. IEEE Transactions on Smart Grid, vol. 9, n. 6, p. 6206-6213, 2018.

3. Palate BO, Guedes TP, Grilo-Pavani A, Padilha-Feltrin A and Melo JD. Aggregator units allocation in low voltage distribution networks with penetration of photovoltaic systems. International Journal of Electrical Power & Energy Systems, vol. 130, p. 107003, 2021.

4. Naamane A and M'sirdi NK. Towards a Smart Grid Communication. Energy Procedia, vol. 83, p. 428-433, 2015.

5. Boccardo D, *et al.* Energy footprint framework: A pathway toward smart grid sustainability. IEEE Communications Magazine, vol. 51, n. 1, p. 50-56, 2013.

6. Lopes D, *et al.* An Algorithm for the Generation of Routes For The Collection Of Solid Waste In The City Of Manaus Using IoT Data. Sodebras, v. 16, n.174, p. 50-53, 2020. ISSN 1809-3957.

7. Nepomuceno C, *et al.* A Theoretical and Practical Approach to an IoT protocol. Sodebras, v. 15, n.169, p. 113-118, 2020. ISSN 1809-3957.

8. Aalamifar F, Shirazi GN, Noori M and Lampe L. Cost-efficient data aggregation point placement for advanced metering infrastructure. In: 2014 IEEE International Conference on Smart Grid Communications (SmartGridComm), Venice, Italy, p. 344-349, 2014.

9. Wang Y, Chen Q, Hong T and Kang, C. Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges. IEEE Transactions on Smart Grid, vol. 10, n. 3, p. 3125-3148, 2019.

10. Martins JFA, *et al.* Smart Meters and Advanced Metering Infrastructure. In: TAŞCIKARAOĞLU, A and ERDINÇ O. (Ed.). Pathways to a Smarter Power System. Academic Press, 2019, p, 89-114.

11. Balachandran K, Olsen RL and Pedersen JM. Bandwidth analysis of smart meter network infrastructure. In: 16th International Conference on Advanced Communication Technology, Pyeongchang, South Korean, p. 928-933, 2014.

12. Rolim G, Albuquerque CVND and Moraes IM. Modelo e solução para o problema de posicionamento de agregadores em redes elétricas inteligentes (in Portuguese). In: XXXIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, Vitória, Espírito Santo, Brazil, 14 pp, 2015.

13. Carniel AZ and Mestria M. A Chemical Reaction Optimization Algorithm for Phase Measurement Units Placement. In: XXI ENMC - Encontro Nacional de Modelagem Computacional e IX ECTM – Encontro de Ciências e Tecnologia de Materiais, Búzios, Rio de Janeiro, Brazil, 10 pp. 2018.

14. Mohammadi MB, Hooshmand RA and Fesharaki, FH. A New Approach for Optimal Placement of PMUs and Their Required Communication Infrastructure in Order to Minimize the Cost of the WAMS. IEEE Transactions on Smart Grid, vol. 7, n. 1, p. 84-93, 2016.

15. Santos JPV. Localização de agregadores de dados em redes elétricas inteligentes (in Portuguese). Dissertação de Mestrado Integrado em Engenharia Electrotécnica e de Computadores. Faculdade de Ciências e Tecnologia, Universidade de Coimbra, Portugal, 60 pp, 2019.

16. Alquthami T, *et al.* Analytics framework for optimal smart meters data processing. Electrical Engineering, vol. 102, n. 3, p. 1241–1251, 2020.

17. Sharma K and Saini LM. Performance analysis of smart metering for smart grid: An overview. Renewable and Sustainable Energy Reviews, vol 49, p. 720-735, 2015.

18. Tavasoli M, Yaghmaee MH and Mohajerzadeh, AH. Optimal placement of data aggregators in smart grid on hybrid wireless and wired communication. In: 2016 IEEE Smart Energy Grid Engineering (SEGE), Oshawa, Canada, p. 332-336, 2016.

19. Beasley JE. A Lagrangian Heuristic for Set-Covering Problems. Naval Research Logistics, vol. 37, n. 1, p. 151–164, 1990.

20. Fernandes BM and Mestria M. Modelo Matemático para o Planejamento de Agregadores em Smart Grids (in Portuguese). In: Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, Campinas, São Paulo, Brazil, v. 6, n. 2, 2 pp., 2018.

21. LINDO. LINDO User's Manual. Chicago, Illinois: LINDO Systems Inc, 2003, 298 pp. Available in: <https://www.lindo.com/downloads/PDF/LindoUsersManual.pdf>. Accessed in September 18, 2020.

22. Siva BN, Khan M and Han, K. Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities. Sustainable Cities and Society, vol. 38, p. 697-713, 2018.

23. Araújo ML and Mestria M. GRASP Method For Planning Wireless Mesh Networks. Sodebras, v. 12, n.144, p. 171-176, 2017. ISSN 1809-3957.

24. Jiang H, Luo W and Zhang Z. A privacy-preserving aggregation scheme based on immunological negative surveys for smart meters. Applied Soft Computing, vol. 85, p. 105821, 2019.

25. Shrestha M, Johansen C, Noll J and Roverso, D. A Methodology for Security Classification applied to Smart Grid Infrastructures. International Journal of Critical Infrastructure Protection, vol. 28, p. 100342, 2020.

26. Yu JJQ, Lam AYS and Li VOK. Chemical reaction optimization for the set covering problem. In: 2014 IEEE Congress on Evolutionary Computation (CEC), Beijing, China, pp. 512-519, 2014.