# On the Applicability of Ant Colony Optimization to Non-Intrusive Load Monitoring in Smart Grids

Antonio Gonzalez-Pardo[1,2]([✉]), Javier Del Ser[2], and David Camacho[3]

[1] Basque Center for Applied Mathematics (BCAM), Alameda de Mazarredo, 14, 48009 Bilbao, Basque Country, Spain
`agonzalezp@bcamath.org`
[2] OPTIMA Area, TECNALIA, 48160 Derio, Spain
`javier.delser@tecnalia.com`
[3] Computer Science Department, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Madrid, Spain
`david.camacho@uam.es`

**Abstract.** Along with the proliferation of the Smart Grid, power load disaggregation is a research area that is lately gaining a lot of popularity due to the interest of energy distribution companies and customers in identifying consumption patterns towards improving the way the energy is produced and consumed (via e.g. demand side management strategies). Such data can be extracted by using *smart meters*, but the expensive cost of incorporating a monitoring device for each appliance jeopardizes significantly the massive implementation of any straightforward approach. When resorting to a single meter to monitor the global consumption of the house at hand, the identification of the different appliances giving rise to the recorded consumption profile renders a particular instance of the so-called source separation problem, for which a number of algorithmic proposals have been reported in the literature. This paper gravitates on the applicability of the Ant Colony Optimization (ACO) algorithm to perform this power disaggregation treating the problem as a Constraint Satisfaction Problem (CSP). The discussed experimental results utilize data contained in the REDD dataset, which corresponds to real power consumption traces of different households. Although the experiments carried out in this work reveal that the ACO solver can be successfully applied to the Non-Intrusive Load Monitoring problems, further work is needed towards assessing its performance when tackling more diverse appliance models and noisy power load traces.

**Keywords:** Non-intrusive load monitoring · Ant colony optimization · Power consumption disaggregation

## 1 Introduction

Smart meters lie at the core of the Smart Grid ecosystem by providing the distribution operator with fine-grained information about the energy consumption

of the house, facility, industry or whatever energy consumer it monitors. The knowledge that can be extracted from these data can be used for a wide variety of purposes. As to mention, depending on the regulatory context power distribution companies or energy dealers could extract energy consumption patterns of their clients by analyzing their daily traces, and consequently propose personalized contracts taking into account these patterns. Moreover, companies could use these data to predict, or estimate, the future power demand of a specific area or building, which calls for inherent opportunities in short-term demand side management schemes. But benefits do not fall only on the side of the energy distributor. Customers could take advantage from the data extracted by the smart meters as well: this data allow users to timely know their power consumption and what appliances have generated this consumption, towards eventually changing their consumption habits.

The main problem underlying beneath this last hypothesized utility of disaggregating the power profile registered by the meter is the way these data can be obtained. One possible way consists of deploying a metering device per installed appliance. However, the costs associated with this approach are much higher than that incurred by using only one meter that captures the overall power consumption associated with the house, which makes the former a metering layout restricted to very particular scenarios.

When using this last approach (i.e. a single meter per house), a procedure for disaggregating the power signal is required in order to discriminates what appliances (model, number and on/off time instants) correspond to the observed consumption. This disaggregation is widely known as *Non-Intrusive Load Monitoring* (NILM) [13]. Unfortunately, the evergrowing variety of appliances and the different power consumption modes for a given appliance makes it extremely difficult to solve NILM problems from a generalistic standpoint, fact that has unchained a flurry of research on algorithmic approaches to this family of paradigms. As such, research works focused on NILM problems can be grouped in non-supervised and supervised methods. Non-supervised approaches exploit unlabeled data, hence there is no training process for the identification model. Most of the related work refer to *blind source separation* [10], *Hidden Markov Models* (HMM) [10,30], or *Factorial Hidden Markov Models* (FHMM) [6,16]. On the other hand, supervised approaches require a labeled dataset and a training process to adjust some of the parameters of the model. The application of supervised approaches to NILM problems mainly concentrates on pattern recognition and optimization tasks [17,22]. Pattern recognition relies mostly on *Neural Networks* (NN) [25] or *Support Vector Machines* (SVM) [18,25]. Likewise *hierarchical clustering* [3,14], *fuzzy C-means* [4], *Self-Organizing Maps* (SOM) [28,29] or *Hopfieldś networks* [19,20] are among the most popular clustering techniques used for NILM problems.

Notwithstanding the activity in this area, the application of computational intelligence algorithms for NILM problems has not been explored in depth yet. There are several works that instead of applying bio-inspired algorithms to disaggregate power consumption traces, such techniques optimize part of the

parameters of the underlying model such as the weights of the NN, or the definition of the clusters that contains the consumption profiles of the different appliances. In this sense, *Particle Swarm Optimization* (PSO) [2], Genetic Algorithms (GA) [1,7], Ant Colony Optimization (ACO) algorithms [26] or *Honey Bee Mating Optimization* [9] have been used for this purposes.

In this work the application of ACO for NILM problems is studied from a different perspective. The disaggregation of power consumption can be viewed as a Constraint Satisfaction Problem (CSP), for whose resolution several ACO approaches are available. This paper presents an initial study about the performance of a new ACO model for complex graph-based problems, which was first proposed in [11]. The model is described in detail particularly in what regards to the formulation of the NILM problem as a variant of the CSP. Simulation results will preliminarily show that the proposed ACO solver excels at disaggregating power signatures corresponding to different appliances generated from real consumption traces.

This paper is structured as follows: first, Sect. 2 shows how CSP problems can be tackled via ACO, whereas Sect. 3 describes the new ACO model proposed for the CSP problem modeling the power disaggregation paradigm undertaken in this manuscript. Section 4 discusses the obtained simulation results over real datasets and, finally, Sect. 5 ends the paper by drawing concluding remarks and outlining lines of future research.

## 2    Solving CSP Problems Using ACO Algorihtms

Nowadays, there is a huge number of problems that can be modeled as Constraint Satisfaction Problems (CSP). This family of problems is defined by a triple $\{X, D, C\}$, where $X$ is the set of variables that compose the problem, $D$ contains the possible values for the variables described in $X$, and $C$ is a set of constraints that relate the values of the different variables [8,27]. The NP-hard complexity featured by CSP problems motivates the widely reported use of heuristic algorithms for solving them efficiently, among which ACO schemes outstand prominently.

ACO algorithms are based on the foraging behaviour of ants [5]. Ants take different decisions during their execution that allow them to build their own solutions to the problem. The utilization of ACO algorithms to solve CSP problems requires the representation of the problem as a graph, over which the ACO is executed. This graph, called *construction graph*, is defined as $G = (V, E)$ where $V$ represents the nodes of the graph and $E$ is the set of edges that connect the nodes. Roli *et. al.* proposed this construction graph for solving CSP problems in [21], where they established a fully connected graph where nodes are pairs ⟨*variable*, *value*⟩. There are other different approaches from the one just described. Examples are the models proposed by Solnon [23,24] and Khan et. al. [15]. But in all of them, the resulting graph is extremely big due to the high number of nodes and edges created in the system.

In this work, a new model proposed in [11] is used to perform the disaggregation power consumption. This model has been used to solve the *N-Queens* [11] problem and the *Resource-Constraint Project Scheduling Problem* [12].

## 3  Description of the New Model for CSP Problems

This section describes the model proposed in [11] for solving CSP problems. This model is composed by three important aspects:

1. *Smaller Decision Graph*: in this model the decision graph is smaller than then ones generated in the approaches described in the previous section. This is due to the fact that in the new model the resulting graph contains a node per variable involved in the problem (independently of the different values that can be assigned to this variable).
2. *New Ants Behaviour*: the behaviour of the ants in the classical approaches (Roli et al., Solnon and Khan et al.) is simple: ants only travel through the graph. The action of visiting a node corresponds to an assignment of a value to a variable, because both (variable and value) are encoded in the node. In the new approach, ants have a slightly more complex behaviour, because when the ants arrive to a certain node, they have to select a value for the corresponding variable taking into account their own local solution and the different restrictions involved in the problem.
3. *The Oblivion Rate Meta-Heuristic*: the simplification in the size of the decision graph entails a fast growth in the number of pheromones created by the ants. For this reason, a new meta-heuristic has been included in the system for controlling the number of created pheromones.

This model can be used to solve any CSP problem as the one behind the NILM problem central to this paper. The rationale for formulating the NILM problem as a CSP instance lies on the triple $\langle X, D, C \rangle$ where, in the NILM case, $X$ is the set of appliances contained in the house and $D$ denotes the times of the day at which the different appliances can be turned on. Finally, $C$ limits that the maximum consumption at an specific time can not exceed the power consumption registered by the smart meter or, for example, that two different ovens are not intuitively expected to be used at the same time. Taking into account this description of the NILM problem as a CSP, the resulting graph contains a node per appliance. The edges of the graph connects the different nodes to allow ants to explore the start of different appliances and its compliance with the metric. Note that by using the classical approach (i.e. to represent that a single appliance can be turned on at e.g. any minute of a day), a total of 1440 nodes per appliance would be obtained. Nowadays, smart meters that register the power consumption per seconds are available in the market, and tens of appliances can be easily found in domestic houses. Consequently, the size of the resulting graph would be unmanageable.

Ants selects randomly the node their going to visit taking into account their local solution built so far, the heuristic function of the problem, and the

pheromones deposited in the graph. In this work, the heuristic function corresponds to the difference between the power registered by the smart meter and the power consumption generated by the solution built by the ant. More formally, given the power consumption registered by the smart meter at time $t$ (denoted as $P(t)$) and the historic power consumption generated by the ants solution up to time $t$ (denoted as $(S(t))$, the heuristic function used in this work favours those appliances whose power consumption at time $t$ $(a(t))$ approximates $S(t)$ to $P(t)$, i.e. $P(t) - (S(t) + a(t))$ is minimum.

Regarding the *Oblivion Rate* meta-heuristic, in this initial work there is no need to control the number of pheromones because in the experimental phase controlled experiments have been carried out focused on the study about the applicability of ACO algorithm to solve NILM problems.

## 4 Experiments and Discussion

This section presents the different experiments carried out in this work. As previously highlighted, the goal of this work is to test whether the ACO model proposed in [11] can be applied to NILM problems. To this end, the first step is to obtain a valid dataset with real measurements. We have selected the REDD database [16] from the set of open databases available in the Internet. This database stores the power consumption of 6 different houses during 1 month. For each house, the database contains the power readings of the two power mains and the individual circuits for the house. The individual circuits (the ones used in this work) contains the apparent power logged once every three seconds. In this work, we have used the dataset corresponding to `house2`, from which we have identified the individual power consumption for some of the appliances, specifically the refrigerator, the stove, the dishwasher, the disposal and the kitchen outlets. Figure 1(a) exemplifies this set of extracted signatures by depicting that obtained for the dishwasher via a thresholded amplitude gain detection procedure.

Once these individual signatures have been extracted, we have combined them to create several artificially aggregated power consumption of 20 houses during 1 day (i.e. 28800 measurements). For the simplicity of this initial study, we have assumed that there is not more than 1 appliance of each class in the house. This means that, for example, two refrigerators are not allowed to turn on at the same time. Each appliance can be turned on a certain maximum number of times in a day: disposal (2), dishwasher (3), outlets (5) and stove (2). Note that the refrigerator is assumed to be always on. In this work, we have generated a synthetic consumption of 20 instead of using the real data provided in the REDD database, because the goal of this preliminary work is to test the applicability of the described ACO model to NILM problems. For this reason, we need some controlled and reduced experiments to understand the practical behaviour of the system. Figure 1(b) shows one of the produced daily power consumption profiles.

Once we have generated the global power consumption of the houses, the ACO model is built. As shown in Fig. 2 the decision graph is composed by 6 nodes, each representing one of the 5 different appliances taken into account
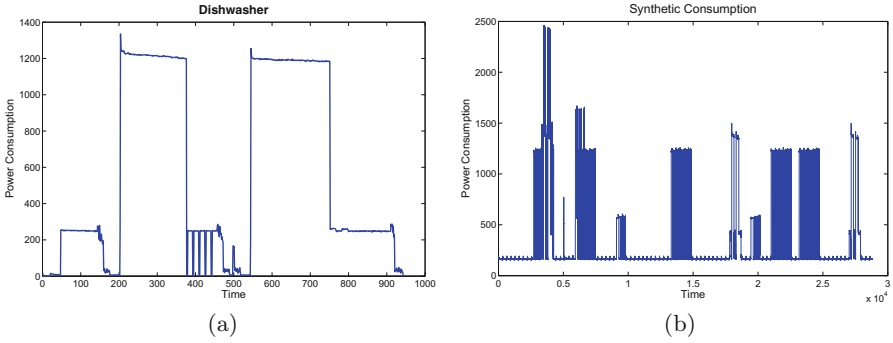
**Fig. 1.** (a) Evolution of the power consumption of the dishwasher; (b) One of the 20 synthetically generated power consumption profiles.

plus an additional appliance that stands for the initial status of the house when no appliance is on. Since we are dealing with 5 appliances and 28800 different timesteps (3 second sampling rate), by using the classical representation and the previous reasoning the resulting graph would amount up to 144000 nodes and $2.073 \times 10^{10}$ edges.
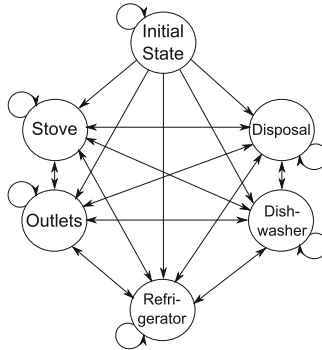


**Fig. 2.** Decision graph created in this work for the NILM problem using the approach described in [11].

Finally, the ACO algorithm is executed with 100 ants, 200 iterations, evaporation rate equal to 0.05, $\alpha = 1$, $\beta = 2$ and 10 repetitions. The evaporation rate defines the decrease speed of the pheromone values per step to permit bad solutions to disappear from the graph. Finally, $\alpha$ and $\beta$ are parameters that measures the influence of the heuristic value ($\alpha$) and the pheromones ($\beta$) on the decision process.

The evaluation of the quality of the produced solutions is given by the *mean square error* and the *precision* of the algorithm. The mean square error (MSE)

measures the different between the registered power consumption $P(t)$ and the power consumption $S(t)$ furnished by the solution of the ant over the $N = 28800$ samples measured by the smart meter. The precision is defined as the percentage of appliances correctly identified. The correct identification of any appliance results in the identification of the appliance (i.e. the kitchen outlets) but also the timestep when the appliance is turned on. A precision of 1 means that all the appliances have been correctly identified. Mathematically speaking:

$$\text{MSE} \doteq \frac{1}{N} \sum_{i=1}^{N} |P(t) - S(t)|^2 \tag{1}$$

$$\text{Precision} \doteq \frac{\text{Number of correctly detected appliances}}{\text{Total number of appliances in the daily profile}} \tag{2}$$

The performed experimental results over the synthetically produced profiles revealed that the proposed ACO model was able to fully solve 18 out of the posed 20 disaggregation problems with MSE = 0 and Precision = 1. In general the system gets the true solution in the early phase of the algorithm (initial steps). This was expected due to the following conditions:

1. The system operates on a reduced dataset in terms of number of appliances.
2. The global power consumption is generated using the individual signatures of the appliances and we do not consider different consumption profiles for the same appliance, nor do we address any variable-length power signature.
3. The global power consumption is generated only for one day (28800 measurements).
4. The generated power consumption is not subject to background noise.

All these characteristics ease significantly the identification of the appliances, yet still permits to conclude that the proposed ACO model can be applied to the NILM problem. Furthermore, the obtained results shed light on the practical applicability of this approach to the NILM problem and its expected performance when the above assumptions are removed.

## 5    Concluding Remarks and Future Research

The disaggregation of power consumption profiles is a area that is gaining momentum in the research community due to the plethora of applications and benefits for companies and customers stemming from the identification of customers' behavioral profiles in terms of energy consumption. In particular this paper has focused on the so-called Non-Intrusive Load Monitoring problem, which can be casted as to find the set of appliances of a certain energy consumer and their on/off instants based exclusively on its overall measured consumption.

This paper has elaborated on the applicability of the ACO algorithm for NILM problems. First of all, the problem has been shown to yield a Constraint Satisfaction Problem (CSP) where the set of variables are the different appliances, the values corresponds to the timestep when the appliances can be turned

on and the constraint limits the maximum power consumption per timestep. The ACO model for solving CSP proposed in [11] has been put to practice in this application scenario. This model creates a decision graph with a node per each appliance in the house being analyzed.

An individual power consumption signature for each appliance has been extracted from the selected REDD dataset, from which the power consumption of 20 different houses during one day has been artificially generated by taking into account the different appliances. Experimental results reveals that the ACO model can correctly disaggregate 18 out of 20 power consumptions.

Future work will be devoted to (1) the utilization of real power consumption during a longer time; and (2) the usage of Big Data functionalities (e.g. SPARK) to implement distributed, computationally efficient versions of the ACO algorithm aimed at accommodating the disaggregation of concurrently arriving meter data. Also will be of interest to assess the influence of measurement noise on the performance of the ACO algorithm and whether the Oblivion Rate metaheuristic is needed.

# References

1. Baranski, M., Voss, J.: Genetic algorithm for pattern detection in nialm systems. In: Proceedings of IEEE International Conference System, Man Cybernetics, vol. 4, pp. 3462–3468 (2004)
2. Chang, H.-H., Lin, L.-S., Chen, N., Lee, W.-J.: Particle-swarm-optimization-based nonintrusive demand monitoring and load identification in smart meters. IEEE Trans. Ind. Appl. **49**(5), 2229–2236 (2013)
3. Chicco, G., Napoli, R., Piglione, F.: Comparison among clustering techniques for eleelectric customer classification. IEEE Trans. Power Syst. **21**, 993–940 (2006)
4. Chicco, G., Napoli, R., Piglione, F., Postolache, P., Scutariu, M., Toader, C.: Emergent electricity customer classification. IEEE Gener. Trans. Distrib. **152**, 164–172 (2005)
5. Dorigo, M.: Optimization, Learning and Natural Algorithms (in Italian). Ph.D. thesis, Dipartimento di Elettronica, Politecnico di Milano, Milan, Italy (1992)
6. Egarter, D., Bhuvana, V.P., Elmenreich, W.: Paldi: online load disaggregation via particle filtering. IEEE Trans. Instrum. Meas. **64**(2), 467–477 (2015)
7. Egarter, D., Sobe, A., Elmenreich, W.: Evolving non-intrusive load monitoring. In: Esparcia-Alcázar, A.I. (ed.) EvoApplications 2013. LNCS, vol. 7835, pp. 182–191. Springer, Heidelberg (2013)
8. Eiben, A.E., Ruttkay, Z.S.: Constraint Satisfaction Problems. IOP Publishing Ltd. and Oxford University Press, New York (1997)

9. Gavrilas, M., Gavrilas, G., Sfintes, C.V.: Application of honey bee mating optimization algorithm to load profile clustering. In: 2010 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA), pp 113–118, September 2010
10. Goncalves, H., Ocneanu, A., Berges, M., Fan, R.H.: Unsupervised disaggregation of appliances using aggregated consumption data. In: Proceedings of KDD Workshop Data Mining Applications Sustainability (SustKDD) (2011)
11. Gonzalez-Pardo, A., Camacho, D.: A new csp graph-based representation for ant colony optimization. In: 2013 IEEE Conference on Evolutionary Computation, 20–23 June 2013, vol. 1, pp. 689–696 (2013)
12. Gonzalez-Pardo, A., Camacho, D.: A new csp graph-based representation to resource-constrained project scheduling problem. In: 2014 IEEE Conference on Evolutionary Computation (2014, in press)
13. Hart, G.W.: Residential energy monitoring and compcomputer surveillance via utility power flow. IEEE Technol. Soci. Mag. **8**(2), 12–16 (1989)
14. Jota, P.R.S., Silva, V.R.B., Jota, F.G.: Building load management using cluster and statistical analyses. Int. J. Electr. Power Energ. Syst. **33**(8), 1498–1505 (2011)
15. Khan, S., Bilal, M., Sharif, M., Sajid, M., Baig, R.: Solution of n-queen problem using aco. In: IEEE 13th International Multitopic Conference, INMIC 2009 (2009)
16. Kolter, J.Z., Johnson, M.J.: Redd: a public data set for energy disaggregation research. In: SustKDD Workshop on Data Mining Applications in Sustainability (2011)
17. Liang, J., Ng, S.K.K., Kendall, G., Cheng, J.W.M.: Load signature study - part i: Basic concepts, structure, and methodology. IEEE Trans. Power Deliv. **25**(2), 551–560 (2010)
18. Lin, G.-Y., Lee, S.-C., Hsu, J.Y.-J., Jih, W.-R.: Applying power meters for appliance recognition on the electric panel. In: Proceedings of 5th IEEE Conference Industrial Electronics Applications (ICIEA), pp. 398–405 (2010)
19. López, J.J., Aguado, J.A., Martín, F., Munoz, F., Rodríguez, A., Ruiz, J.E.: Electric customer classification using hopfield recurrent ann. In: Proceedings of 5th International Conference on the European Electricity Market (EEM08) (2008)
20. López, J.J., Aguado, J.A., Martín, F., Munoz, F., Rodríguez, A., Ruiz, J.E.: Hopfield-k-means clustering algorithm: a proposal for the segmentation of electricity customers. Electr. Power Syst. Res. **81**(2), 716–724 (2011)
21. Roli, A., Blum, C., Dorigo, M.: ACO for maximal constraint satisfaction problems. In: Proceedings of MIC'2001 - Metaheuristics International Conference, vol. 1, pp. 187–191. Porto, Portugal (2001)
22. Shaw, S.R., Leeb, S.B., Norford, L.K., Cox, R.W.: Nonintrusive load monitoring and diagnosis in power systems. IEEE Trans. Instrum. Meas. **57**(7), 1445–1454 (2008)
23. Solnon, C.: Ants can solve constraint satisfaction problems. IEEE Trans. Evol. Comput. **6**, 347–357 (2002)
24. Solnon, C.: Ant Colony Optimization and Constraint Programming. Wiley-ISTE, UK (2010)
25. Srinivasan, D., Ng, W.S., Liew, A.C.: Neural-network-based signature recognition for harmonic source identification. IEEE Trans. Power Deliv. **21**(1), 2254–2259 (2010)
26. Sum, Y.M., Wang, C.L., Zhang, Z.S., Liu, S.: Clustering analysis of power systems load series based on ant colony optimization algorithm. In: Proceedings of the Csee, pp. 40–45 (2005)

27. Tsang, E.P.K.: Foundations of Constraint Satisfaction. Computation in Cognitive Science. Academic Press, London (1993)
28. Tsekouras, G.J., Kotoulas, P.B., Tsirekis, C.D., Dialynas, E.N., Hatziargyrou, N.D.: A pattern recognition methodology for evaluation of load profiles and typical days of large electricity customers. Electr. Power Syst. Res. **78**, 1494–1510 (2008)
29. Valero, S., Ortiz, M., Senabre, S., Gabaldón, A., García, F.: Classification, filtering and identification of electrical customer load patterns through the use of self-organizing maps. IEEE Trans. Power Syst. **21**, 1672–1682 (2006)
30. Zaidi, A.A., Kupzog, F., Zia, T., Palensky, P.: Load recognition for automated demand response in microgrids. In: Proceedings 36th Annual Conference IEEE Industrial Electronics Society (IECON), pp. 2242–2447 (2010)