Renewable energy prosumers clustering towards target aggregated prosumption profiles based on recursive predictions

Ioannis Mamounakis(1), Dimitrios J. Vergados(1), Prodromos Makris(1) (2), Emmanouel Varvarigos(1) (2)
(1) Computer Technology Institute & Press "Diophantus"
Patras, Greece
(2) Institute of Communication and Computer Systems,
Athens, Greece
{mamounakis, vergados, pmakri, manos}@ceid.upatras.gr

Nikolaos D. Doulamis
Institute of Communication and Computer Systems
Athens, Greece
ndoulam@cs.ntua.gr

Abstract— **This paper introduces a decision making framework for aggregating Microgrids and/or other small energy producers and consumers (i.e. prosumers) into groups, whose purpose is to participate in liberalized electricity markets as single entities. The aggregator is able to offer aggregated Renewable Energy Source (RES) units to the wholesale market, in ways that are more efficient than individual prosumers acting alone. We first present the proposed framework and information flow among the involved market entities. We then focus on the problem of finding the set of prosumers whose aggregate prosumption profile can best fit a given target pattern requested by a market actor. We propose a linear autoregressive forecasting algorithm and a genetic clustering algorithm, which can easily adapt to the requirements set by the various use cases. Numerical results show that the aggregator can produce clusters in real time improving the average deviation from the target pattern by up to 50%.**

Keywords—aggregator; clustering; virtual microgrid; prediction;

I. Introduction

During the last few years, several countries have used subsidy policies in order to incentivize the generation of renewable energy. In order to allow small energy prosumers to participate and compete in the electricity market against traditional power generators, the concept of virtual micro-grids (VMG) is investigated in the framework of the Virtual Microgrids for Smart Energy Networks (VIMSEN) EU project [1]. A VMG is defined as an association of rather small and distributed energy prosumers, created for the purpose of participating in a liberalized electricity market [2]. These small or very small actors (the prosumers) agree to operate on a common basis, without the need to install advanced smart grid equipment, such as storage, advanced PVs, advanced DR equipment or advanced communication routers. Instead, their functionalities and capabilities are enhanced by forming coalitions and collaboratively participating in the market under the efficient management of a Virtual Microgrid Aggregator (VMGA). In the VIMSEN context, an aggregator (VMGA) is responsible for orchestrating the prosumers located in a given area (whose definition and size depends on the use case) into clusters, thus creating energy groups that are better placed to offer reliable services to different markets, and serve different market players (i.e. Market Operator, Distribution System Operator etc.). As a result, a Virtual Microgrid in VIMSEN is slightly different from the traditional Microgrid and the Virtual Power Plant concepts [3], even though there are common functionalities, as it involves both production and consumption (prosumption) and tries to exploit statistical multiplexing and market participation advantages coming from the aggregation of very small actors [4].

The different electricity markets typically operate at various time intervals prior to the time of delivery. After a buy or sell bid has been accepted in the market, the participant is committed to consume or deliver the traded amount of energy. Thus, efficient market participation requires an estimation of the amount of energy that the aggregator (i.e. on behalf of its registered prosumers) will need to consume, and also the amount of energy that it will be able to contribute to the system at each time interval. Therefore, accurate forecasts of the production and consumption for each prosumer must be available at the aggregator's side at times prior to the closing time of each market [5].

The large amount of recent literature in the area shows that the accuracy of the energy forecasts is not only an interesting research problem, but is also of great economic significance. In the case of consumption forecasting, models based on historical data may be used, as these load patterns are relatively easily predictable. However, in the case of RES production forecasting, more advanced and accurate forecasting methods are required [6]. As presented in [7], most forecasting tools for renewable energy production are based on Neural Networks. As a result, they pose large computational burden, which would prohibit the aggregator's participation in near-real-time events, given the relatively light processing capabilities of the low-cost energy gateways residing at the prosumer's premises. In a related context, in order to reduce RES production fluctuations during the day, several studies propose load prediction techniques accompanied by batteries and/or demand response (DR) solutions to reduce the hourly deviations. In [8], the authors propose a system to program the day-ahead output of a hybrid RES system comprising a wind-turbine and PV panels based on output prediction, using a battery to smoothen deviations. In [9], a battery is used to minimize the grid power peaks created by a RES microgrid. In order to achieve that, the system forecasts the day-ahead power generation and the consumption load in the microgrid and runs in an operational setting using 1-min resolution data, to incorporate a dynamic load target threshold towards which the load can optimally be reduced. The system limitations include the battery performance efficiency and the actual and predicted power outputs, which may lead to a loss of battery control. In [10], authors propose a technique that makes use of forecasting data, and apply a central moving average (CMA) strategy, eliminating any lag in the response and reducing the battery needs or, in other words, getting a better grid power profile with the same battery system. This approach, however, is subject to the accuracy of the forecast data. In order to handle forecast errors, the proposed strategy makes corrections in the power exchanged with the grid, taking into account not only the forecasting error, but also the battery state of charge. Authors in [11] propose a framework based on a unit commitment (UC) problem aiming to minimize system cost. The proposed demand model captures the aggregated behavior of a large population of small-scale RES loads equipped with PV-battery systems, including appropriate energy management systems. The effect of the demand model on the load profile of each residential and commercial prosumer shows that increasing the penetration of DR, improves controllability in the presence of increased intermittent supply penetration in the grid. In [12], a Demand-Side Management (DSM) optimization-based model is proposed that is applicable to Electric Vehicles (EVs), considering that EV load can be shifted based on charging prices. In the proposed model, agents are responsible for load, generation and storage management. Finally in [13] a methodology to assess the impact on the demand profiles due to the charging of plug-in electric vehicles in distribution networks is presented.

The above studies assume that end users are equipped with smart meters, home energy management systems for resources scheduling, and batteries in order to achieve the DR target. Such devices, however, increase the total cost for each prosumer. Despite the aforementioned methods, it is essential for the energy market to focus on low-cost (e.g. simple smart meters, communication gateways) system solutions for the low end market. The premise of the current paper is that the use of lightweight systems by small prosumers, combined with the decentralized intelligence offered by aggregators, is sufficient to enable them to participate in the electricity market and compete with traditional power generators. Towards this end, we present (i) an auto regressive forecasting method and (ii) an optimized matching genetic algorithm schema for dynamically selecting a set of prosumers that matches, on aggregate, a given (target) distribution published by the wholesale market operator (MO) or some other actor such as the Balance Responsible Parties (BRP) or DSO. The reliability of the energy delivery is increased through both aggregation and forecasting. Since the forecasts' accuracy heavily depends on the amount of time between the calculation of the forecast and the delivery time, a near real time prediction technique is executed for the efficient estimation of RES prosumption. The main contributions of this paper are:

- The applicability of the proposed solution to small-scale RES prosumers who operate with basic smart grid equipment installed at their premises, such as low-cost gateways that are able to acquire energy prosumption datasets. There is no need for complex and expensive equipment for DR, batteries for storage, advanced micro grid controllers or any advanced communication between the aggregator and the prosumers.
- Through the dynamic clustering, the smoothening of the profiles through aggregation and accurate forecast estimations make the application of storage devices or DR flexibility less necessary to reduce fluctuations.
- The implementation of complex auctioning techniques in not necessary as we assume that a private fixed SLA exists between the VMGA and each one of the prosumers [14].
- Based on the prosumption forecasts, the aggregator executes the target matching algorithm in order to quantify the cluster of prosumers that will respond to a specific demand response event. Genetic clustering algorithm is preferred among other methods as its performance has already been tested in our previous work [4].

II. SYSTEM MODEL & PROBLEM STATEMENT

RES generation is qualitatively very different compared to conventional fossil energy production, mainly due to the positive externalities (e.g. green credits, locality benefits) but also the negative externalities (increased variations) that the former creates. An aggregator acts as a middleware layer between the electricity market and the small RES prosumers providing an interface of determining RES price based on the current demand-supply requirements and the value of RES externalities. This results in an independent operation between the traditional and the RES market. In the following, we assume that RES market acts as a split market from the conventional electricity wholesale market [15]. In this way, it is possible to internalize the green credits stemming from RES production, while simultaneously accounting for the demand and supply requirements. In the context of a split (independent but interacting) RES energy market, a request of prosumption $r(t), \forall t \in [t_a, t_b]$ RES units is assumed to be asked by the wholesale MO. The same procedure is followed during the energy request by DSO and BRPs. In most cases, demand and generation bids are collected in the wholesale electricity market, which finally determines the generation to activate. However, special conditions in the distribution network lead the MO to request specific DR in order to keep the network stable. Variables t_a and t_b denote start and finish time of forward time intervals.

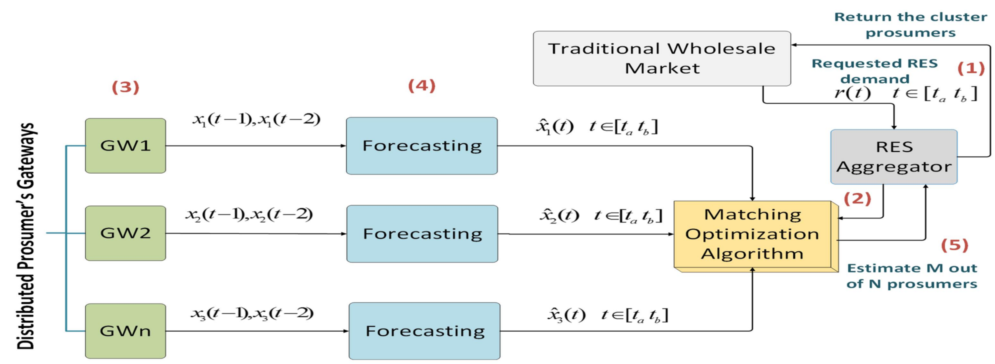

Figure 1. The proposed architecture for prosumer clustering in order to satisfy the requested RES demand for the ahead time interval $r(t), \forall t \in [t_a\ t_b]$.

This amount is requested to the RES market operator (or to the RES aggregator if such a market operator does not exist), which aims at matching the demand bid. This request may involve a specific area (in which case only prosumers in that area will be allowed to contribute) or be nationwide. This means that the request of $r(t), \forall t \in [t_a, t_b]$ is a time series of RES request for the time ahead interval $[t_{a,}\ t_b]$. In general, the $r(t)$ value could be a time series of near-real-time DR units request by the DSO to deal with an imminent congestion management problem or in other cases a time series of near-real-time DR units request by the BRP to deal with an imminent balancing problem. In this paper, we only deal with the demand decrease scenario. The RES market operator tries to match the requested demand by employing the services of a RES aggregator (or of multiple aggregators). Let us assume that a set of N small prosumers (i.e., units that are capable of both producing and consuming electricity) are available (nationwide or in the area specified in the request). The ultimate goal of the RES aggregator is to apply a matching algorithm in order to estimate the M<N out of N available prosumers that collectively better fit the requested ahead target profile. In order to perform demand matching, the RES aggregator needs to know the forecasted prosumption (excess of energy) for each prosumer i=1 ,.., N.

Table 1: Proposed Algorithm steps

Step #1: A demand prosumption request $r(t), \forall t \in [t_a, t_b]$ for the time ahead interval $[t_{a,}\ t_b]$ is provided to the RES aggregator by the wholesale traditional electricity market.
Step #2: The aggregator forwards the demand request to the matching optimized interface.
Step #3: The gateways measure the excess of energy until the current time instance t_c.
Step #4: The forecasting interface estimates $\hat{x}_i(t)$ for the time ahead interval $[t_{a,}\ t_b]$ exploiting previous actual measurements $x_i(t), \forall\ t \leq t_c$.
Step #5: The matching algorithm, exploiting a genetic optimization schema, selects a set of optimal *M* out of *N* available prosumers whose aggregate prosumption profile best satisfies the requested target profile.

As a result it is necessary to know:

$$x(t) = cons_i(t) - prod_i(t)\ \forall\ t \in [t_a, t_b] \qquad (1)$$

where $prod_i(t)$ and $cons_i(t)$ denote the energy production and consumption, respectively, of prosumer p_i versus time, for the time ahead interval $[t_{a,}\ t_b]$. The amount $x_i(t)$ is forecasted by the gateway GW_i of prosumer p_i. In particular, each gateway GW_i is aware of the prosumption $r_i(t)$ at time *t*. Then, the forecasting module embedded in the gateway is responsible for providing estimates of $x_i(t)$ for the time ahead interval $[t_{a,}\ t_b]$. It is clear that the better the forecast is, the more closely the matching algorithm will be able to satisfy the requested target profile $r(t), \forall t \in [t_a, t_b]$. We introduce a forecasting method and an optimized matching algorithm for dynamically selecting (clustering) a set of *M<N* out of *N* available prosumers, that collectively best satisfy the requested target profile $r(t)$ based on forecasting estimates of the prosumption $\hat{x}_i(t), \forall t \in [t_a, t_b]$. The clustering process is performed after a signal from the MO has received and before the delivery time. The forecasted estimates $\hat{x}_i(t)$ are related with the actual prosumption measurements $x_i(t)$ through $x_i(t) = \hat{x}_i(t) + \varepsilon(t), \forall t \in [t_a, t_b]$ where $\varepsilon(t)$ denotes the prediction error. Then, the forecasting approach aims at minimizing the error $\varepsilon(t)$ based on an adopted model. The proposed architecture depicted in Table I, provides a description of the main steps of the proposed algorithm for the dynamic clustering of prosumers in order to collectively satisfy the requested target profile.

III. PROBLEM FORMULATION

A. The Forecasting Model

Let us recall that $x(t)$ is the measured prosumption (excess energy) from a prosumer's gateway. Based on previous actual measurements of $x_i(t),\ \forall\ t \leq t_c$ the forecasting interface provides time ahead predictions. In the following, a linear autoregressive moving average model is adopted to relate the previous and the ahead energy excess units. Formally, an autoregressive moving average model is written as follows:

$$x(t) = \sum_{i=1}^{p} a_i x(t-i) + \sum_{i=0}^{q-1} b_i y(t-i) + \varepsilon = \hat{x}(t) + \varepsilon \qquad (2)$$

In Eq. (2), variables p and q denote the model order both for the autoregressive and the moving average term, while variables a_i and b_i denote the model coefficients. Actually, Eq. (2) expresses that the values of the energy prosumtion at a time instance t is related with the values at previous time instances plus some external factor [modeled through variables $y(t-i)$. In Eq. (2), we have denoted as $\hat{x}(t)$ the predicted prosumption at the time ahead instance t by the model. Usually, coefficients a_i and b_i are estimated to minimize the error $\varepsilon(t)$ of the prediction, thus the error between $x(t)-\hat{x}(t)$ the actual and the predicted value. Several optimization algorithms exist in the literature for estimating the coefficients a_i and b_i of the model. Examples include the Yule-Walker [16] approach where a system of linear equations are exploited to estimate a_i and b_i, or the Burg's lattice method [16] that solves the lattice filter using the harmonic mean of forward/backward squared prediction errors, or even the geometric lattice approach, where the geometric mean instead of harmonic mean is exploited. The model of (2) actually assumes a temporal coherency among the prosumption values. This implies a temporal smoothness in the energy production, which is actually a valid statement for RES generators. Usually, there are no abrupt changes in the RES generation and therefore assumption of temporal continuity is valid.

B. Improving Clustering Through Forecasting

Forecasting can greatly improve clustering decisions and consequently the efficiency of the overall scheme. This is because the grouping of prosumers so as to achieve a given target prosumption function is performed, in the absence of forecasts, based on previous prosumption measurements. Thus, variations in the actual prosumption (mainly due to high RES production variability) make it highly probable that the selected prosumer group (VMG) will not be able to satisfy the targeted function. Forecasting allows better clustering decisions, by basing the grouping decisions on the predicted prosumption rather than on its previous measurements. In case that no forecasting is exploited, the matching optimization interface adopts a differentiated pulse-coded modulator (DPCM) approach. This means that the error between two successive samples of $x_i(t)$ are encoded. The difference is used to predict future samples of $x_i(t)$. In case of incorporation of a better prediction method, we exploit better forecasts and therefore better performance of the optimization matching interface.

C. Genetic algorithm approach for optimal clusters

Prosumer i, $1 \leq i \leq N$ is characterized by his prosumption equation (1) where t is the index of the time period. Variables $cons_i(t)$ and $prod_i(t)$ correspond to the energy consumption and production, respectively, for the given prosumer and time period. We define an indicator variable $z_i \in \{0,1\}$, that indicates whether prosumer i is part of the cluster or not. The aggregate prosumption for the cluster is:

$$R(t) = \sum_{i=1}^{N} z_i \times x_i(t) \qquad (3)$$

The objective of the clustering algorithm is to search for a cluster that has an aggregate prosumption close to a given target pattern. If we denote the target pattern as $g(t)$, then we try to find a cluster where the mean square error between the target pattern and the aggregate prosumption of the cluster is as small as possible. In other words, given the sequences $x_i(t)$ and $g(t)$, we obtain values z_i that minimize the expression:

$$\sum_{t=1}^{T}(R(t)-r(t))^2 = \sum_{k=0}^{n}\left(\sum_{i=1}^{N}(z_i \times x_i(t)) - g(t)\right)^2 \qquad (4)$$

This problem appears to be NP-hard at least for the general case of VMGA's operation in which a huge prosumers' portfolio and near-real-time responses are assumed [4].

1) The genetic algorithm

We can use a genetic algorithm to search for solutions to the above combinatorial problem. The genetic representation follows logically from the structure of the problem. We define the structure of a chromosome as a vector z of binary values

$$z_i \in \{0,1\}, 1 \leq i \leq N \qquad (5)$$

If z_i is equal to zero, then the corresponding prosumer i is not a member of the cluster, and its prosumption is not counted. Otherwise prosumer i is part of the cluster, and its prosumption is aggregated in the total prosumption of the cluster.

2) The fitness function

The fitness function also arises logically from the problem definition. However, since the optimization engine that was used for the simulations tries to maximize the fitness function (instead of minimizing it), we use the opposite of Eq. (4). Thus we define the fitness function of a solution z as

$$-\sum_{t=1}^{T}(\sum_{i=1}^{N}(z_i \times x_i(t)) - g(t))^2 \qquad (6)$$

If the cluster's aggregate prosumption matches the target exactly, then the above fitness function will produce a value of zero, which is the largest possible value. If a perfect matching is not found, the fitness of a solution will be negative, with a value proportional to the mean square error between the solution and the target.

3) The mutation function

The mutation function defines the way that new solutions will be generated from an existing solution. In the mutation function that we implemented, mutation corresponds to the random swapping of two consecutive genes.

4) The reproduction function

The reproduction function defines the way a new offspring solution is generated from two parent solutions.

In our implementation we used the two-point crossover method.

IV. Numerical Results

A. Simulation Scenarios and setup

In order to evaluate the performance of the proposed algorithm, we implemented it in the Decision Support System (DSS) component of the VMGA that was developed for the VIMSEN project. The DSS system is already able to run and execute in an efficient way the matching algorithm. This allowed us to use data from 40 real life solar prosumers to measure the efficiency of the algorithm. The methodology of our evaluation is the following. We generated "demand response" (DR), events with a random starting time and duration of 15-minute blocks. Each DR event specifies an amount of energy (kWh), for each 15-minute block, which is requested by the DSO, the BRP or any other market participant. The request is made available 15-minutes before the start of the first 15-minute interval of the delivery period. The aggregator requests the gateways for their forecasts of their energy availability for the duration of the DR event. Using these forecasts, aggregator executes the target matching algorithm in order to create the cluster of prosumers that will respond to the event. For the purpose of this evaluation, we used DR events that were not concurrent. In an actual implementation, one would have to take into account any commitments that the prosumers have already made in different markets or other (previous) DR events. For every request, we evaluated the response of the system under two scenarios: using an AR model for generating the forecasts, and using the "latest known" measurement of the meter as estimate for the next timeslots (no forecasting).

We tested different values for the order of the AR model, from 10 up to 100. We performed 50 DR events for every value of the order parameter that was tested. To provide a graphical view, in Fig. 2 the response of the system for three different DR events is presented. The red line represents the demand pattern that is requested. The green and purple lines are the predicted patterns of the selected cluster, using and not using the AR forecasting model respectively. The dark and light blue lines are the actual achieved prosumption over the period with and without forecasts respectively. We observed that in most experiments the dark blue line is closer to the red line than the light blue line is to the red line. This indicates that the forecasting algorithm helps in achieving a better match of the requested demand pattern than the case when no forecasting algorithm was used. In Fig. 3, we can see the mean absolute error of the actual prosumption vs. the target for each 15-minute block. When the forecast values are used, the system achieved a smaller mean absolute error, with a smaller variance. We can see that the order parameter affects the accuracy of the forecasts, with an order of 100 resulting to the best accuracy on these experiments. The forecasting algorithm can improve the deviation from the target on average up to 50%, but considering the large variance in the estimation error, the gains may be much higher in some cases.

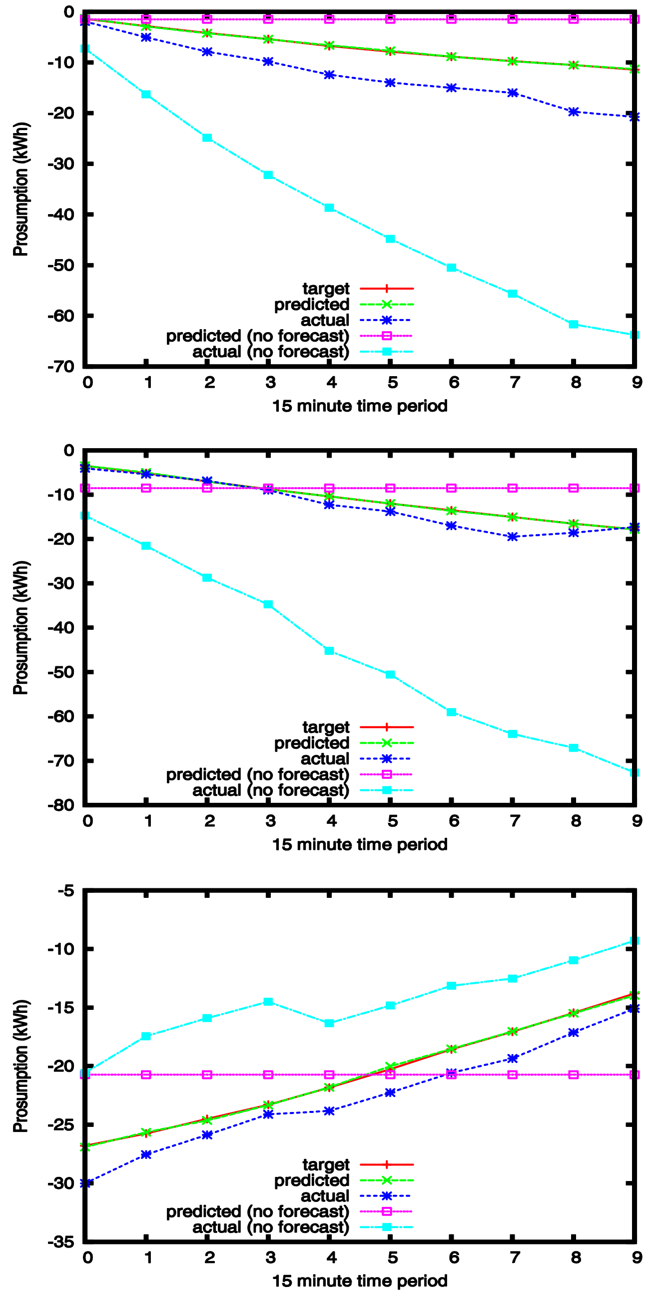

Fig. 2: The requested demand versus the demand achieved by the selected cluster, with and without forecasts

I. Conclusion

In this paper, we presented a scheme for enabling energy prosumers to participate in real-time energy markets. We introduce an entity called aggregator, which is responsible for representing the prosumers in the energy markets. The aggregator is capable of receiving "Demand Response" events from other energy market participants, and to respond to them, by creating a cluster of prosumers that will respond to the published event. We used the autoregressive (AR) model for generating the energy prosumption forecasts.

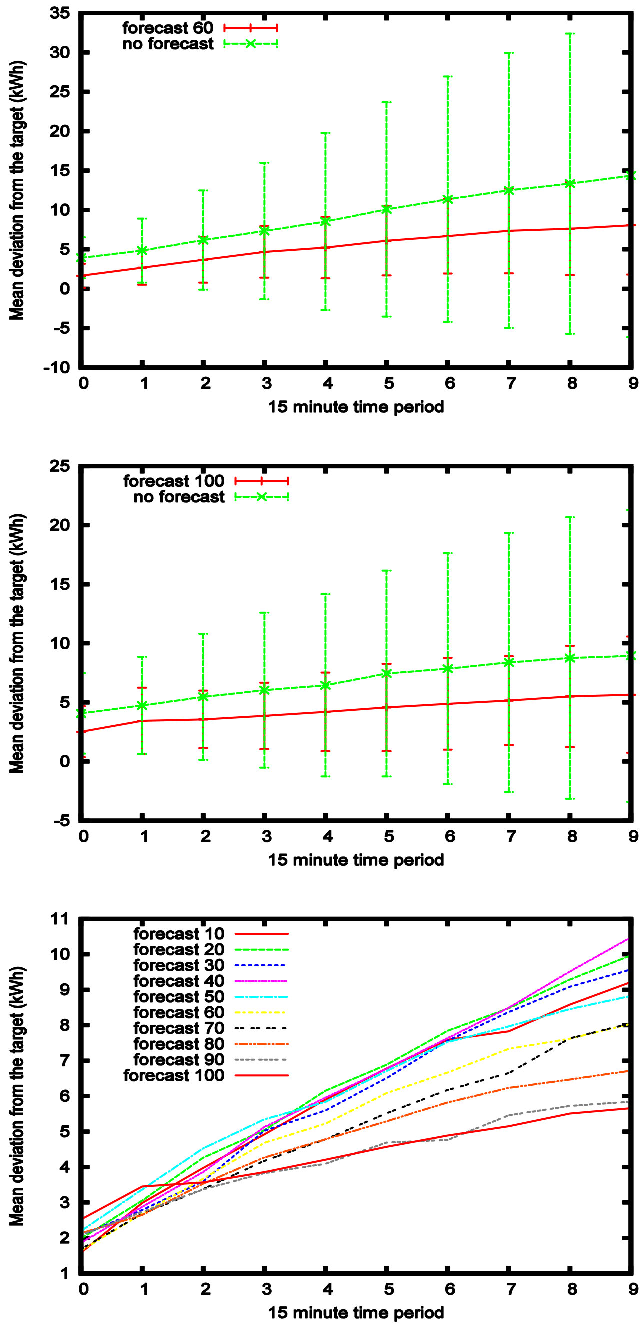

Fig. 3: The mean absolute error of the actual prosumption vs. the target for different values of the order parameter.

The identification of the optimal set of prosumers to respond to the event is a computationally complex problem, so we introduce a heuristic genetic algorithm for finding near-optimal solutions. Using real prosumption data, we evaluated the algorithms performance and found that the target matching algorithm is generally capable of finding a good set of prosumers, given accurate forecasts. Due to the inaccuracy of the forecasting algorithm, a deviation exists between the target prosumption pattern and the one actually achieved by the prosumers. This deviation, however, is smaller than what would be achieved if no forecasting algorithm is used.

Acknowledgment

The work presented in this paper received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 731767 in the context of the SOCIALENERGY project. The authors would also like to acknowledge the contributions of their colleagues from the EU-funded FP7 VIMSEN consortium (Grant Agreement No. 619547).

References

[1] https://www.ict-vimsen.eu

[2] Ioannis Mamounakis et al. , "A virtual microgrid platform for the efficient orchestration of multiple energy prosumers", Proceedings of the 19th Panhellenic Conference on Informatics, pp. 191-196, 2015.

[3] Peter Asmus, Microgrids, "Virtual Power Plants and Our Distributed Energy Future", The Electricity Journal, vol. 23, pp. 72-82, 2010.

[4] Dimitrios J. Vergados et al. , "Prosumer clustering into virtual microgrids for cost reduction in renewable energy trading markets", Sustainable Energy, Grids and Networks 7, pp. 90-103, 2016.

[5] N. Damsgaard et al. , "Study on the effective integration of distributed energy resources for providing exibility to the electricity system", Tech. Rep. 54697590000, European Commission, 2015.

[6] Rich H. Inman, et al., "Solar forecasting methods for renewable energy integration Progressin Energy and Combustion", Science 39, pp.535-576, 2010.

[7] Comodi G, Giantomassi A, Severini M., "Multi-apartment residential microgrid with electrical and thermal storage devices: experimental analysis and simulation of energy management strategies", Appl Energy vol. 137, pp.854–66, 2015.

[8] Marinelli M, Sossan F., "Testing of a predictive control strategy for balancing renewable sources in a microgrid", IEEE Transactions Sustainable Energy vol. 5, pp. 1426– 33, 2014.

[9] Hanna R, et al., "Energy dispatch schedule optimization for demand charge reduction using a photovoltaic-battery storage system with solar forecasting", Sol Energy, vol. 103, pp. 269–87, 2014.

[10] Julio Pascual, et al., "Energy management strategy for a renewable-based residential microgrid with generation and demand forecasting", Applied Energy, vol.12 , 2015.

[11] Hesamoddin Marzooghi, GregorVerbič ,David J. Hill, "Aggregated demand response modeling for future grid scenarios", Sustainable Energy, Gridsand Networks vol. 5, pp. 94–104, 2016.

[12] M.A. López , et al., "Demand-side management in smart grid operation considering electric vehicles load shifting and vehicle-to-grid support", Electrical Power and Energy Systems vol. 64, pp. 689–698, 2015.

[13] Andres F. Botero, and Mario A. Rios, "Demand Forecasting Associated with Electric Vehicle Penetration on Distribution Systems", IEEE PowerTech, Eindhoven, The Netherlands, 29 June-2 July 2015.

[14] J. van der Burgt, G. Sauba, E. Varvarigos and P. Makris, "Demonstration of the Smart Energy Neighbourhood Management System in the VIMSEN Project", IEEE PowerTech, Eindhoven, The Netherlands, 29 June-2 July 2015.

[15] J. van der Burgt, "Intermediate Version of Business Models and Pricing Mechanisms within the Virtual Microgrids", VIMSEN project, Available:ict-vimsen.eu/images/Open_Data/public%20deliverables_pdf /vimsen_d8.2.2_final_v2_24112015.pdf, 2015.

[16] Petre Stoica, Randolph Moses, "Introduction to Spectral Analysis", Taylor & Francis, Ltd., American Statistical Association, American Society for Quality, 2005.