Coalition Formation Algorithm of Prosumers in a Smart Grid Environment

Nicolas Gensollen, Vincent Gauthier, Michel Marot and Monique Becker
CNRS UMR 5157 SAMOVAR,
Telecom SudParis/Institut Mines Telecom
Email: {nicolas.gensollen, vincent.gauthier, michel.marot, monique.becker}@telecom-sudparis.eu

Abstract—In a smart grid environment, we study coalition formation of prosumers that aim at entering the energy market. It is paramount for the grid operation that the energy producers are able to sustain the grid demand in terms of stability and minimum production requirement. We design an algorithm that seeks to form coalitions that will meet both of these requirements: a minimum energy level for the coalitions and a steady production level which leads to finding uncorrelated sources of energy to form a coalition. We propose an algorithm that uses graph tools such as correlation graphs or clique percolation to form coalitions that meet such complex constraints. We validate the algorithm against a random procedure and show that, it not only performs better in term of social welfare for the power grid, but also that it is more robust against unforeseen production variations due to changing weather conditions for instance.

I. INTRODUCTION

One of the key ideas in the smart grid revolves around the introduction of communication means inside the power grids. This could enable complex improvements in the energy management and leads progressively to a greener energetic system [1] [2]. Distributed energy resources (DER) such as wind turbines or photovoltaic panels are not supposed to emerge only in remote farms, but also in residential areas. Together with electric vehicles, and demand side management tools, they will constitute the building blocks which will help to turn the today pure energy consumers into true actors of the grid operation [1]. Such agents that both consume and produce energy are ready to make concessions (appliances delays, V2G) to ensure grid stability, and are commonly called prosumers [3] [1].

There seems to have a clear consensus on the benefit of having bidirectional communication flows between the prosumers and the grid, as demand side management relies on such an architecture [2]. Nevertheless as the number of active prosumers is expected to rise, it is safe to assume there is a need for more complex communication patterns (such as coalitions) that will help to decrease the communication burden and satisfy the multiple requirements of the power grid management. The formation of coalitions inside the smart grid environment could be applied to various types of agents such as self sufficient micro-grids [4], sizable and adjustable virtual power plants (VPP) [1], [5], or fleet of electric vehicles that back up the grid in some emergency situations (V2G) [1]. These are only a few use cases where the coalition of multiple agents can enhance the grid reliability and in the mean time reduce the communication burden.

In this paper, we focus on how to statistically improve the production stability by carefully forming coalitions of prosumers that have greater probabilities of staying in acceptable range of production. By using prediction techniques, we form coalitions that meet the contract values ranges proposed by the grid, enabling it to schedule its production accordingly. Meeting contract values for the prosumers is especially relevant in power trade conditions, where energy is traded based on day-ahead predictions. In this context, participants should try to minimize their prediction errors in order to maintain a stable state for the grid. They may even occur some penalties if the productions deviate from the initial contract values. However, it remains pertinent for both side to maintain the production as stable as possible : for the prosumer, the stability (with renewable energy sources) will impact its net gain, and for the power grid, it will help to maintain its reliability.

In this paper, we seek to form coalitions that are able to announce "high contract values with high reliability". We thus aim at:
- Building a realistic prosumer model with renewable production sources based on weather data (see section III-A).
- Define a coalition formation model that enables the grid to set some requirements under which any group of prosumers will be allowed to sell its production (see section III-B)
- Define a utility function that will satisfy the grid requirements and maximize the stability of the prosumer coalition productions (see section III-C)
- Define a coalition formation algorithm (see section IV)

In section III, we consider that the agent's production depends on meteorological conditions (wind turbines and PV as generators), various energy mix and preferences, and their appliances (loads). We use meteorological data (see section III-A) to account for seasonal and daily variations in the prosumer's energetic profiles as well as realistic geographical correlations between different agents. With these ingredients, we are able to run realistic simulations and record the different output profiles as time-series.

In order to select groups of prosumers that lower the volatility of the coalition's productions, we define an algorithm based on a popular approach for stocks market clustering [6]. We use a modified clique percolation algorithm (see section IV) that will enable us to expand the cliques as needed in order

to form proper coalitions that fulfill both the grid requirements and lower the production volatility. Finally, section V provides some results and a conclusion in section VI.

II. RELATED WORK

Traditionally, forming coalitions in a pool of agents can be done either in a centralized way where a single central unit is responsible for all the computations or in a distributed way where agents have only local knowledges and take actions accordingly. It is of common use to represent the situation and assess the stability of the solution by using game theoretic tools. Some papers [7] [8] focus then on finding an optimal coalition structure giving a pool of autonomous self interested agents using distributed merge/split algorithms.

Tackling the stability issues of renewable DER, the TradeWind project [9] simulated the impact of wind power on electricity exchange and cross-border congestions by using a flow-based market model. The idea revolves around identifying key european interconnections (already existant or not) in order "to make optimal use of the european spacial de-correlation of wind power". It was indeed shown that geographical aggregation provides smoothening effects and that the amount of prediction errors for wind power in a geographical region diminishes as the region size increases, especially for short forecast horizons.

On a narrower scale, the authors of [10] study the formation of virtual power plants (VPP) composed of multiple self-interested DER. On the grid side, two requirements for the formation of virtual power plants are considered : the reliability of supply and the minimization of entities the grid has to deal with. From this, [10] builds a pricing mechanism that encourages VPP to report true estimates of their aggregated production and penalizes prediction errors. A redistribution scheme of the VPP to the DER is also constructed such that the payoff allocation lies in the core of the game, meaning that no DER has an incentive to leave the coalition.

In this paper we ought to form coalitions utilities that depend on statistical properties of time-series values. The setting is thus similar with some financial studies on stock exchanges, where researchers tried to find relevant clusters of stocks based on their daily prices variations. The problem of clustering is usually approached by means of a similarity measure and completed by a clustering technique such as K-means or hierarchical clustering. Nevertheless, in [6] the author introduced an approach where stocks time-series of their daily log returns, are organized in a graph such that stocks exhibiting similar price fluctuation patterns are close to each other. This closeness notion is formalized with a similarity measure based on Pearson correlation coefficient ($d_{ij} = \frac{1}{2}\sqrt{2(1 - \rho_{ij})}$ or $d_{ij} = 1 - \rho_{ij}^2$) that enables to weight the edges of the graph.

ϵ-graphs, consists in filtering edges based on their weight, only keeping edges whose weights are less than ϵ. In [11], [12], the authors studied the topological properties (average clustering, connectivity, relative number of cliques) of the correlation graph against those of growing random graphs, depending on the threshold ϵ. However, there is no well defined method to select a right tradeoff ϵ as a function of the network topology.

Presented in this way, the time-series clustering task seems very close to graph community detection. Communities in networks are indeed often seen as groups of nodes exhibiting high internal densities of links as well as a low density across communities [13]. Although several techniques exists (based on different graph properties : modularity [13], edge betweenness [14], spectrum of graph Laplacian [15]), the clique percolation algorithm [16] uses directly this observation and employs a greedy expansion rule based on a fitness function. When the algorithm stops, some coalitions may be partially overlapping, meaning that a node can belong to multiple communities. Detection of overlapping communities is actually a very active field of research, especially in social networks where a person might belong to several communities.

III. MODEL

A. Prosumer model

Our major concern was to design a prosumer model with realistic patterns of consumption, production as well as realistic geographical correlations between them. We based our prosumer model on French's weather data from 2006 to 2012 sampled every 3 hours [17] (similar data are available for the United States covering 2010 [18]).

As shown in the first part of figure 1, the first step consists in discretizing the studied zone around well chosen weather stations and gathering traces for these stations. We consider here that prosumers can only produce through wind turbines and photovoltaic panels (PV). On the other hand, we built their consumption patterns according to two major cycles :

- **Daily cycles** : Consumption is low during night, and higher during the day with two picks in the morning and evening. Some noise is added so that prosumers have similar but not identical cycles.
- **Seasonal cycles** : Consumption is higher in the winter because of heating and low in the summer (air conditioning is not considered). Temperature traces were the principal ingredient for modeling these cycles.

We thus collect for all chosen weather stations three kinds of data (average wind speed, cloudiness, and average temperature), which we will call a climate vector in the following. We consider that climate vectors are constant over their area, meaning that, if two agents are in the same area, they are exposed to the same climate vector.

At this point, modeling agents consists in fixing a few parameters (most of the time drawn from random distributions) such as the geographical position, the number of wind turbines, PV, appliances, the temperature of confort, and so on. The objective is that, depending on the weather of his zone, the DER and appliances he owns, and the way he decides to

Figure 1: Process diagram

of our process. Thereby, we remove these seasonal effects for all agents (see second block of figure 1). We denote by $\mathcal{T}_i = \{P_i(t_0), ..., P_i(t_K)\}$ the resulting time-series for agent i, and by $\mathcal{F}(P_i)$ the corresponding probability distribution.

We now extend these notations to any coalition $s \subset \mathcal{A}$:

- $P_s(t) = \sum_{i \in s} P_i(t)$
- $\mathcal{T}_s = \{P_s(t_0), ..., P_s(t_K)\}$ with $\mathcal{F}(P_s)$ the corresponding probability distribution
- A coalition S has the possibility to announce a contract value P_s^{CRCT} on the market.

Due to the volatile nature of DER productions or user consumptions, the available power of a coalition will oscillates around P_s^{CRCT}. It is of the responsibility of the information system that forms coalitions to select agents in order to provide a production as stable as possible. Based on the $\mathcal{F}(P_s)$ distribution, the probability of s value being below its contract value $Pr[P_s \leq P_s^{CRCT}]$ appears therefore as a quantity of interest that we would want to keep as low as possible.

More formally, the rest of the paper derives a heuristic that aims at solving the following problem (Table I summarizes the main notations):

$$\underset{\substack{S \subset CS \\ |S| = N_{COAL} \\ \forall s \in S, \ |s| \neq 0 \\ P_s^{CRCT} \geq P^{MIN}}}{\arg \min} \sum_{s \in S} \Pr\big[P_s \leq P_s^{CRCT}\big] \qquad (1)$$

We explain now the two requirements that will help us solve this problem. Let $\phi \in [0, 1]$ be the **reliability** threshold imposed to the coalitions. It stipulates that for any coalition s willing to join the market, the probability of s value (at any instant t) being below its contract value should at most be ϕ. That is, $\forall s \in 2^N$, $Pr[P_s \leq P_s^{CRCT}] \leq \phi$ (ϕ being restricted to small values for consistency).

We also consider the **minimum contract value** constraint denoted by P^{MIN}, which states that only coalitions with higher contract values ($P_s^{CRCT} \geq P^{MIN}$) will be accepted.

At this point, if a coalition s wishes to join the market, it has to choose a contract value that fulfill the two conditions. For simplification, we consider that coalitions will always apply the same economically consistent strategy of announcing the highest possible contract value that obeys the reliability rule (a value we denote by $P_\phi(s)$). If $P_\phi(s) \geq P^{MIN}$, meaning that it also obeys the minimum contract value rule, then s announces this value on the market : $P_s^{CRCT} = P_\phi(s)$. Otherwise, coalition s is not able to enter. Basically, a coalition s is valid if and only if :

$$\begin{cases} Pr[P_s \leq P_\phi(s)] \leq \phi \text{ (reliability rule)} \\ and \ P_\phi(s) \geq P^{MIN} \text{ (min value rule)} \end{cases} \qquad (2)$$

C. Utility function

We now choose a utility function (eq. 3) that derives directly from the above remarks. If a coalition cannot provide a valid contract value, it receives naturally a utility of zero. Furthermore, it seems obvious that the utility should increase with the contract value. The $1/|s|$ term in eq. 3 indicates that

heat his home, a given prosumer is able to compute his production/consumption at any time.

We thus used existing models enabling conversion between weather variables and output powers (see power curves in [10] [19]). [20] provides also a convenient way of using cloudiness traces as realistic degradation factors in a *"clear blue sky"* solar model, which enables us to rebuild realistic solar irradiance traces for stations that do not provide such information directly. Any agent is thus able to compute at any time how much he produces and consumes.

B. Coalition formation model

More formally, we denote by $\mathcal{A} = \{a_1, ...a_N\}$ the set of agents (the prosumers) and $P_i(t)$ the instantaneous power value at agent i at instant t (its instantaneous production minus his instantaneous consumption, meaning that $P_i(t)$ represents agent i available power at instant t). During the simulation (from t_0 to t_K), all agents record their values with an hour time interval. As expected, the prosumer's time-series exhibit high seasonal patterns, but completely different from agent to agent because they depend both on the energetic mix and the habits. Nevertheless, these macro variations can hide the interesting information in the correlation coefficients and twist the rest

$\mathcal{A} = \{a_1, ..., a_N\}$	Set of agents
$P_i(t) \in \mathbb{R} \ (P_s(t))$	Instantaneous production of agent i at time t
$\mathcal{T}_i = \{P_i(t_0), ..., P_i(t_K)\}$	Time series of agent i available production values
$\mathcal{F}(P_i) \ (\mathcal{F}(P_S))$	Probability distribution of the production of the agent i
$\phi \in [0, 1]$	Reliability threshold
$P^{MIN} \in \mathbb{R}^+$	Min value threshold
$P_s^{CRCT} \in \mathbb{R}^+$	Contract value of coalition s
$P_\phi(s) \in \mathbb{R}^+$	Max contract value under ϕ constraint
$\mathcal{U}_{\phi, \ PMIN}(s)$	Utility function
$\tau_i(s) \in \mathbb{R}$	Decision rule for mapping overlapping into disjoint coalitions
$\epsilon \in [0, 1]$	Pruning threshold for the correlation graph
$N_{COAL} \in \mathbb{Z}^+$	Number of desired coalition
CS	Set of all coalition structures

Table I: Notations

we favor small coalitions, mainly because they are easier to maintain in terms of communications.

$$\mathcal{U}_{\phi, \ PMIN}(s) = \mathbf{1}_{s \ valid} \frac{P_\phi(s)}{|s|} \qquad (3)$$

Obviously, maximising this utility function amounts to maximizing the coalition contract value with the minimum possible number of agents.

In order to illustrate what is done in the following, lets consider a very simple example with two agents, say i and j, with gaussian value probability distributions $\mathcal{F}(P_i) = \mathcal{N}(\mu_i, \sigma_i)$ and $\mathcal{F}(P_j) = \mathcal{N}(\mu_j, \sigma_j)$, such that the joint probability distribution $\mathcal{F}(P_{ij})$ of the coalition $\{ij\}$ is also a gaussian with the following parameters :

$$\begin{cases} \mu_{ij} = \mu_i + \mu_j \\ \sigma_{ij} = \sqrt{\sigma_i^2 + \sigma_j^2 + \rho_{ij}\sigma_i\sigma_j} \end{cases} \qquad (4)$$

with ρ_{ij} the Pearson correlation coefficient between P_i and P_j. We can easily write the reliability condition $Pr[P_{ij} \leq P_\phi(ij)] \leq \phi$ as :

$$\frac{1}{2}\left[1 + erf\left(\frac{P_\phi(ij) - \mu_{ij}}{\sigma_{ij}\sqrt{2}}\right)\right] \leq \phi \qquad (5)$$

The strategy of $\{ij\}$ consists in maximizing its contract value as long as it respects the inequality 5, e.g to annonce $P_\phi(ij) = \mu_{ij} - \sigma_{ij}\sqrt{2}erf^{-1}(1-2\phi)$. It thus appears (as it was intuitively understandable) that, for equivalent sizes, coalitions with low relative standard deviations (σ_{ij}/μ_{ij}) are able to announce higher contract values.

What this paper investigates in the following is the developement of an algorithm that organizes prosumers such that the synergy term of the standard deviation ($\rho_{ij}\sigma_i\sigma_j$ in the example above) is minimized. In such settings, and through the clique percolation procedure (section IV), we will show that coalitions with low relative standard deviations, e.g high utility coalitions, can be computed.

IV. COALITION FORMATION

This section explains the process with which we form the coalitions (see the third block of Figure 1). First, we need to simulate the timeseries of available power (first two blocks of Figure 1). We consider a pool $\mathcal{A}$ of 200 agents, whose parameters were chosen randomly. The prosumers are positioned (also randomly) on a square lattice previously filled with climate vectors obtained from the french data sets (see section III). Simulations were run from february 2006 to december 2010 such that we are dealing with 200 hourly sampled timeseries of available power over this period. Removing season trends finally leads us to the formation of coalitions.

The model we used in order to simulate timeseries of available production provides some diversity because of the combination between the energetic mix (renewable generators combination) and the climate vectors. Nevertheless, as the number of agents grows, the timeseries tend to exhibit similar patterns. This is apparent when creating a correlation graph $G_1(\mathcal{A}, E_1, \omega_1)$ with the same kind of metric as [11] or [12] ($d_{ij} = 1 - \rho_{ij}^2$), where well defined clusters appear in the ϵ-graph for any values of ϵ.

However, these clusters of strongly correlated timeseries are the exact opposite of what we are seeking. We can indeed consider them directly as coalitions and compute their utilities, and the results show (see the green curves on figure 4), as expected, terrible values (far worse than a random split of the agents in the same number of coalitions).

We thus opt for reversing the metric ($d'_{ij} = \rho_{ij}^2$) such that uncorrelated timeseries are close to each other and (anti)correlated timeseries are distant in $G_2(\mathcal{A}, E_2, \omega_2)$. As expected [12], independently of the ϵ parameter, the resulting ϵ-graph exhibits henceforth much less clustering than in G_1 and communities seem hardly visible. Therefore, using classical clustering or community detection algorithms seem to provide poor results.

However, as seems intuitively understandable, cliques of this graph tend to exhibit very good utility values. Such structures contain indeed a link between every two nodes, meaning that the overall correlation is quite small. Obviously, the ϵ parameter is indirectly responsible for the sizes of the cliques, if it is too low, the ϵ-graph of G_2 will not provide enough cliques, conversely, if ϵ is too high, we loose important information as the graph becomes very dense and cliques tend to overlap strongly. For large values of ϵ and for non trivial number of agents, finding cliques can even become computationally repulsive. Despite being direct and simple, improving the sizes of the cliques by increasing ϵ seems too brutal. Furthermore, as the utility function is also focused on maximizing $P_\phi(s)$, it might be the case that some cliques benefit from additional agents even if they don't form a clique anymore.

Cliques for small values of ϵ appear thus as good seeds for stable coalitions. However we have to find a way to bring the coalitions above the grid requirements as well as increasing the social welfare by incorporating more nodes to
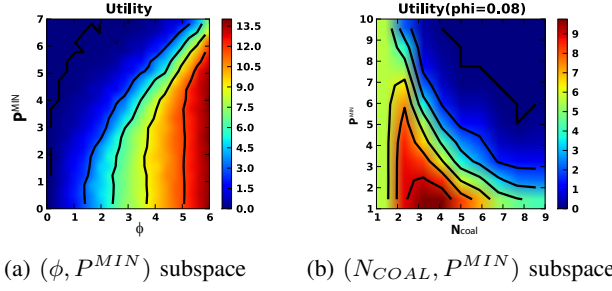
(a) (ϕ, P^{MIN}) subspace   (b) (N_{COAL}, P^{MIN}) subspace

Figure 2: Utility function as function of the parameters space $\{\phi \in [0,1], N_{COAL} \in \mathbb{Z}^+, P^{MIN} \in \mathbb{R}^+\}$: In $\{\phi \in [0,0.6], P^{MIN}\}$ when N_{COAL} is fixed (fig 2a), and in $\{N_{COAL}, P^{MIN}\}$ for a fixed ϕ (fig 2b). For readability, P^{MIN} is expressed in tenth of MW.

the coalitions. Clique percolation algorithms have proven to be a useful heuristic for such complex problems. The main idea behind this algorithm is to make the seeds grow as long as they increase a fitness function based on internal and external degrees of the seeds [16]. The main difference between our case study and that algorithm is that we chose this fitness function as our utility function (cf. eq.(3)) that both maximizes the production and minimizes its volatility.

As explained in section II, clique percolation leads generally to overlapping communities. For simplicity, in this paper, we wish to keep the coalitions disjoint and leave the management of overlapping coalitions for future work. We thus implemented a simple heuristic that considers nodes in multiple seeds one by one and chooses its final coalition as the one that "needs it the most" in terms of utility loss. More formally, for a coalition s and a node $i \in s$, we define :

$$\tau_i(s) = \frac{\mathcal{U}_{\phi,\ P^{MIN}}(s) - \mathcal{U}_{\phi,\ P^{MIN}}(s - \{i\})}{\mathcal{U}_{\phi,\ P^{MIN}}(s)} \qquad (6)$$

If node i belongs to multiple coalitions, the only coalition including node i is the one that maximizes τ_i.

At this point, we have three degrees of freedom : the reliability (ϕ), the required power to enter the market (P^{MIN}), and the prunning parameter (ϵ). For convenience, we introduce the number of desired coalitions (N_{COAL}) and we choose ϵ as the smallest value such that the ϵ-graph contains at least N_{COAL} non overlapping cliques. The overall utility depends now on the grid requirements and on the number of desired coalitions.

Next section shows the utility behavior within the parameters space and provides the results of our 200 agents test.

V. RESULTS

As shown in figure 2a, the parameters ϕ and P^{MIN} shape the utility function such that, if ϕ is close to zero, the reliability requirement is very high. Only small values of P^{MIN} could then potentially lead to valid coalitions (and positive utility values). Conversely, the higher ϕ, the less constraints are imposed to the coalitions and more valid coalitions can arise

for a larger spectrum of P^{MIN}. Obviously, the highest utility values are found for high ϕ, because the coalitions formed are able to announce higher contract values, yielding higher utilities.

In figure 2b, we fix the reliability value to a given empirical value ($\phi = 0.08$). We can observe how the social welfare $\sum_{s \in S} \mathcal{U}_{\phi,\ P^{MIN}}(s)$ evolves according to P^{MIN} and N_{COAL}. In our setting we can observe a maximum point around $N_{COAL} = 3$, followed by a decrease. There is clearly a maximum number of coalitions sustainable for a given set of parameters. After this value we will form more coalitions that would be unable to pass the grid requirements. Moreover, reckon that increasing N_{COAL} means also increasing ϵ, leading to denser graphs, meaning that the algorithm performance will also decrease.
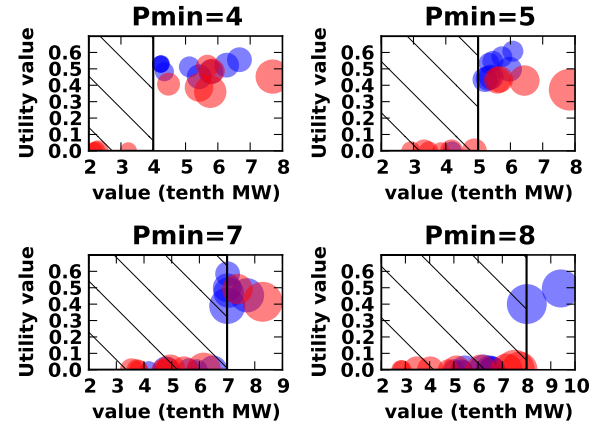

Figure 3: Evolution of the coalitions for different values of P^{MIN} for clique percolation (blue dots) and random process (red dots). The diameters of the dots are proportional to the corresponding coalition sizes. The hatched zone corresponds to the *"under-requirement space"*, and results in zero utility for any coalition in this area.

In figure 3, we perform a simulation with 200 prosumers, and we fix the reliability $\phi = 0.1$ meaning that coalitions should produce more than their contract values at least 90% of the time. We observe the result of our algorithm for a fixed number of coalitions of $N_{COAL} = 10$ and let the algorithm form the coalitions as the grid changes according to P^{MIN}. We compare our algorithm with a random process that splits the agents in N_{COAL} coalitions. Figure 3 shows this evolution for our algorithm (blue dots) and for the random process (red dots). The diameter of a dot is proportional to the number of agents in the coalition and the higher the dot is, the higher its utility. The P^{MIN} values of the x axis are expressed in tenth of MW for readability and the hatched zone corresponds to the *"under-requirement space"*, meaning that whenever a coalition is in this zone, it has a null utility.

Looking at figure 3, we can see first that our algorithm performs better than the random formation algorithm at finding high valued coalitions. The blue dots allowed to enter the
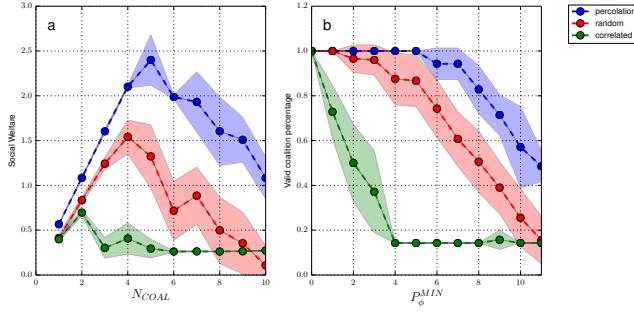
Figure 4: Evolution of the social welfare **a)** and the percentage of coalitions able to enter the market **b)** for different values of P_{MIN}. Blue curves represent clique percolation, red curves, the random process, and the green ones, correlated coalitions.

market (non hatched zone) are indeed outnumbering the red dots, especially when the grid requirements are neither too low nor too high ($P^{MIN} = 5$ in figure 3).

In more details, figure 4a presents the evolution of social welfare as the number of coalitions increases (all other parameters are kept constant) for random (red curve), clique percolation (blue curve), and correlated coalitions (green curves) that stands for a worst case scenario. As for figure 4b, it shows the percentage of coalitions able to enter the market for different values of P^{MIN}. For consistency, we average the results of both plots of the random procedure over 100 realizations.

When the grid requirements are constant (figure 4b), and the number of desired coalitions is low, clique percolation generally performs better than random case. Moreover, when N_{COAL} gets bigger, the performance of a random split tumble down rapidly while our clique percolation is able to maintain efficiently the social welfare of the coalitions formed.

When the grid requirements vary, for very low P^{MIN}, all coalitions for all algorithms are able to enter the market, yielding an acceptance percentage of 100%. But as P^{MIN} increases, we see the percentage of the correlated coalitions is going down very quickly. After a few increases in P^{MIN}, the percentage of the random procedure starts dropping while it stays almost constant for our algorithm. For $P^{MIN} = 8$, we see that only a little more than half of the coalitions for the random case are able to enter while approximately 85% of them enters for the clique percolation algorithm. Finally, when the grid requirements becomes too high, the acceptance percentage of our algorithm tends slowly to zero (not shown in this plot for readability).

VI. Conclusion

We believe that the originality of this paper lies in its willingness to exploit the de-correlation of prosumer profiles in order to build stable coalitions. In this direction, we presented a model based on meteorological traces, that captures the complex "*energetic mix/climate vectors*" combination and generates realistic production and consumption patterns. We then built a framework that enables the grid to specify stability and minimum production requirements for filtering the coalitions.

On this basis, we proposed a simple algorithm that seeks for uncorrelated prosumer patterns as potential seeds and expand them in coalitions able to rise above the grid requirements.

The algorithm is validated against a random process, where we show that it performs better (coalitions are more stable and the overall production is more important) and that it exhibits a higher robustness/flexibility against grid requirements changes.

Interesting leads for future works would be the use of correlated clusters to reduce the number of entities the algorithm has to deal with, or the introduction of a payoff allocation towards the prosumers such that the stability against player defection could be analyzed through game theory. Showing that maintaining the coalitions formed with our algorithm necessitates less communication and less storage capacity could also conduct to a stimulating project. Besides, we believe that not restricting the algorithm to non overlapping coalitions and studying the strategies and weights of nodes with multiple options could lead to interesting works.

References

[1] S. D. Ramchurn, P. Vytelingum, A. Rogers, and N. R. Jennings, "Putting the 'smarts' into the smart grid: A grand challenge for artificial intelligence," *Commun. ACM*, vol. 55, no. 4, pp. 86–97, Apr. 2012.

[2] C. Wu, A.-H. Mohsenian-Rad, and J. Huang, *Smart Grid Communications and Networking*, E. Hossain, Z. Han, and V. Poor, Eds. Cambridge University Press, 2012.

[3] A. J. D. Rathnayaka, V. Potdar, and M. H. Ou, "Prosumer management in socio-technical smart grid," in *Proceedings of the CUBE International Information Technology Conference*, 2012, pp. 483–489.

[4] S. Pahwa, C. Scoglio, and N. Schulz, "Topological Analysis and Mitigation Strategies for Cascading Failures in Power Grid Networks."

[5] M. Braun, "Virtual Power Plants in Real Applications Pilot Demonstrations in Spain and England as part of the European project FENIX."

[6] R. Mantegna, "Hierarchical structure in financial markets," *The European Physical Journal B*, vol. 11, no. 1, pp. 193–197, 1999.

[7] W. Saad, Z. Han, M. Debbah, and A. Hjorungnes, "A distributed coalition formation framework for fair user cooperation in wireless networks," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 9, pp. 4580–4593, September 2009.

[8] X. Luan, J. Wu, S. Ren, and H. Xiang, "Cooperative power consumption in the smart grid based on coalition formation game," in *Advanced Communication Technology (ICACT)*, Feb 2014, pp. 640–644.

[9] F. Van Hulle *et al.*, "Integrating Wind, European project report."

[10] G. Chalkiadakis *et al.*, "Cooperatives of distributed energy resources for efficient virtual power plants," in *The 10th International Conference on Autonomous Agents and Multiagent Systems*, 2011, pp. 787–794.

[11] A. Garas, P. Argyrakis, and S. Havlin, "The structural role of weak and strong links in a financial market network," *The European Physical Journal B*, vol. 63, no. 2, pp. 265–271, 2008.

[12] J.-P. Onnela *et al.*, "Clustering and information in correlation based financial networks," *The European Physical Journal B - Condensed Matter and Complex Systems*, vol. 38, no. 2, pp. 353–362, 2004.

[13] M. E. J. Newman, "Modularity and community structure in networks," *Proceedings of the National Academy of Sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.

[14] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[15] M. E. J. Newman, "Finding community structure in networks using the eigenvectors of matrices," *Phys. Rev. E*, vol. 74, p. 036104, Sep 2006.

[16] A. Lancichinetti, S. Fortunato, and J. Kertész, "Detecting the overlapping and hierarchical community structure in complex networks," *New Journal of Physics*, vol. 11, no. 3, p. 033015, 2009.

[17] [Online]. Available: http://www.infoclimat.fr

[18] [Online]. Available: http://www.ncdc.noaa.gov/

[19] [Online]. Available: http://www.wind-power-program.com

[20] G. Dans and F. Science, "Rayonnement solaire," vol. 360, 2007.