# Allowing Stable Groups of Prosumers to Enter the Market through Proper Aggregation


Nicolas Gensollen
Institut Mines Telecom
Email: http://www.michaelshell.org/contact.html

Vincent Gauthier
Twentieth Century Fox
Springfield, USA
Email: homer@thesimpsons.com

Michel Marot
and Monique Becker
Starfleet Academy
San Francisco, California 96678-2391
Telephone: (800) 555–1212
Fax: (888) 555–1212



***Abstract*—There is a growing interest for coalitional concepts in the smart grid, spanning multiple objectives such as power losses reduction, benefits maximization or market stability... In this paper, we study coalitions of prosumers (agents that produce but also consume depending on time) which aim at selling energy to the grid. We reckon that the grid is able to specify two requirements (stability and minimum production) that enable a coalition to enter the market. Based on probability distributions of units available power, coalitions which are able to achieve low standard deviations are considered as stable coalitions. As it seems intuitively understandable that maintaining stable coalitions would be less expensive in terms of communication load than highly versatible ones, our objective consists in forming coalitions that satisfy grid requirements both in terms of stability and minimum production. Our definition of stability being implicitely linked to correlation between agents timeseries (see section III), we organize them over a correlation graph and use a variation of clique percolation to construct valid coalitions in reasonable time.**


## I. Introduction

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

It is rather clear that there should be communication flows between prosumers and the grid as most of the demand side management concepts suppose such a link [2]. Nevertheless as the expected number of active prosumers in the future seems quite large, the literature assume more and more communication between prosumers as this is a key ingredient to maintain more complex structures (such as coalitions), and thereby decrease the communication load on the grid's side. Furthermore, it has been shown that coalitions in the smart grid provide nice ways of improving stability while posing exciting challenges in terms of network infrastructure and management. Self sufficient microgrids that can disconnect from the main grid [4], virtual power plants (VPP) that map small DER in sizeable and adjustable plants [5] [1], or coalitions of electric vehicles that back up the grid in emergency situations (V2G) [1] are just a few examples of how coalitional concepts can enhance grid reliability.

In this paper, our focus is on how coalitions of prosumers, aiming at selling energy to the grid, should be formed in order to support it efficiently. Namely, improving statistically the production stability while decreasing the communication needs for staying in a stable state. The idea is that, by using prediction techniques, coalitions can propose contract values to the grid, enabling it to schedule production on its side accordingly. Actually, contract values are particularly relevant in power exchanges contexts where energy is traded based on predictions (day-ahead estimations for instance). It seems obvious that, in such systems, participants should try to minimize prediction errors in order to maintain a stable state of grid operation, and that some penalty rules should be applied to ensure that coalitions are willing to report contract values correctly. Furthermore, as the coalition productions are supposed to come from (uncertain) renewables, the notion of reliability comes into play with a visible tradeoff between contract values worth and reliability.

In this paper, our concern is on how prosumers should be aggregated in order to optimize this tradeoff, that is, how coalitions should be formed with the intention of announcing "high contract values with high reliability". We thus aim at:

- Realistic representations (prosumers productions are non independent variables and come (indirectly) from real traces)
- Minimizing the standard deviations of coalitions production probability distributions through the use of a proper utility function (see section III)
- Being (relatively) scalable with the number of prosumers

In more details, and as will be explained in section III, we consider agents that, depending on meteorological conditions, personal preferences, and their appliances (loads and generators), consume and produce more or less energy (using only wind turbines and PV as generators, which can be quite simply and efficiently modeled). We used real meteorological traces (see section III) to account for seasonal and daily variations in the prosumers energetic output profiles as well as realistic correlations. With these ingredients, we are able to run simulations and record the different output profiles as timeseries.

As correlations play a significant role in our mechanism, we organize the prosumers in a correlation graph, a quite popular approach in market stocks clustering since [6] (see section II). We will see that in such graphs, cliques represent structures of interest. In our settings, it means that cliques are “good” (from the utility point of view) foundations for high valued coalitions. We thereby use a modified clique percolation algorithm (see section V) that will enable us to expand the cliques as much as needed in order to form proper coalitions that fulfill grid requirements.

The rest of the paper is organized as follows, section II locates our work in the related litterature. Section III explains the model we used to obtain realistic prosumers available power timeseries and introduces some notations. Section IV relates the method for building coalitions based on the agent timeseries. Finally, section V provides some results.

## 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 focus on finding an optimal coalition structure giving a pool of autonomous self interested agents. In this direction, [7] and [8] design a distributed merge and split algorithm based on modified Pareto order with utility functions aiming at optimizing a parameter of interest such as energy losses or communication costs.

Attacking 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, namely 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 and as will be explained in section III, coalitions utilities depend on statistical properties of historical values. The goal being to form coalitions that, statistically speaking, are more likely to exhibit stable behaviors. We will see that, when it comes to the formation of the coalitions, a prosumer is completely interchangeable with the timeserie representing its available production.

The setting is thus similar with some fiancial 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] Mantegna introduced another approach where stocks, represented by the timeseries of their daily log returns, are organized in a graph such that stocks (vertices) exhibiting similar price fluctuation patterns are close to each other. This closeness notion is formalized with a similarity measure based on Pearson correlation coefficient ($d_{ij} = \frac{1}{2}\sqrt{2(1-\rho_{ij})}$ but sometimes also $d_{ij} = 1-\rho_{ij}^2$ ) that enables to weight the edges of the graph. Because there is a distance between every two nodes, [6] obtains a complete graph where precious information is flooded. In order to discard unsignificant information, [6] used a minimum spanning tree oriented technique and achieves a hierarchical clustering of the stocks. Since then, several other filtering techniques have been proposed, such as k-nearest graphs, or $\epsilon$-graphs, that relax the tree assumption [11].

The idea of $\epsilon$-graphs consists in filtering edges based on their weight, only keeping edges whose weights are less than $\epsilon$. Despite the procedure simplicity, the choice of $\epsilon$ (or k for k-nearest graphs) is not trivial and may influence strongly the results. There is indeed to our knowledge no optimal rule for chosing $\epsilon$. In [12] and [11], 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 $\epsilon$. One of the conclusions was that "strong links" (i.e links between strongly correlated timeseries) are responsible for clustering while "weak links" provide network connectivity.

Presented in this way, the timeseries 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 betweeness [14], spectrum of Laplacian [15]...), the clique percolation algorithm [16] uses directly this observation and the idea behind it is actually quite simple. It starts indeed by searching for cliques (a complete subgraph) in the given graph as potential community cores. Because other nodes can naturally belong to a community with a less strong connectivity requirement, the algorithm uses a fitness function that quantifies the quality (according to some criterions) of a community structure inside the graph. Expansed cliques alternatively look in their respective neighborhood, select the node that yields the best increase in fitness, and finally incorporate it. The percolation algorithm stops when seeds cease to grow or when all nodes are affected to at least one seed. At this point, 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 persons often belongs to several communities (family, friends, work colleagues...).

The following section of this paper explains the model we built in order to obtain sufficient historical values for our prosumers, such that forming coalitions based on statistics is actually meaningfull.

## III. Model

A major concern while designing our model was its ability to translate correctly realistic patterns (consumption and production) as well as realistic geographical correlations between patterns. Largely because of these reasons (but also because they are easily, and sometimes freely, available online) we chose to use meteorological traces as inputs. In this paper we used namely both french data from 2006 to 2012 sampled every 3 hours [17] (similar data for the United States covering year 2010 can also be found at [18]). The weather stations already offer a quite good sampling of a given territory, and, with the growth of personal weather stations constantly updating data bases, the mesh becomes finer and finer. 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. In this paper, we consider 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 (we did not consider air conditioning in this paper). Temperature traces were the principal ingredient for modeling these cycles.

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

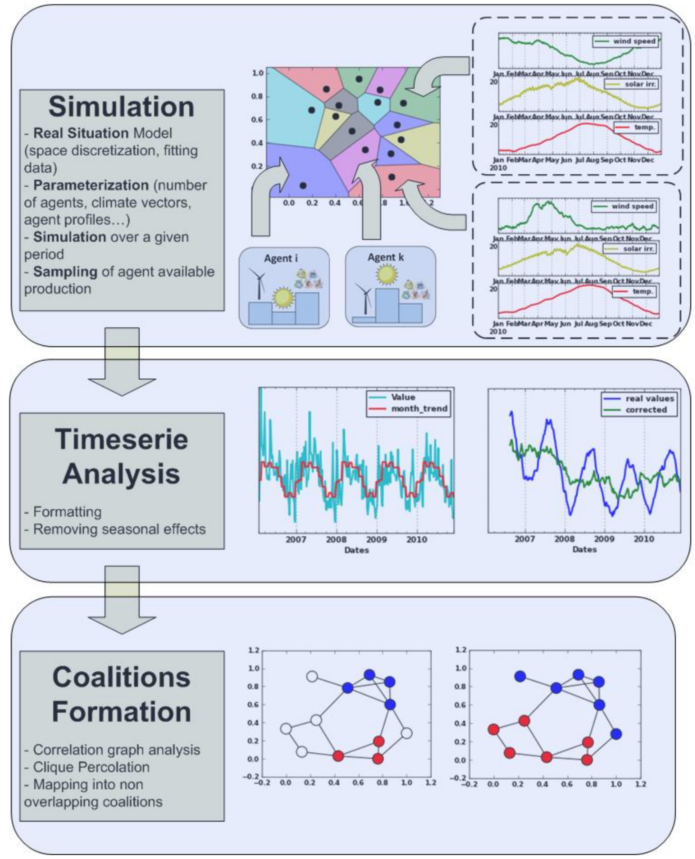


Figure 1: Process diagram

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 her zone, the DER and appliances she owns, and the way she decides to heat her home, a given prosumer is able to compute her production/consumption at any time. We thus used simple and known power curve models for linking weather variables to output power, such that a wind turbine (or a PV for instance) takes a wind speed (a solar irradiance) as input and gives an instantaneous output power [10] [19]. Any agent is thus able to compute at any time (within the traces bounds) how much she produces and consumes.

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 of agent i at instant t (its instantaneous production minus 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 (because we introduced them), the timeseries exhibit high seasonal patterns, but completely different from agent to agent because they depend both on the energetic mix and habits of a given prosumer. Nevertheless, these macro variations can hide the interesting information in the correlation coefficients and twist the rest 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 timeserie for agent i. For any agent i, we note $\mathcal{P}_i$ the probability distribution drawn from $\mathcal{T}_i$ and by $P_i$ a random variable following $\mathcal{P}_i$.

We now extand 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 $\mathcal{P}_S$ its probability distribution
- A coalition S has the possibility to announce a contract value $P_S^{CRCT}$ on the market.

In this paper, we consider that coalition announcements are constrained by the grid. Namely, the grid has the possibility of fixing two parameters :

- the **reliability** (denoted by $\phi \in [0, 1]$)
- the **minimum contract value** ( $P_\phi^{MIN}$).

The reliability 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 $\phi$. That is, $\forall S,\ \forall t,\ Pr[P_S(t) \leq P_S^{CRCT}] \leq \phi$. Furthermore, for consistency, we restrict $\phi$ to small values ( $\phi << 0.5$). The minimum contract value under $\phi$ states that only coalitions with higher contract values ($P_S^{CRCT} \geq P_\phi^{MIN}$) will be accepted.

At this point, if a coalition S whishes 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 $P_\phi(S) \geq P_\phi^{MIN}$, meaning that it also obeys the minimum contract value rule, then S annonces 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 :

$$\begin{cases} \forall t,\ Pr[P_S(t) \leq P_\phi(S)] \leq \phi \ \textit{(reliability rule)} \\ and\ P_\phi(S) \geq P_\phi^{MIN} \ \textit{(min value rule)} \end{cases}$$

We now choose a very simple utility function 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 indicates that we favorise small coalitions, mainly because they are easier to maintain in terms of communications.

$$\mathcal{U}_{\phi,\ P_\phi^{MIN}}(S) = \mathbf{1}_{S\ valid} \frac{P_\phi(S)}{|S|}$$

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 $\mathcal{P}_i = \mathcal{N}(\mu_i, \sigma_i)$ and $\mathcal{P}_j = \mathcal{N}(\mu_j, \sigma_j)$, such that the joint probability distribution $\mathcal{P}_{ij}$ of the coalition $\{ij\}$ is also a gaussian with the following parameters :

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

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

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

The strategy of $\{ij\}$ consists in maximizing its contract value as long as it respects this inequality, 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 ( $\sigma_{ij}/\mu_{ij}$ ) 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 deviation, e.g high utility coalitions, can be computed.

## IV. Forming coalitions

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 provide some diversity because of the combo between the energetic mix and the climate vectors. Nevertheless, as the number of agents grows, the timeseries tend to exhibit similar patterns. This is particularly visible when creating a correlation graph $G_1(\mathcal{A}, E_1, \omega_1)$ with the same kind of metric as [12] or [11] ($d_{ij} = 1 - \rho_{ij}^2$). Indeed, very well defined clusters appear in the corresponding $\epsilon$-graph for any values of $\epsilon$.

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, 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 [11], independently of the $\epsilon$ parameter, the resulting $\epsilon$-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 ideed a link between every two nodes, meaning that the overall correlation is quite small. Obviously, the $\epsilon$ parameter is indirectly responsible for the sizes of the cliques, if it is too low, the $\epsilon$-graph of $G_2$ will not provide enough cliques, conversely, if $\epsilon$ is too high, we loose important information as the graph becomes very dense and cliques tend to overlap strongly. For large values of $\epsilon$ and for non trivial number of agents, finding cliques can even becomes computationally repulsive. Despite being direct and simple, improving the sizes of the cliques by increasing $\epsilon$ 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 $\epsilon$ appear thus as good seeds for stable coalitions, but we still have to find a way to bring them above the grid requirements as well as increasing the social welfare, both by incorporating more nodes. This is exactly where clique percolation proves useful for our concerns. The main idea of making the seeds grow as long as it increases a fitness function is exactly what we need. The only difference being that our fitness function is simply chosen as the utility function.

As explained in section II, clique percolation leads generally to overlapping communities. For simplicity, in this paper, we wish to keep the coalitions separated and leave the management of overlapping coalitions for future work. We thus implemented a simple heuristic that consider 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 :

$$\tau_i(S) = \frac{\mathcal{U}_{\phi,\ P_\phi^{MIN}}(S) - \mathcal{U}_{\phi,\ P_\phi^{MIN}}(S - \{i\})}{\mathcal{U}_{\phi,\ P_\phi^{MIN}}(S)}$$

If node i belongs to multiple coalitions, the only coalition retaining node i is the one that maximizes $\tau_i$.

At this point, we have three degrees of freedom, namely the reliability ($\phi$), the required power to enter the market ($P_\phi^{MIN}$), and the prunning parameter ($\epsilon$). For conveniance, we introduce the number of desired coalitions ($N_{COAL}$) and we fix $\epsilon$ as the smallest value such that the resulting $\epsilon$-graph contains at least $N_{COAL}$ non overlaping cliques. The overall utility depends now on the grid requirements and on the number of desired coalitions.

The next section shows how the utility behave within the parameters space and provides the results of our 200 agents test.

## V. Results

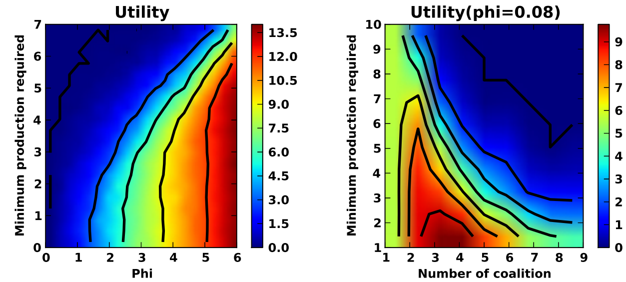


Figure 2: Utility function in parameters space. Left in $(\phi, P_\phi^{MIN})$ when $N_{COAL}$ is fixed. Right in $(N_{COAL}, P_\phi^{MIN})$ for a fixed $\phi$

As visible on the left plot of figure 2, the parameters $\phi$ and $P_\phi^{MIN}$ shape the utility function such that, if $\phi$ is close to zero, the reliability requirement is very high and only small values of $P_\phi^{MIN}$ could potentially lead to valid coalitions (and positive utility values). Conversely, the higher $\phi$, the less constraints are imposed to the coalitions and valid ones can arise for a larger spectrum of $P_\phi^{MIN}$. Obviously, the highest utility values are found for high $\phi$, because coalitions are able then to announce higher contract values, yielding higer utilities.

In the following, we fix the reliability to a given empirical value ($\phi = Cste$) and observe how the coalitions evolves for different values of $P_\phi^{MIN}$. Note that the opposite is also possible although a little less intuitive, but not shown here because of space limitations and because it leads to similar conclusions.

On the right plot of figure 2, we can see that when the grid constraints are fixed ($\phi = 0.08$ and $P_\phi^{MIN} = 2$ for instance) and only the number of coalitions varies, there is an increase in social welfare (sum of coalitions utilities) up to a maximum point (for $N_{COAL} = 3$ here) before a decrease. The reason is that increasing the number of coalitions allows more coalitions to reach stability and enter the market, but there is a point where nodes bringing stability are not sufficient inside the coalitions to make them pass the grid requirements, and some coalitions start to fail with zero utility. Moreover, reckon that increasing $N_{COAL}$ means also increasing $\epsilon$, leading to denser graphs where information is flooded, meaning that the algorithm performances also decrease. Naturally, it is also visible that small values of $P_\phi^{MIN}$ lead to higher utilities because coalitions are able to announce higher contract values.

For our 200 prosumers, we suppose that the grid fixes a reliability $\phi = 0.1$ (meaning that coalitions should produce more than their contract values at least 90% of the time). We also fix the number of coalitions to 10 and have a look at how these coalitions evolve as the grid changes the $P_\phi^{MIN}$ requirement. As a comparison, we use a completely random algorithm that only asks for a number of coalitions and partitions the agents in a random fashion. 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 $P_\phi^{MIN}$ 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 seems to make generally a better job at finding high valued coalitions, but also that the results seem far more robust against how the grid positions its requirements. The blue dots allowed to enter the market seem indeed to outnumber the red dots, especially when the grid requirements are neither to low or to high ($P_\phi^{MIN} = 5$ in figure 3).

In more details, the left plot of figure 4 presents the evolution of social welfare as the number of coalitions increases (all other parameters are kept constant) for both random (red curve) and clique percolation (blue curve). As for the right plot, it shows the percentage of coalitions able to enter the market for different values of $P_\phi^{MIN}$. For consistency, we average the results of both plots of the random procedure over 100 realizations (5 realizations for clique percolation, but the results are relatively stable), and the errorbars in the plots stands for the standard deviations of the results.

When the grid requirements are constant (left plot), and the number of desired coalitions is low, clique percolation generally performs only a little better than random. Nevetheless, when $N_{COAL}$ gets bigger, the performance of a random split tumble down rapidely while clique percolation social welfare decreases slowly.

When the grid requirements vary, for very low $P_\phi^{MIN}$ (close to zero), all coalitions for both algorithms are able to enter the market, yielding an acceptance percentage of $100\%$. But as $P_\phi^{MIN}$ increases, we see the percentage of the random procedure quickly dropping while it stays constant for our

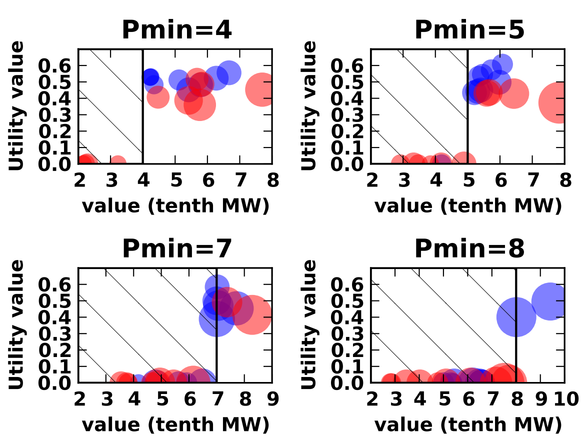


Figure 3: Evolution of the coalitions for different values of grid requirement $P_\phi^{MIN}$. Blue dots represents clique percolation results while red dots stands for random process results.

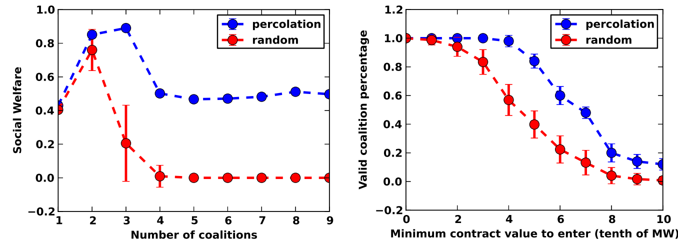


Figure 4: Percentage of coalitions able to enter the market for different values of grid requirement $P_{MIN}$. Blue curve represents clique percolation results while red curve stands for random process results

algorithm. For $P_\phi^{MIN} = 4$, we see that only a little more than half of the coalitions for the random case are able to enter while all of them enters for the clique percolation algorithm. Finally, when the grid requirement becomes to high, the acceptance percentage of our algorithm tends slowly to zero (only partly shown in this plot for readability).

## VI. Conclusion

In this paper we collected meteorological traces for weather stations sampling a given territory. After discetization of the space, each zone receives the climate vector corresponding to its mother weather station. Prosumers are then distributed randomly between zones and each prosumer receives a random number of DER and loads as well as diverse consumption habits. Simulations are then runned and allow us to obtain, for each prosumer, the timeserie describing the instantaneous power that he could potentially be wanted to sell on the market. Nevertheless, as entering on the energy market is constrained by the grid operator, such that only sufficiently reliable and sufficiently productive units are allowed to enter, prosumers have usually no other choice than forming coalitions. We then introduced a framework such that the grid requirements take a precise meaning and such that coalitions announce contract values accordingly.

Through this paper, our interest was on how we could circumscribe groups of prosumers that can rise above the grid requirements. In this direction, we removed seasonal variations from the prosumers timeseries and proposed to organize them over a correlation graph with a reversed metric. A greedy clique percolation process on the $\epsilon$-graph enables us to finally establish coalitions of agents. We showed in section V that our algorithm was actually finding coalitions with high utility values but also that it was far more robust than a random search over coalition structures when the grid operator increases his requirements (Figure 4).

An interresting lead for future work would be the introduction of a payoff allocation towards the prosumers such that the stability against player defection could be analysed. 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.

## ACKNOWLEDGMENT

The authors would like to thank...

## 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."

[2] C. Wu, A.-H. Mohsenian-Rad, and J. Huang, *Vehicle-to-grid systems: ancillary services and communications*. Cambridge University Press, 2012.

[3] a. J. D. Rathnayaka, V. Potdar, and M. H. Ou, "Prosumer management in socio-technical smart grid," *Proceedings of the CUBE International Information Technology Conference on - CUBE '12*, p. 483, 2012. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2381716.2381808

[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," vol. 49, no. 561.

[6] R. N. Mantegna, "P HYSICAL J OURNAL B," vol. 197, pp. 193–197, 1999.

[7] W. Saad and Z. Han, "A Distributed Coalition Formation Framework for Fair User Cooperation in Wireless Networks," no. May 2008, 2009.

[8] X. Luan, J. Wu, S. Ren, and H. Xiang, "Cooperative power consumption in the smart grid based on coalition formation game," *16th International Conference on Advanced Communication Technology*, pp. 640–644, Feb. 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6779040

[9] D. Europe, "Integrating Wind."

[10] R. Kota, A. Rogers, and N. R. Jennings, "Cooperatives of Distributed Energy Resources for Efficient Virtual Power Plants," no. Aamas, pp. 787–794, 2011.

[11] J.-P. Onnela, K. Kaski, and J. Kertesz, "Clustering and information in correlation based financial networks," *The European Physical Journal B - Condensed Matter*, vol. 38, no. 2, pp. 353–362, Mar. 2004.

[12] 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, Jun. 2008. [Online]. Available: http://www.springerlink.com/index/10.1140/epjb/e2008-00237-3

[13] M. E. J. Newman, "Modularity and community structure in networks."

[14] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 12, pp. 7821–6, Jun. 2002.

[15] M. E. J. Newman, "Finding community structure in networks using the eigenvectors of matrices."

[16] A. Lancichinetti, "Detecting the overlapping and hierarchical community structure in complex networks."

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

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

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