Computer Networks 103 (2016) 15–32

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Dealing with uncertainty in the smart grid: A learning game approach☆

Hélène Le Cadre[a,*], Jean-Sébastien Bedo[b]

[a] ENSTA ParisTech, Applied Mathematics Unit (UMA), 828, boulevard des Maréchaux, 91120 Palaiseau, France
[b] Orange Labs, 78, rue Oliver de Serre, 75015 Paris, France

ARTICLE INFO

Article history:
Received 23 January 2015
Revised 9 March 2016
Accepted 14 March 2016
Available online 7 April 2016

Keywords:
Learning
Regret
Optimization under uncertainty
Game theory

ABSTRACT

We model the smart grid as a decentralized and hierarchical network, made up of three categories of agents: suppliers, generators and captive consumers organized in microgrids. To optimize their decisions concerning prices and traded power, agents need to forecast the demand of the microgrids and the fluctuating renewable productions. The biases resulting from the decentralized learning could create imbalances between demand and supply leading to penalties for suppliers and for generators. We analytically determine prices that provide generators with a guarantee to avoid such penalties, transferring risk to the suppliers. Additionally, we prove that collaborative learning, through coalitions of suppliers among which information is shared, minimizes the sum of their average risk. Simulations run for a large sample of parameter combinations, using external and internal regret minimization, show that the convergence of collaborative learning strategies is clearly faster than that resulting from individual learning. Finally, we analyze the suppliers' incentives to organize in a grand coalition versus multiple coalitions, and the tightness of the learning algorithm's theoretical bounds.

1. Introduction

In Europe, and especially France, power networks rely heavily on nuclear-based technology. With this type of nonrenewable technology, generations can be adapted by the plant operator who alternates openings and closings and optimizes the duration of the switches between modes. The objective is thus to adapt the generations so as to meet the uncertain demand level [13]. For renewables, generations can only be partially controlled, for instance, by lowering the wind turbine speed [17]. Renewable integration in the power network requires deploying smart Information and Communication Technologies (ICTs) to supervise the grid operations [18]. Indeed, renewable generation is highly unpredictable since it depends on uncontrollable exogenous factors like wind, sun, swell, etc. Furthermore, the new active role of end users, who can become power generators, dynamically adapt their consumption and fit into a multitude of microgrids [24,25], dramatically increases the volume of exchanged data flows. ICTs appear to be a means to retrieve the most salient information from this large amount of data and to train forecasters to provide efficient predictions regarding fluctuating generations such as renewables. These predictions will then be used as inputs to optimize the smart grid operations [2].

In practice, it is increasingly apparent that current forecasting methods cannot properly handle extreme situations corresponding to either severe weather phenomena or critical periods for power system operations. For example, forecasting methods used to predict wind power were mostly designed to provide single value forecasts to estimate the generations. Only recently, probabilistic methods have been introduced to provide estimations of the entire distribution of future generations [3] or predictions based on intervals [12]. In such methods, forecasts may take the form of either quantile estimations or density estimations [5]. The difficulty to process massive, heterogeneous and dynamic volumes of data emanating from decentralized and heterogeneous sources has favored the launching of automatic data processing methods grouped under the umbrella of *Machine Learning* [12]. Learning based on regret minimization [4] belongs to this latter category. This class of method provides the forecaster with a density function that

☆ The authors would like to thank the two anonymous referees and the Editor for their helpful comments which have helped us to greatly improve their model and the interpretation of the results.

* Corresponding author. Tel.: +33686976408.
E-mail addresses: helene.le-cadre@ensta-paristech.fr, helene.lecadre@gmail.com (H. Le Cadre), jeansebastien.bedo@orange.com (J.-S. Bedo).

http://dx.doi.org/10.1016/j.comnet.2016.03.002
1389-1286/

associates a weight to each possible output. The density function is updated by merging (online) information from various experts' reports. As a result, these methods are more robust in the face of extreme events and appear particularly well suited to model erratic processes such as renewables.

In the framework of the smart grid, learning is performed in a decentralized manner since each agent primarily learns the hidden information using its own observations. Existing literature on distributed learning mostly focuses on distributed learning algorithms that are suitable for implementation in large-scale engineering systems [14,15,23]. Many articles concentrate on games of potential [26]. This class of games is of particular interest since they have inherent properties that can provide guarantees on the convergence and stability of the system but remain very difficult to use in a full systemic framework. Indeed some systems may not be modeled as potential games [17]. A signaling game is also introduced in [14] to model competition among geographic demand markets constrained by limited transmission capacities.

The learning game studied in this paper belongs to the category of repeated uncoupled games since one agent cannot predict the forecasts and so actions of the other agents at a given time period. To take its decision i.e., optimal prices and power orders, each agent is aware of the forecast history of all the agents in its coalition and of its utility. For finite games with generic payoffs, recent work has shown the existence of completely uncoupled learning rules i.e., rules where the agents observe only their own prediction history and their utility, leading to Pareto-optimal Nash equilibria [23]. Marden et al. exhibit a different class of learning procedures that lead to a Pareto-optimal vector of actions that do not necessarily coincide with Nash equilibria [17]. Close to the work exposed in our article, Zheng et al. propose an online algorithm that simultaneously updates the weight given to each forecaster using regularized sequential linear regression, while allowing each forecaster to be retrained based on the latest observations in an online manner [27]. The updating of the individual forecasters to accommodate the online observations relies on a gradient-descent algorithm. In the same spirit, another approach based on reinforcement learning has been proposed by Khan et al. in [11]. Their heterogeneous learning algorithm allows the users to learn their own optimal payoff at a certain cost while simultaneously optimizing their strategy. Expert system coordination can also be used to aggregate the set of predictors into a better global predictor [10].

Most collaborative mechanisms studied in literature lead to price or quality of service alignment. In addition, the group composition provides an additional state space in which information about the environment can be accumulated [19]. To our knowledge, no study has so far been made of the impact of collaboration through information sharing, when prices are individually determined, on the underlying system performance. Of course, collaboration might not emerge due to the agents' natural incentives to cheat and deviate from the cooperative equilibrium and also, most frequently, due to the regulator's intervention. There are a number of well-understood reasons why regulators often do not allow horizontal collaboration: if suppliers are allowed to collaborate, they might cooperate to raise the price i.e., reduce quantity below the efficient baseline, and create market power [6]. Alternatively, suppliers might cooperate to reduce quality of service. Courts punish agreements that explicitly aim to decrease competition.

In this article, we answer the following questions:

(1) How will the biases caused by the errors made by the agents in their predictions affect the agents' average risk?
(2) Does collaborative learning improve the smart grid's overall performance?
(3) Do the agents have greater incentives to organize in a grand coalition than in a multiplicity of smaller size coalitions?

The article is organized as follows. In Section 2, we introduce the foundation of our model i.e., the agents, their utility and their optimization program. Complete information Stackelberg game is then solved in Section 3, proceeding by backward induction. We analytically derive the optimal prices and power orders for the agents. Partial information is introduced in Section 4 where the interacting agents learn hidden individual sequences in a distributed fashion. To illustrate the theoretical results derived in the previous sections, in Section 5 we compare: firstly, the time of convergence of suppliers' learning strategies under external and internal regret minimization in cooperative and non-cooperative scenarios, secondly, which behaviors should emerge depending on the game parameters value and finally, analyze the tightness of the bounds derived theoretically in various scenarios.

Notation and modeling assumptions

Agents and utility functions

$\mathcal{M}_i$	Microgrid i
S_i	Supplier i
G_k	Generator k
TSO	Transmission System Operator
$\Pi_i(t)$	S_i utility
$\tilde{\Pi}_k(t)$	G_k utility
$B_i(t)$	Net benefit for $\mathcal{M}_i$
$\Pi_i^0(t)$	S_i utility evaluated in unbiased forecasts

Parameters

t	Generic time period
n	Number of suppliers
K	Number of generators
θ_i	Parameter depending on the composition of $\mathcal{M}_i$
b_0, b_1	Generic positive parameters
γ_i	S_i unitary penalty cost
$\tilde{\gamma}_i$	G_k unitary penalty cost in the provision of S_i

Decision variables

$\tilde{p}_k(t)$	G_k unitary price
$q_{ik}(t)$	Power order from S_i to G_k
$p_i(t)$	S_i unitary price
$a_i(t)$	$\mathcal{M}_i$ decision variable
$\alpha_{ki}(t)$	Proportion of power allocated by G_k to S_i

Random variables and forecasters

$\nu_i^S(t)$	$\mathcal{M}_i$ generation
$\nu_k^G(t)$	G_k production
$\mathcal{E}_G$	Set of all the possible production levels for one generator
$\mathcal{E}_S$	Set of all the possible generation levels for one microgrid
$\mathbf{f_i(t)}$	Vector of the predictions made by S_i
$f_i(\nu_i^S, t)$	S_i forecast of $\mathcal{M}_i$ generation
$f_i(\nu_k^G, t)$	S_i forecast of G_k production
$\mathbf{f(t)}$	Vector of forecasts of all the suppliers
$\mathbf{f_{-i}(y, t)}$	Vector of forecasts of all the suppliers except S_i whose prediction is replaced by $\mathbf{y}$
$\nu(\mathbf{t})$	Vector of productions of microgrids and generators
$d_{ij}^k(t)$	Disagreement between S_i and S_j on the prediction of G_k production
$\underline{D}_{SS}(i)$	Lower bound on the disagreements between S_i and the other suppliers
$\overline{D}_{SS}(i)$	Upper bound on the disagreements between S_i and the other suppliers
$d_t(.)$	Supplier's learning strategy