A Game Theoretical Approach to Smart Grid Energy Cooperation

Max Henning Höth
2055977

Alejandro León Castell
2046806

Walter Martemucci
2057975

January 19, 2023

Abstract

We present two game-theoretic models for energy transfer between service gateways of a smart grid, and find the conditions on the payoffs for reaching an energy cooperation scenario. In particular, we consider a model with an exogenous centralized agent that both subsidizes energy prices and restricts the action choice; and a decentralized model where we penalize the smart grid going down. We then numerically simulate these models and compare the utilities, battery energy levels and energy prices as the game stages progress.

1 Introduction

We study the problem of energy cooperation between service gateways that share some data collecting and transmission task, as presented on [2].
Our goal is to introduce and simulate a model that keeps this smart grid running, by considering rational energy consumption and providing the necessary incentives or behaviour restrictions for cooperation to occur.

2 Models

Here we present the two models: a centralized one, which we also call non-cooperating for reasons that will be clear after the numerical results; and a decentralized one, which we also call cooperating.

2.1 Centralized Model

First, we consider a centralized model in which an exogenous agent restricts the action set available for every player at every game stage, following [1]. Formally we consider the following stochastic game:

Let $N = \{A, B\}$ be the set of players.
Let $\mathcal{A} = \{\text{Sell, Buy, Idle}\}$ be the action set.
Let $M = [0, C_{max}] \times [0, C_{max}]$ be the battery energy level state-space, which will condition the available actions and payoffs.
The game consists of $t = 0, 1, ..., T$ stages of the following stage game:

> Let $\{h_t^{(i)}\} \sim \mathcal{U}(0, H)$ i.i.d be the renewal energy harvested by player i at stage t.
> Let $\{c_t^{(i)}\} \sim \mathcal{N}(\mu, \sigma^2)$ i.i.d be the energy consumption of player i at stage t.
> Let $\{a_t^{(i)}\}$ be the energy exchanged by player i at stage t with the grid and/or the other player.
> All the above random variables are 0 at $t = -2$ and $t = -1$ by assumption.
> The battery energy levels $\{X_t^{(i)}\}$ determine the game state, which updates as:
>
> $$X_t^{(i)} = X_{t-1}^{(i)} + h_{t-1}^{(i)} - c_{t-1}^{(i)} + a_{t-1}^{(i)}$$
>
> The price of energy $p(t)$ updates with demand:
>
> $$p(t) = p(t-1) + \kappa \left[a_{t-1}^{(A)} + a_{t-1}^{(B)} - a_{t-2}^{(A)} - a_{t-2}^{(B)} \right]$$

Define $0 < b < 1$ and $s > 1$ as the incentive price rate to buy/sell from/to the other player instead of the regular grid. This will ensure that players exchange as much

energy between them as possible. They are assumed to sell down to C_{down} and buy to C_{max}.
Now we introduce how the centralized agent restricts the actions, given the game state $\left(X_t^{(A)}, X_t^{(B)}\right)$.
Let $\sigma_t^{(i)}$ be the action taken by player i at stage t and define $\Delta_t^{(i)} \equiv h_{t-1}^{(i)} - c_{t-1}^{(i)}$ for convenience. Then:

$$\sigma_t^{(i)} \in \{\text{Buy}\} \text{ if } \Delta_t^{(i)} < 0 \text{ and } X_t^{(i)} + \Delta_t^{(i)} \leq C_{down}$$
$$\sigma_t^{(i)} \in \{\text{Sell}\} \text{ if } \Delta_t^{(i)} < 0 \text{ and } X_t^{(i)} + \Delta_t^{(i)} > C_{down}$$
$$\sigma_t^{(i)} \in \{\text{Sell}\} \text{ if } \Delta_t^{(i)} \geq 0 \text{ and } X_t^{(i)} + \Delta_t^{(i)} \geq C_{max}$$
$$\sigma_t^{(i)} \in \{\text{Sell, Idle}\} \text{ if } \Delta_t^{(i)} \geq 0 \text{ and } X_t^{(i)} + \Delta_t^{(i)} < C_{max}$$

Define for convenience the energy exchanged by player i at stage t due to buying or selling:

$$a_{t,sell}^{(i)} \equiv X_t^{(i)} + \Delta_t^{(i)} - C_{down}$$
$$a_{t,buy}^{(i)} \equiv C_{max} - \left(X_t^{(i)} + \Delta_t^{(i)}\right)$$

Finally we can define the payoffs for the stage game. Using the following notation $v_t^{(i)}\left(\sigma_t^{(i)}, \sigma_t^{(-i)}\right)$ we have:

$$v_t^{(i)}(\text{Sell, Sell}) = v_t^{(i)}(\text{Sell, Idle}) = p(t)\, a_{t,sell}^{(i)}$$
$$v_t^{(i)}(\text{Buy, Buy}) = v_t^{(i)}(\text{Buy, Idle}) = -p(t)\, a_{t,buy}^{(i)}$$

$$v_t^{(i)}(\text{Sell, Buy}) = p(t)\Big\{\Theta\left(a_{t,sell}^{(i)} - a_{t,buy}^{(-i)}\right) \left[s \cdot a_{t,buy}^{(-i)} + \left(a_{t,sell}^{(i)} - a_{t,buy}^{(-i)}\right)\right] + s \cdot a_{t,sell}^{(i)} \left[\Theta\left(a_{t,buy}^{(-i)} - a_{t,sell}^{(i)}\right) + \delta\left(a_{t,buy}^{(-i)} - a_{t,sell}^{(i)}\right)\right]\Big\}$$

$$v_t^{(i)}(\text{Buy, Sell}) = -p(t)\Big\{\Theta\left(a_{t,buy}^{(i)} - a_{t,sell}^{(-i)}\right) \left[b \cdot a_{t,sell}^{(-i)} + \left(a_{t,buy}^{(i)} - a_{t,sell}^{(-i)}\right)\right] + b \cdot a_{t,buy}^{(i)} \left[\Theta\left(a_{t,sell}^{(-i)} - a_{t,buy}^{(i)}\right) + \delta\left(a_{t,sell}^{(-i)} - a_{t,buy}^{(i)}\right)\right]\Big\}$$

$$v_t^{(i)}(\text{Idle, Idle}) = v_t^{(i)}(\text{Idle, Sell}) = v_t^{(i)}(\text{Idle, Buy}) = 0$$

As we can see, the utility in this model is purely the profit made off the energy transfer, and the role of the exogenous central agent is to incentives energy cooperation by subsidizing energy costs inside the smart grid, while restricting the behaviour by imposing the available actions from the game state. In this case the equilibrium is trivial: just take the available action at every stage, and always sell instead of being idle.
This model will now be contra-posed with a decentralized one in which we lose the restrictions and introduce a penalization in the utility if the player goes down. We will see that using some strategy we can keep energy cooperation in this way, and reach a more efficient scenario.

2.2 Decentralized Model

In this model we lift the restrictions and release the exogenous agent. Instead every player decides for their own action. The utility stays unchanged but we introduce a penalization based on their energy reservoir, demand:

$$\text{Buy: } \Delta_t^{(i)} + X_t^{(i)} < C_{down}$$

$$\text{Sell: } \Delta_t^{(i)} + X_t^{(i)} > C_{down}$$

$$\text{Idle: } C_t^{(i)} < C_{down}$$

In this case the player chooses the action which grants her/him the most utility:

$$\sigma_t^{(i)} = max[v_{t,buy}^{(i)}, v_{t,sell}^{(i)}, v_{t,idle}^{(i)}]$$

For convenience the penalization is will be written in brackets and of course only applied if aforementioned conditions are fulfilled:

$$v_{t,buy}^{(i)} = v_{t-1}^{(i)} - p(t) a_t^{(i)}(-1000)$$

$$v_{t,sell}^{(i)} = v_{t-1}^{(i)} + p(t) a_t^{(i)}(-1000)$$

$$v_{t,idle}^{(i)} = v_{t-1}^{(i)}(-1000)$$

with the exchanged energy:

$$a_t^{(i)} \equiv \Delta_t^{(i)} + X_t^{(i)} - C_{down}$$

The utility in this model is not only based on the profit made but as well based on the utility from the day before hence the penalization even effects further time steps. The energy exchanged is now the same for different actions and the utility calculates the benefit for every single one. Due to this there is no trivial equilibrium and the player chooses unprejudiced the move with the biggest interest.
Derived from this model the introduced penalization forces the players to always store enough energy for the next time step and in consequence a blackout is prevented as shown in Figure 8. Furthermore the price stabilizes which indicates a far more stable model as you can see in Figure 10.

3 Numerical Results

In this section, we evaluate the behaviours of the centralised (non cooperative) and decentralised (cooperative) models.
We consider a GT management of two entities, both consumer and renewable producer at the same time, simulating in Python what happens to the battery levels and the energy price in many different fifty days simulations.

$h_t^{(i)}$ boundaries	(0.1,0.8)
Mean consumption	0.5
Minimum storage	0.3
Incentives to sell	1.5
Incentives to buy	0.5

Table 1: simulation parameters

		$P2$		
		B	S	I
	B	(v(B),v(B))	(v(B),v(S))	(v(B), 0)
$P1$	S	(v(S),v(B))	(v(S),v(S))	(v(S), 0)
	I	(0,v(B))	(0,v(S))	(0, 0)

Table 2: normal form of the cooperative game

For the centralised scenario the choice of the numerical parameters is crucial for the outcome of the experiment and concerns the amount of renewable energy harvested at a specific time slot that depends on a uniform distribution, where boundaries are set at the beginning, the energy consumption over the day, given by a normal distribution with $\sigma = 0.05$ and mean set at the start, we also decide the minimum amount of energy the devices need to remain lit and the incentives they have to exchange energy with other players instead of using the main grid.

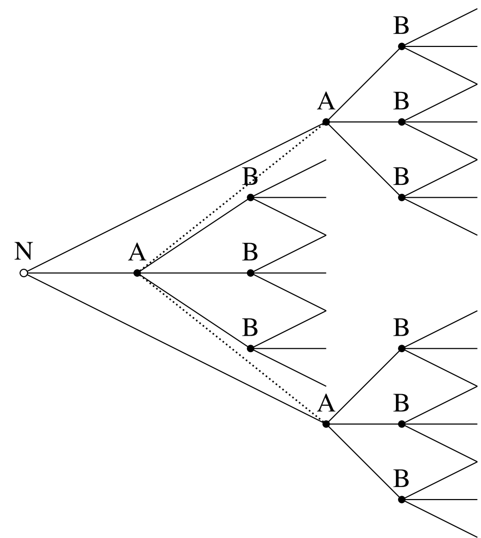

Figure 1: extensive form of the daily static game

In this scenario players are bounded to act accordingly to a specific set of rules we imposed that is mainly deterministic but it also has a certain degree of uncertainty given by the limited possibility of the players to act randomly during some stages of the game; but the system has also to deal with the randomness of some daily parameters that will influence the player attitude towards and we have introduced before.

We have two particular situations to inspect:

- $\sigma_t^{(i)} \in \{\text{Sell, Idle}\}$ if $\Delta_t^{(i)} > 0$ and $X_t^{(i)} + \Delta_t^{(i)} < C_{max}$ with probability $p = \frac{1}{2}$
- $\sigma_t^{(i)} \in \{\text{Buy, Idle}\}$ if $\Delta_t^{(i)} < 0$ and $X_t^{(i)} + \Delta_t^{(i)} > C_{min}$ with probability $p = \frac{1}{2}$

We show the results for the parameters set in Table 1, as expected there is no difference in the energy stocks and exchange behaviours due to the players acting under imposed rules Figure 2.

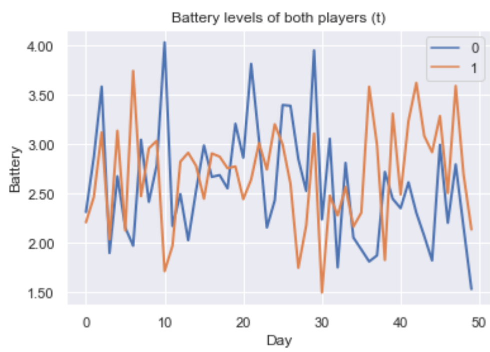

Figure 2: Battery levels, centralised, with incentives

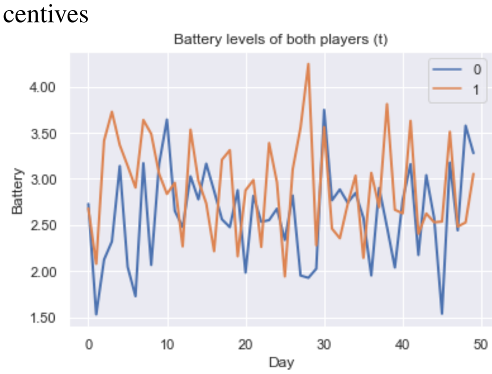

Figure 3: Battery levels, centralised, no incentives

The price of energy Figure 4 does not change under incentives policy considering the players only act on the actual storage and energy harvested so, at least now, the policy represents a waste of resources, while it could be worthy in a cooperative environment where we want for the players to exchange resources instead of relying on the external market (main grid).

The cooperative scenario, instead, gave players the possibility too cooperate in the exchange of energy instead of chasing all they can from the main grid, that is because the price directly depends on the amount of energy purchased

Figure 4: Battery levels, centralised, with incentives

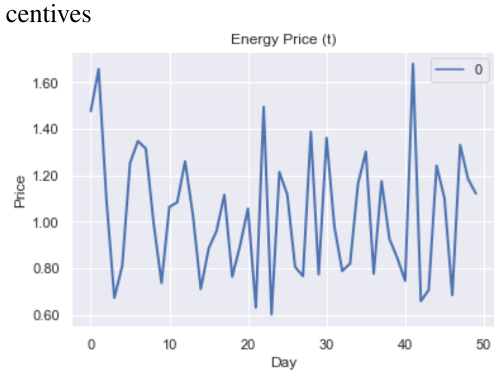

Figure 5: Battery levels, centralised, no incentives

in the previous stage that affects the value of the utility function Table 2. We used the same parameters set for the players and the results show a decrease in the price of the energy Figure 10 as we expected but also see a decrease in the energy storage Figure 8 due to better management of the resources.

4 Conclusion

The two models were simulated. We found the decentralized one to reach a more stable energy cooperation scenario. Thus, the payoff penalization for letting the smart grid go down in the decentralized model is a good choice for creating an energy cooperation scenario with rational energy consumption, while the centralized one fails here.

References

[1] Artiom Blinovas et al. “A Game Theoretic Approach for Cost-Effective Management of Energy Harvesting Smart Grids”. In: *2022 International Wireless Communications and Mobile Computing (IWCMC)*. 2022, pp. 18–23. DOI: `10.1109/IWCMC55113.2022.9825181`.

[2] Ángel Fernández Gambin et al. “Energy cooperation for sustainable IoT services within smart cities”. In: *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. 2018, pp. 1–6. DOI: `10.1109/WCNC.2018.8377450`.

Appendix: figures

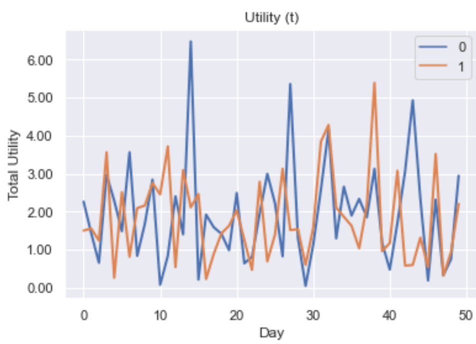

Figure 6: Utilities, centralised, with incentives

Figure 7: Utilities, centralised, no incentives

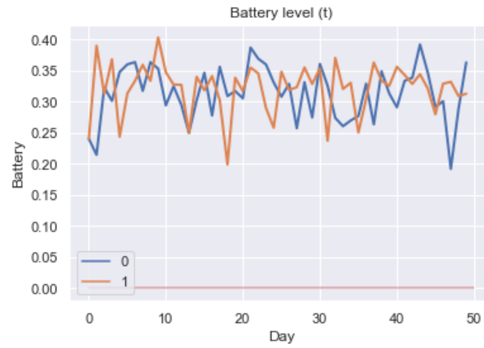

Figure 8: Battery levels, cooperative

Figure 9: Utilities, cooperative

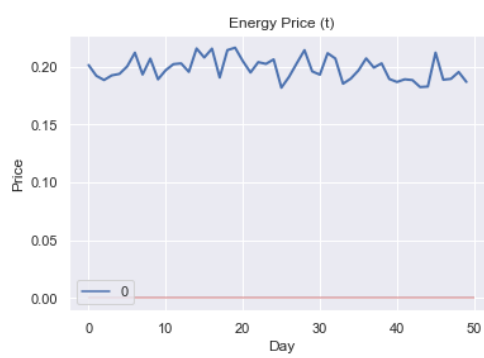

Figure 10: Price, cooperative