POWER AND ENERGY WEBINAR • SERIES

Powered by INESCTEC

Smart meter data sharing in energy systems

Yi Wang

The University of Hong Kong, yiwang@eee.hku.hk

15th May, 2024

Acknowledgment

- Thanks to my Ph.D. student, Yangze Zhou.
- The work is supported by the Research Grants Council of the Hong Kong SAR (HKU 27203723) and the National Key R&D Program of China (2022YFB2403300)

[1] Wang J, Gao F, Zhou Y, et al. Data sharing in energy systems[J]. Advances in Applied Energy, 2023, 10: 100132.

[2] Zhang J, Wang Y, Hug G. Cost-oriented load forecasting[J]. Electric Power Systems Research, 2022, 205: 107723.

[3] Sun Z, Von Krannichfeldt L, Wang Y. Trading and valuation of day-ahead load forecasts in an ensemble model[J]. IEEE Transactions on Industry Applications, 2023.

[4] Zhou Y, Wen Q, Song J, et al. Load data valuation in multi-energy systems: An end-to-end approach[J]. IEEE Transactions on Smart Grid, 2024.

Contents

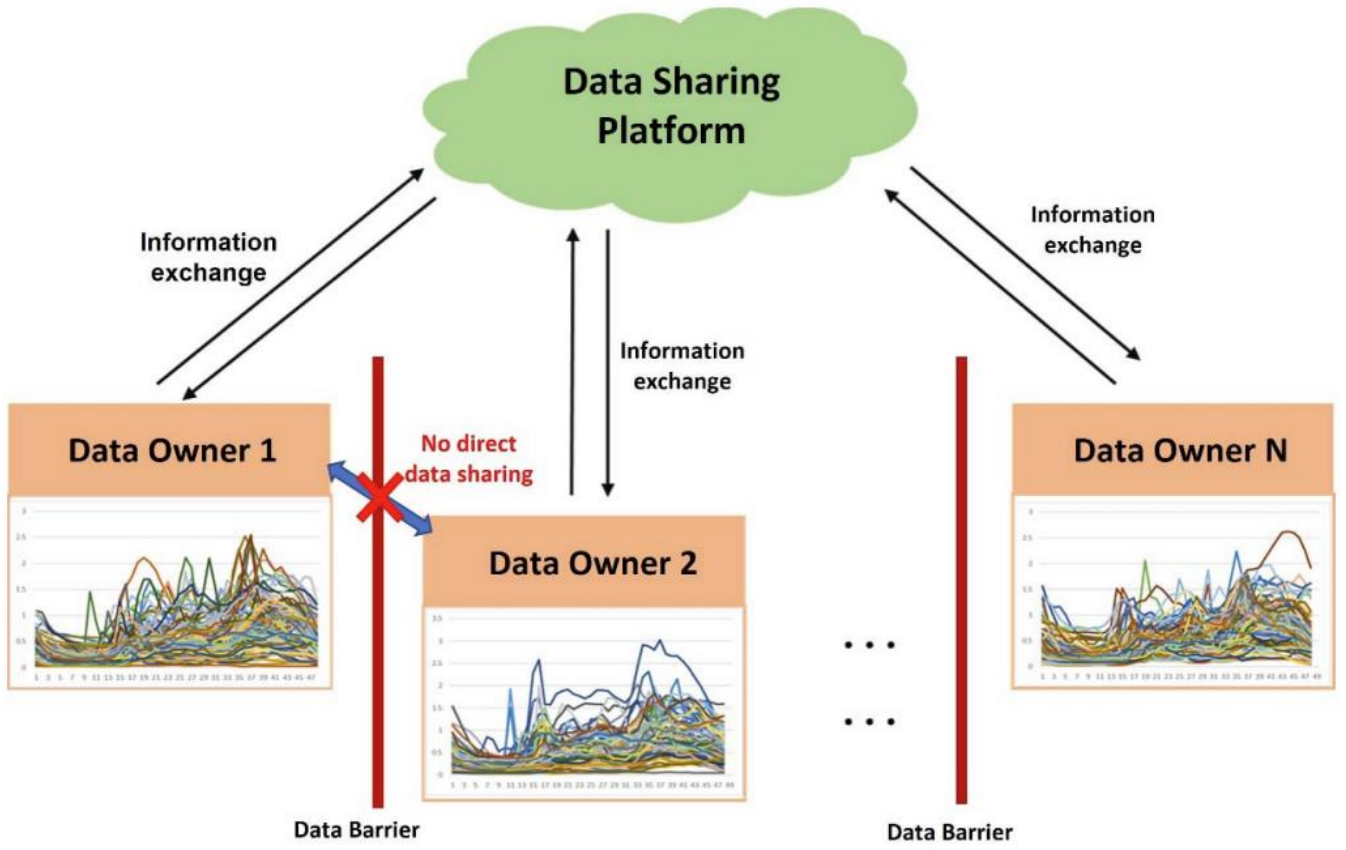
Background

- The aim of **carbon neutrality** pushes the emergence of an increasing number of **distributed energy resources.**
- Consequently, explosive amounts of energy **data** will be generated and **analyzed** in the future **smart cities.**
- Data + Analysis = Smart.

POWER AND ENERGY WEBINAR • SERIES
Powered by INESCTEC

Background

- However, there exits **data barriers…**

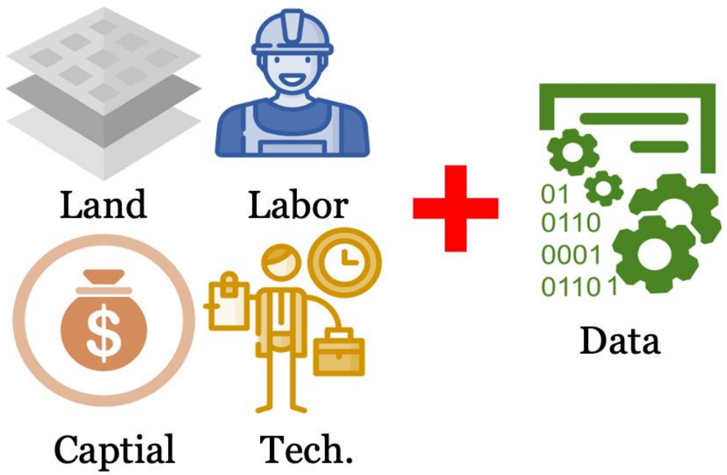


- **Laws & Regulations**
- **Business Competition**

POWER AND ENERGY WEBINAR • SERIES
Powered by INESCTEC

Background

- The Chinese government has issued a series of documents confirming that data is an **emerging production factor** reshaping market economy and driving the great waves of the next-generation industrial revolution.



- Some cities in China have already paid attention to the data barrier problem and set up **data exchange centers**.

Guiyang

Shanghai

Beijing

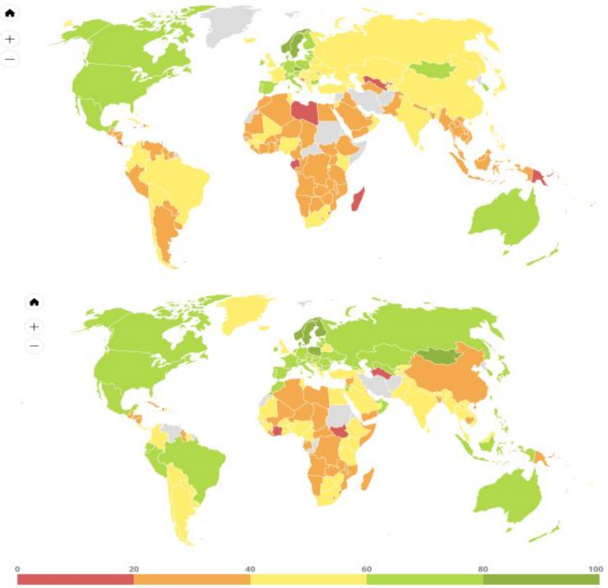
Shenzhen

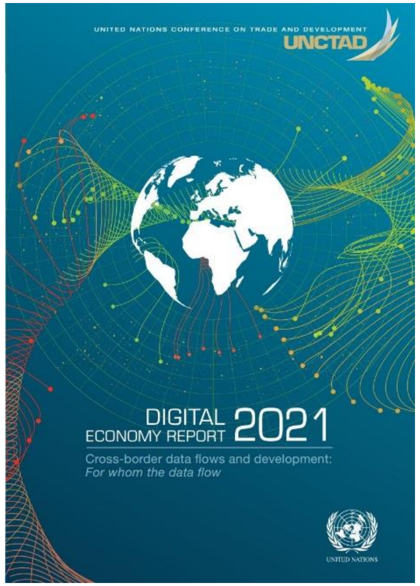
Background

- it has become a **global consensus** to facilitate the free flow of data across borders
- The 2021 Digital Economy Report on Trade and Development calls for a **global data governance approach** to facilitate the free flow of data across borders. The global data market is expected to exceed 22 billion dollars in 2024 (International Data Corporation, IDC).

2016

2022





2021 Digital Economy Report released by the United Nations Conference on Trade and Development

Contents

Overview

Data Products form?

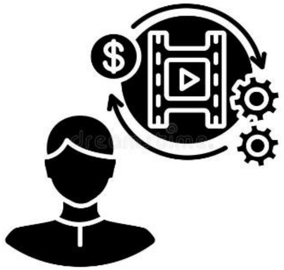
Daap
Raw data
Derived data

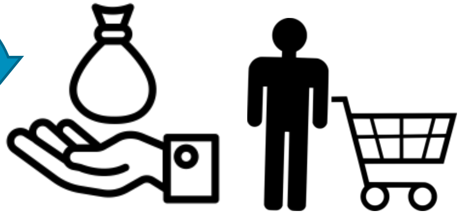
DaaS
Algorithms
Decision support
Automated decision making

Data Commodity
V.S.
Traditional Commodity

Different from traditional data?

Data Transaction





Data Owner

Selling Price?

Data Buyer

Purchasing Price?

- ✓ **Renewable generation forecasting**
- ✓ **Failure detection**
- ✓ **Demand side energy management**

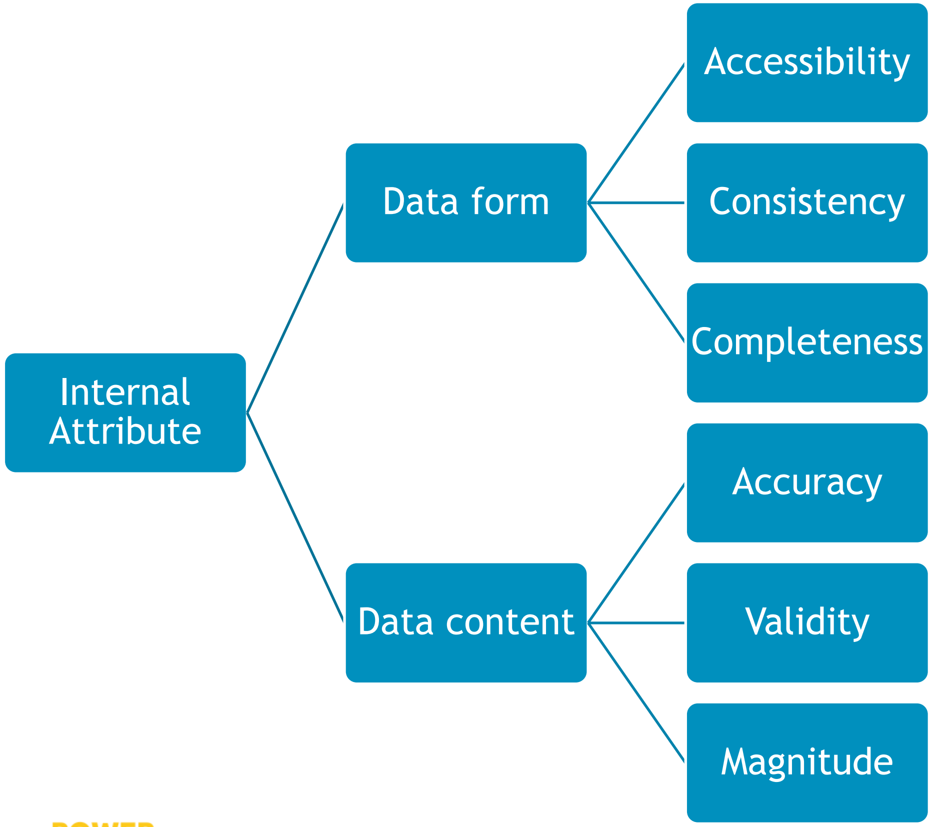
Which scenario?

How to clear the market?
Role of market operator?

POWER AND ENERGY WEBINAR · SERIES
Powered by INESCTEC

Data Attribute

- The attribute of data can be divided into **internal** and **external** attributes.
- The internal attributes include **data form** and **data content**.



Consistency: describe whether the information of same individual is the same in different data sets.

Completeness: describes the degree of how many data information is missing.

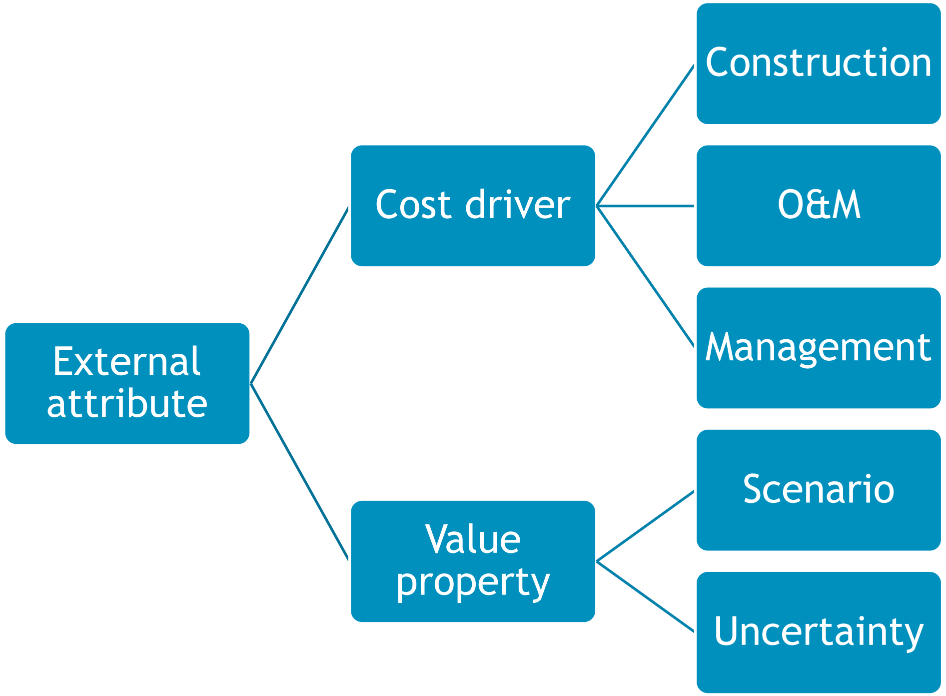
Validity: describes the degree to which the data follows the related rules and whether it conforms to its definition.

e.g., type of data, the format of the data, and the value range of the data.

POWER AND ENERGY WEBINAR · SERIES
Powered by INESCTEC

Data Attribute

- The external attributes are based on cost driver and value property.
- The value of data is dependent on the **applied scenarios**. Specifically, the value of data is obviously different in different scenarios, which brings about value uncertainty.



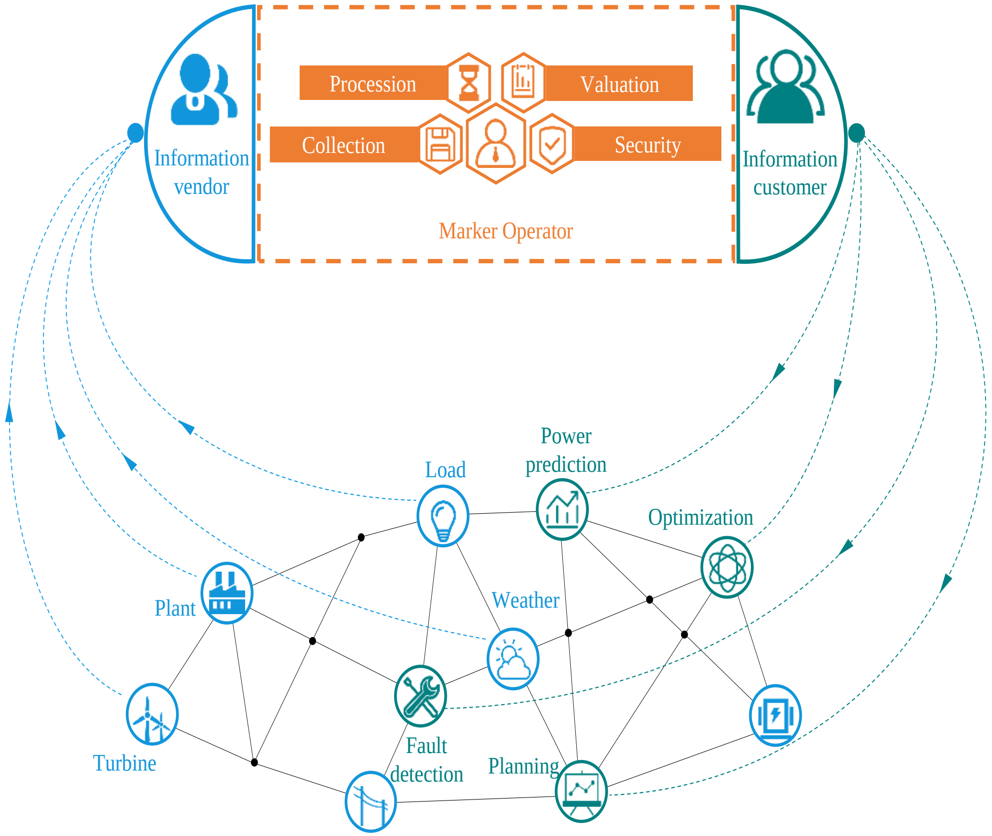
Cost attribute: In general, the costs contain three aspects: construction costs, operation and maintenance costs, and management costs.

Value property: describes the value of data is obviously different in different scenarios, which brings about value uncertainty.

POWER AND ENERGY WEBINAR · SERIES
Powered by INESCTEC

Market Framework

- Data market includes three types of entities: data suppliers/vendors, market operator and data customers.



Data providers are usually responsible for managing, organizing, and making sense of a great deal of data they collect and generate data collection equipment and can collect a large amount of raw data.

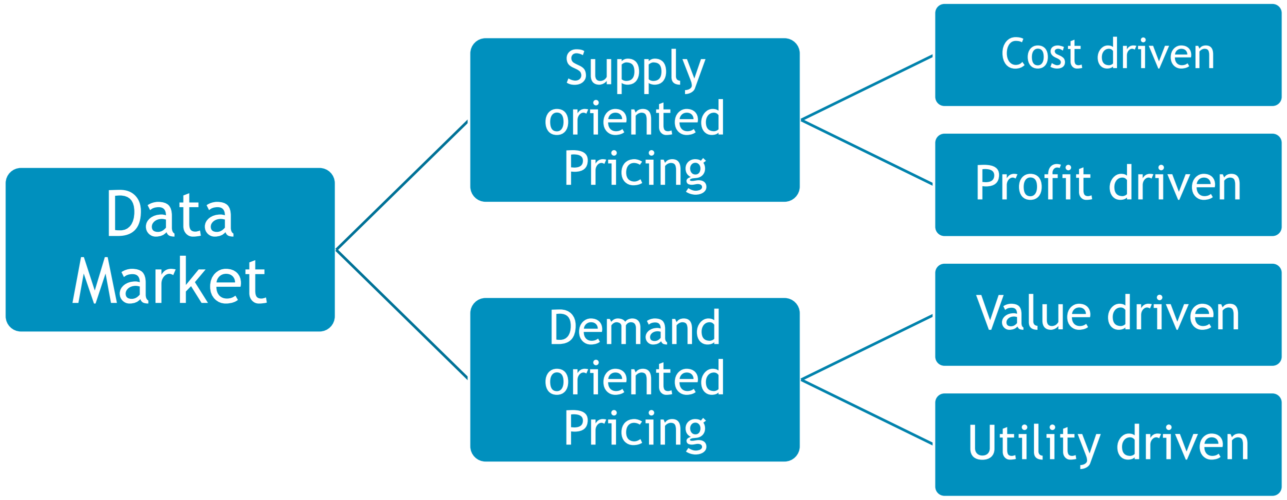
Market operator is responsible for operating the data market, processing the original data, transforming data into commodities and publishing data prices.

Data consumers purchase and extract values from these data products to make smarter decisions and ultimately gain more revenues.

POWER AND ENERGY WEBINAR • SERIES
Powered by INESCTEC

Data Valuation

- **Supply-oriented Strategy**: Supply-oriented pricing strategies are mostly seller-driven, emphasizing supplier profits
- **Demand-oriented strategy**: Related to the expected revenue of the product and the demand intensity of the consumer
- Intensity of demand ~ Quality of the data,
- Expected revenue ~ Benefits that the data bring to the user



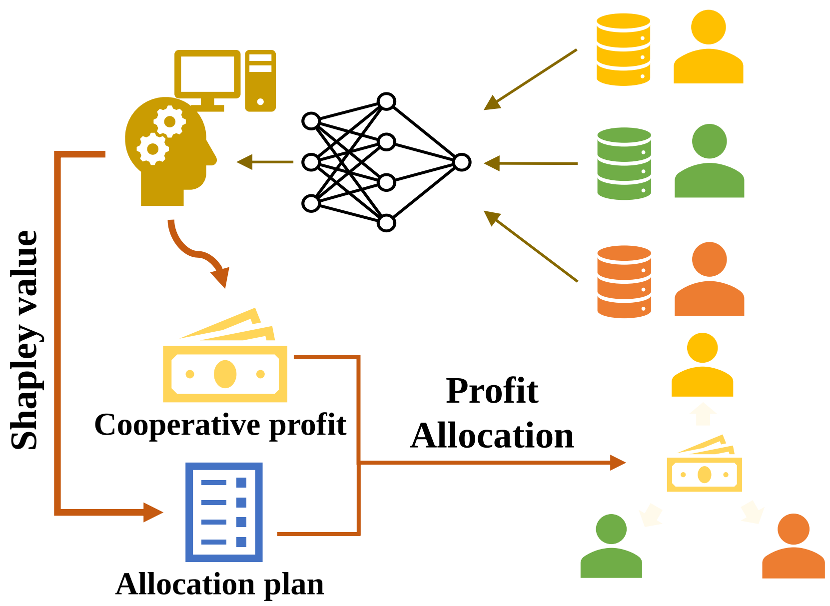
POWER AND ENERGY WEBINAR · SERIES
Powered by INESCTEC

Market Clearing

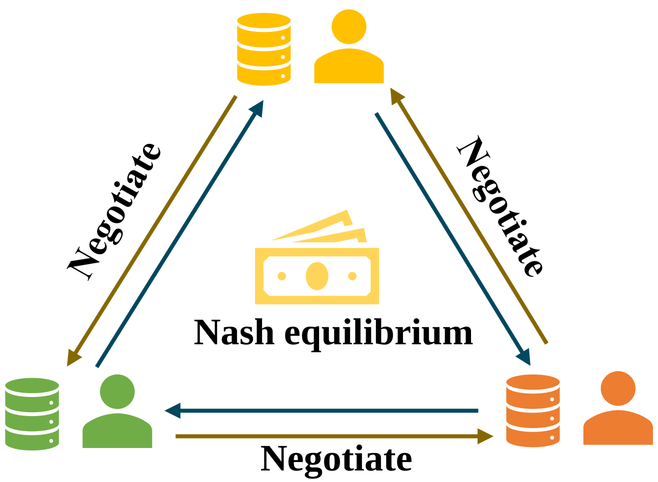
- Different market structure affects the market clearing:

  Game theory data product providers would like to cooperate and accomplish some tasks together.

  Non-cooperative games: competing behaviors and interest conflicts among data product providers and consumers are the main issues.



Cooperative game



Non-cooperative game

POWER AND ENERGY WEBINAR • SERIES
Powered by INESCTEC

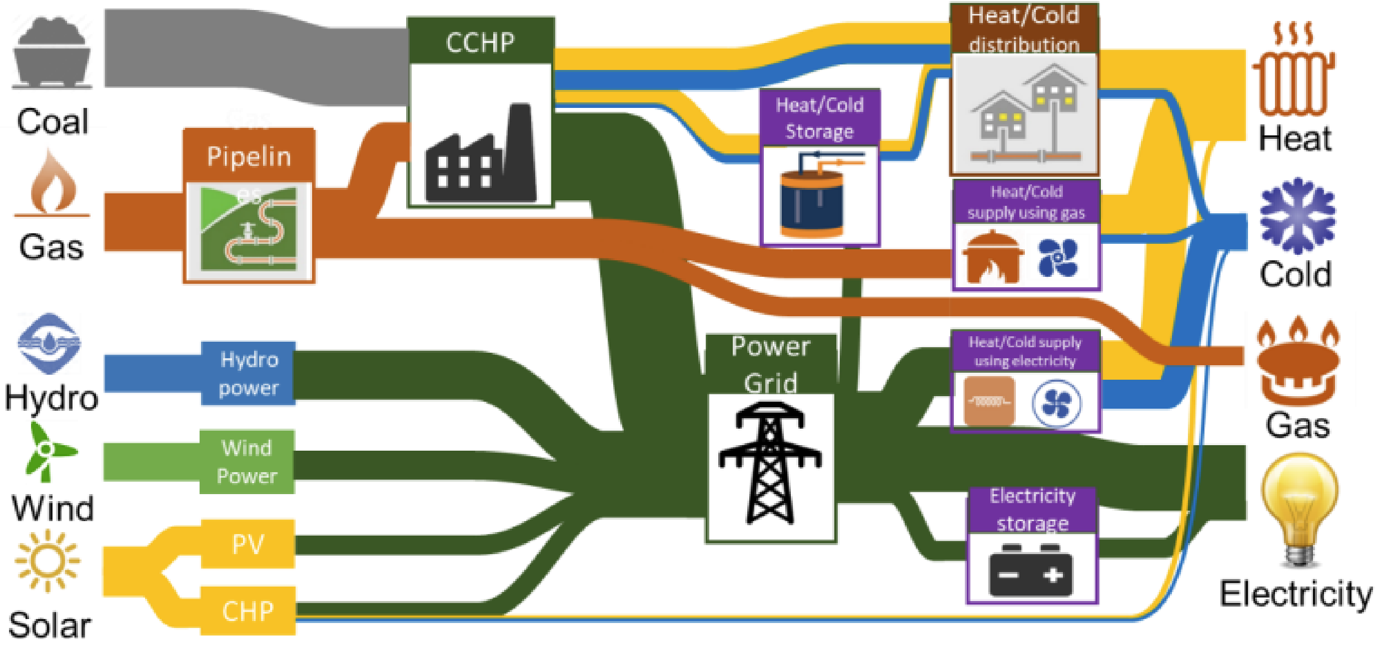
Contents

Forecasting & optimization

Multi-energy loads in MES are deeply couped and considering cross-sector information can improve forecasting accuracy. It means data sharing between different sectors in MES is meaningful.

However, power, gas, and heat/cooling load data are probably owned by different system operators separately. These data owners tend to prioritize their own economic benefits over social benefits when making decisions.

Only by reasonably valuing the data, will they be willing to share their data set.

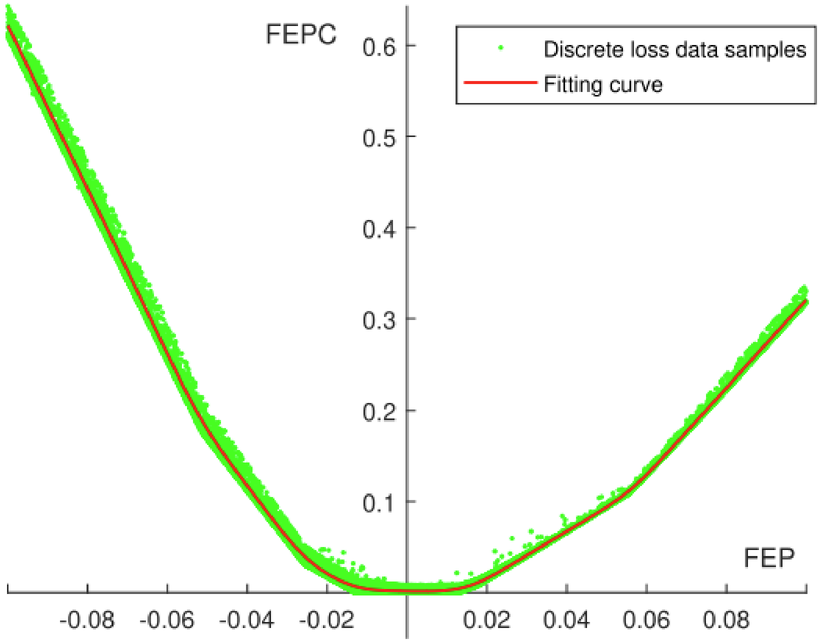


Forecasting & optimization

Economic value of data can be quantified by solving an optimization problem to minimize operation costs, the approach of **forecasting-then-optimization** (FTO) handles forecasting and decision-making as two **separate processes**.

In FTO, the forecasting model is trained with traditional loss functions such as mean square errors (MSE).

Treat the positive prediction errors and negative prediction errors equally using a quadratic function, whereas they exert different impacts on costs.



POWER AND ENERGY WEBINAR • SERIES
Powered by INESCTEC

Problem statement

X_n: Input feature of sector *n*

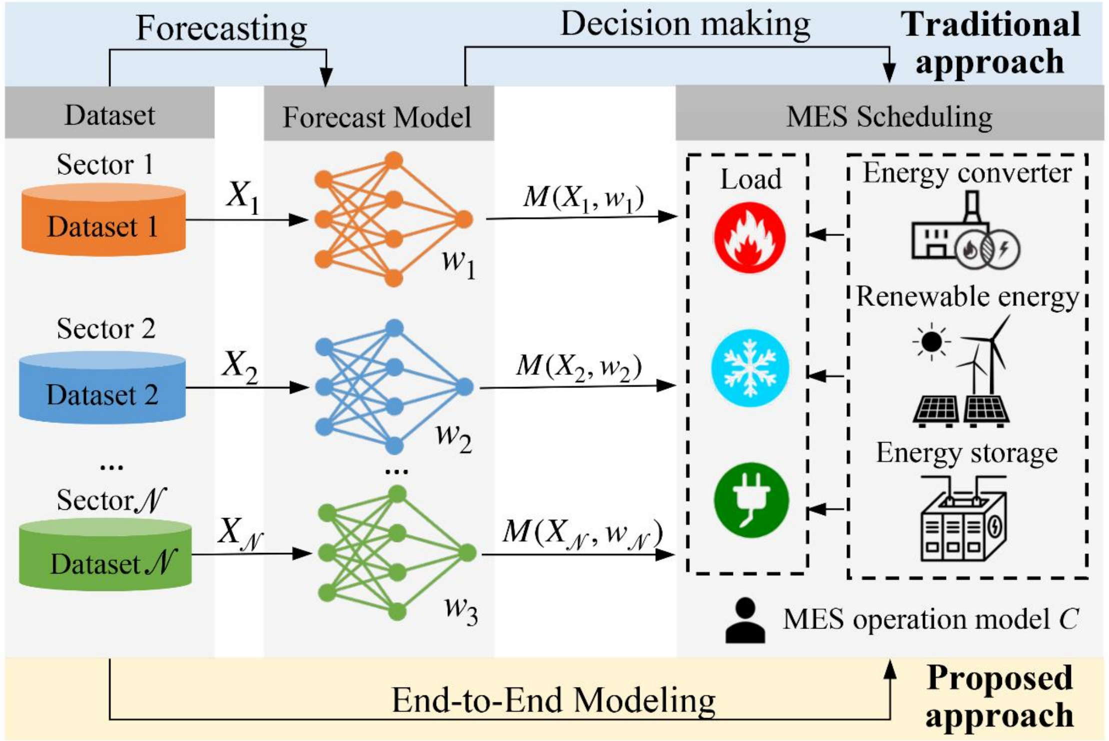
w_n: Model parameters of sector *n*

$M_n(X_n, w_n)|_{n\in N}$: Load forecasts of sector *n*

$$\min_{z} \; C(z, M_n(X_n, w_n)|_{n\in\mathcal{N}})$$

where C and z are the cost and decision variables for the scheduling of MES.

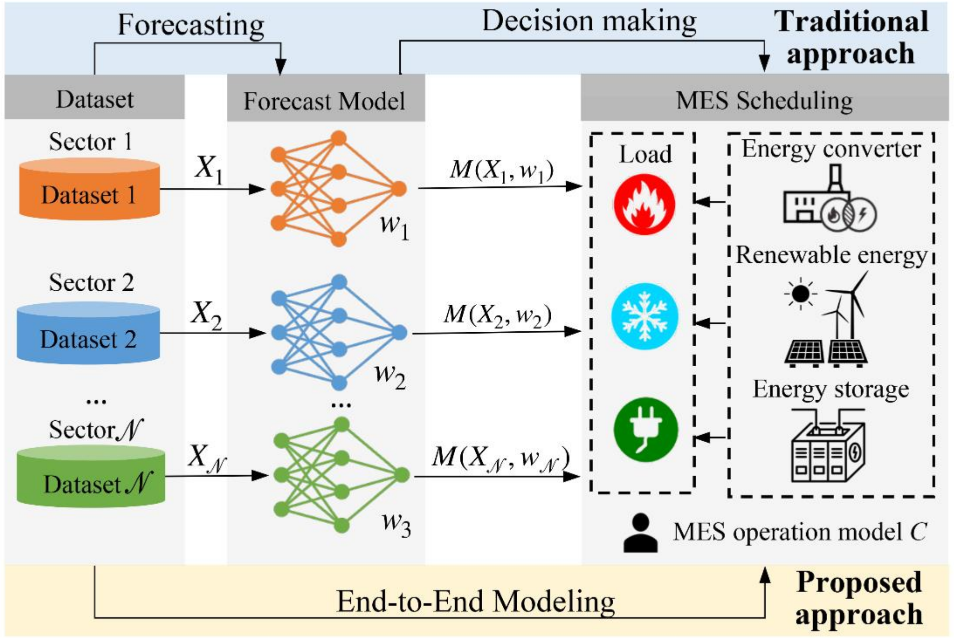
FTO: w_n and z are **determined sequentially.**



- Cross-sector data/information has not been shared and fully utilized to reduce operation costs.
- The forecasting and decision-making processes are treated separately so that data cannot directly serve final decision-making in MES.

POWER AND ENERGY WEBINAR · SERIES
Powered by INESCTEC

Problem statement



$$\min_{z;w_n|_{n\in\mathcal{N}}} C(z, M_n(X_n, w_n)|_{n\in\mathcal{N}})$$

C_N : The operation costs If all sectors cooperate with the MES operator, which means the sectors **share their data X_n with the operator indirectly**.

End-to-End approach: w_n and z are **optimized as a whole.**

To encourage sectors to participate in the end-to-end model, the value of the data owned by various sectors should be quantified:

1) How many additional profits $V(N)$ can be derived from data sharing of various sectors in MES ?

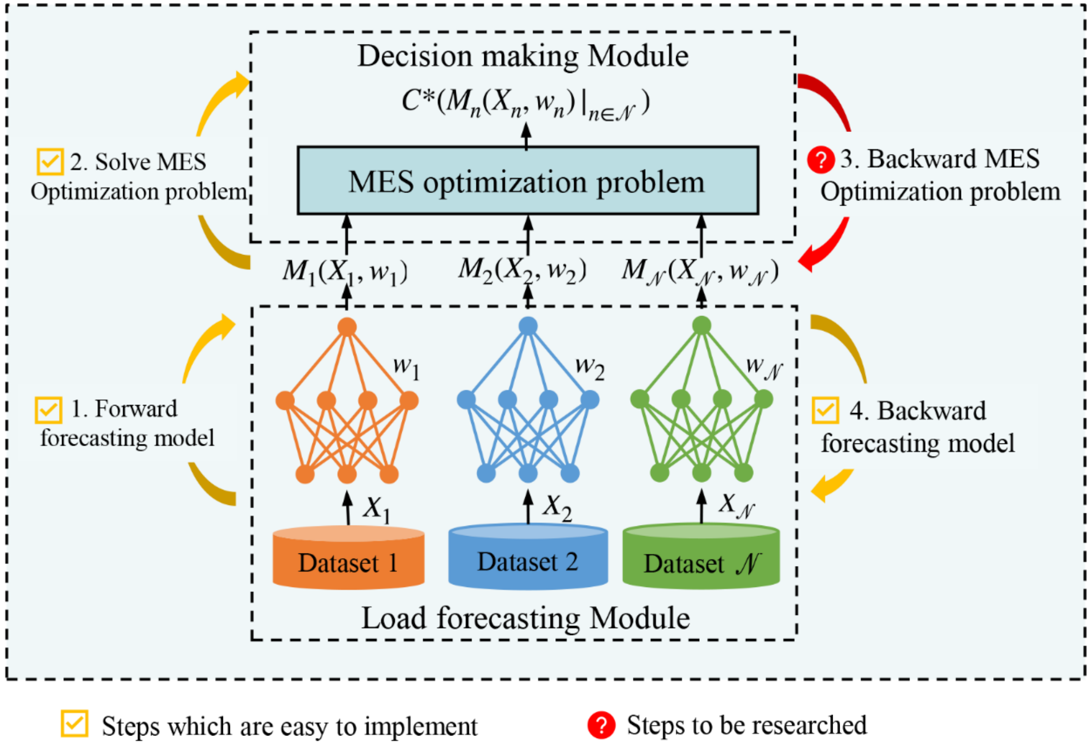
2) How to make a fair plan $\{v_1, v_2, \cdots, v_N\}$ to allocate the profits $V(N)$ to each sector ?

POWER AND ENERGY WEBINAR • SERIES
Powered by INESCTEC

End-to-end Modeling

How to optimize w_n and z as a whole?

An intuitive idea to train the end-to-end model is **forward and backward propagation**, as used for traditional neural network training.



How to obtain the gradient of cost C over load forecasts M ?

Optimization differentiable neural network (OptNet)

POWER AND ENERGY WEBINAR · SERIES
Powered by INESCTEC

End-to-end Modeling

How to obtain the gradient cost C over load forecasts M ? Chain principle $\frac{dC}{dM} = \frac{dC}{dz}\frac{dz}{dM}$

The MES optimization problem can be abstracted:

$$\begin{aligned} \min \quad & C(z, M) \\ \mathrm{s.t.} \quad & f(z, M) \le 0, h(z, M) = 0 \end{aligned}$$

The Lagrange function of the optimization problem

$$\mathcal{L}(z, \lambda, \mu, M) = C(z, M) + \lambda^T f(z, M) + \mu^T h(z, M)$$

The KKT condition of $\mathcal{L}(z, \lambda, \mu, M)$:

$$\begin{cases} f(z, M) \le 0 \\ h(z, M) = 0 \\ \lambda_i \ge 0, i \in \{1,2,\cdots,q\} \\ \lambda_i f_i(z, M) = 0, i \in \{1,2,\cdots,q\} \\ \nabla_z \mathcal{L}(z, \lambda, \mu, M) = 0 \end{cases}$$

Implicit function:

$$G(\tilde{z}, M) = \begin{bmatrix} \nabla_z \mathcal{L}(z, \lambda, \mu, M) \\ \lambda f(z, M) \\ h(z, M) \end{bmatrix}$$

The gradient of $\tilde{z}$ over M can be obtained by the differential principle of implicit function:

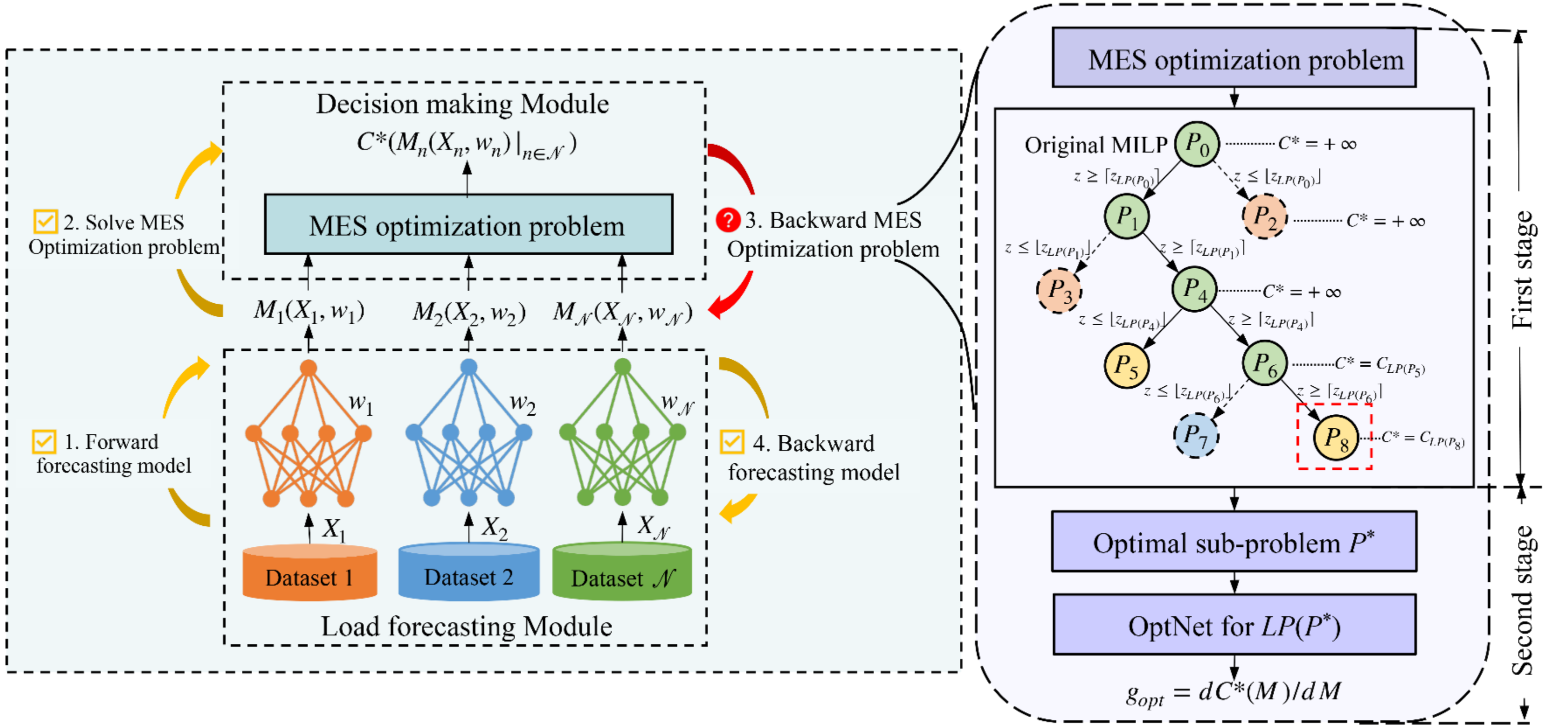
$$\frac{d\tilde{z}}{dM} = G_{\tilde{z}}^{-1}(\tilde{z}, M) G_M(\tilde{z}, M)$$

where

$$\frac{d\tilde{z}}{dM} = \left[\boxed{\frac{dz}{dM}} \quad \frac{d\lambda}{dM} \quad \frac{d\mu}{dM}\right]^T$$

End-to-end Modeling

However, OptNet is designed for LP/QP problem, What if there are integer variables in the optimization problem?



OptNet-embedded branch and bound method: How about incorporating OptNet into the branch and bound search process (Construct OptNet for each yellow node)?

Higher computational complexity and storage requirements.

Data Valuation

Additional Profit Quantification

The reduced operation costs can be regarded as the additional profits derived from the data sharing.

$$V(N) = C_N - C_\emptyset$$

End-to-End data valuation:

1) Each sector $n \in N$ utilizes its own data to develop the basic forecasting model M_n.

2) Computing operation costs of the traditional FTO approach $C_\emptyset$.

3) Integrating the forecasting model with the MES optimization problem for end-to-end model training.

4) Forward-propagating the end-to-end model to calculates the operation costs C_N.

Basic model development:

for *each sector* $n \in \mathcal{N}$ **do**

Random initialize parameters $w_n|_{n\in\mathcal{N}}$

for $k \in [0, E_1]$ **do**

$M_n = M_n(X_n; w_n^{(k)})$

$g_n = \text{Backward}(L_{MSE}(M_n, M_n^{\text{real}}))$

$w_n^{(k+1)} = w_n^{(k)} - lr \cdot g_n$

Return $M_n|_{n\in\mathcal{N}}$

End-to-End data valuation:

$C_\varnothing$ **Calculation:**

$C_\varnothing = \min_z C(z, M_n(X_n, w_n)|_{n\in\mathcal{N}})$

End-to-End modeling:

for $k \in [0, E_2]$ **do**

$M = M_n(X_n, w_n)|_{n\in\mathcal{N}}$

$P^* =$ Optimal sub-problem of $P(z, M)$

Construct OptNet for $LP(P^*)$

for *sector* $n \in \mathcal{N}$ **do**

$g_{\text{opt},n} = \text{Backward}(\text{OptNet})$

$g_n = \text{Backward}(g_{\text{opt},n})$

$w_n^{(k+1)} = w_n^{(k)} - lr \cdot g_n$

$C_\mathcal{N}$ **Calculation:**

$C_\mathcal{N} = \min_z C(z, M_n(X_n, w_n)|_{n\in\mathcal{N}})$

Additioanl profits quantification:

$V(\mathcal{N}) = C_\mathcal{N} - C_\varnothing$

Return $V(\mathcal{N})$

POWER AND ENERGY WEBINAR · SERIES
Powered by INESCTEC

Data Valuation

Additional Profit Allocation

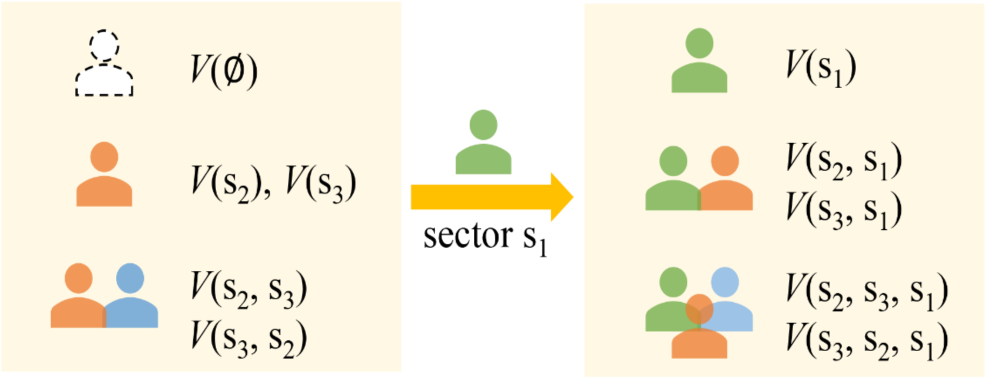
Shapley value has been widely adopted to measure the members' contributions to the collaboration earning.

Zero-Shapley value: $v_n = \frac{1}{|N|} \sum_{S \subseteq N \backslash \{n\}} \frac{1}{\binom{|N|-1}{|S|}} [V(S \cup \{n\}) - V(S)]^+$

Shapley value may be negative

$V(S)$: the value of the cooperation formed by union S

$[\cdot]^+ = \max\{0, \cdot\}$



$V(U)$ for $U \subset S \backslash \{s_1\}$ $V(U \cup s_1)$ for $U \subset S \backslash \{s_1\}$

Two remaining question:

- When some sectors within the MES do not participate in the end-to-end modeling, how to measure $V(S)$?
- The zero-Shapley value does not satisfy the budget balance property:

$$\Gamma(v_n) = \frac{v_n}{\sum_{i \in N} v_i} (V(N) - V(\emptyset))$$

Data Valuation

Additional Profit Allocation

When only sectors in U participate in the cooperation, how to measure $V(U)$?

End-to-End data valuation:

- $C_{\varnothing}$ **Calculation:**
  - $C_{\varnothing} = \min_{z} C(z, M_n(X_n, w_n)|_{n\in\mathcal{N}})$
- **End-to-End modeling:**
  - **for** $k \in [0, E_2]$ **do**
    - $M = M_n(X_n, w_n)|_{n\in\mathcal{N}}$
    - $P^* =$ Optimal sub-problem of $P(z, M)$
    - Construct OptNet for $LP(P^*)$
    - **for** *sector* $n \in \mathcal{N}$ **do**
      - $g_{\text{opt},n} = \text{Backward}(\text{OptNet})$
      - $g_n = \text{Backward}(g_{\text{opt},n})$
      - $w_n^{(k+1)} = w_n^{(k)} - lr \cdot g_n$
- $C_{\mathcal{N}}$ **Calculation:**
  - $C_{\mathcal{N}} = \min_{z} C(z, M_n(X_n, w_n)|_{n\in\mathcal{N}})$
- **Additioanl profits quantification:**
  - $V(\mathcal{N}) = C_{\mathcal{N}} - C_{\varnothing}$
- **Return** $V(\mathcal{N})$

$$\min_{z, w_n|_{n\in N}} C_n(z, M_n(X_n, w_n)\Big|_{n\in N})$$

$$\min_{z, w_n|_{n\in U}} C_n(z, M_n(X_n, w_n)\Big|_{n\in U}, M_n(X_n, w_n)\Big|_{n\in N\backslash U})$$

C_U denotes the operation costs of the **"partially integrated"** end-to-end model.

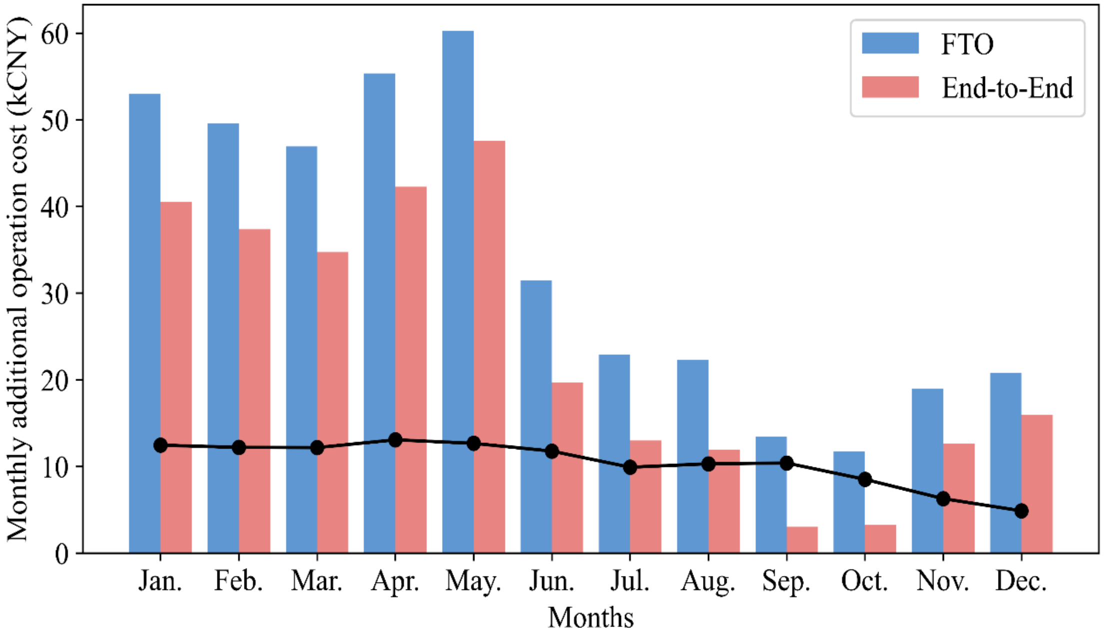
- Only sector $n \in U$ will update their model.
- Sectors in $N\backslash U$ will remain their **model parameters unchanged** (denoted as w_n) and **only submit final forecasts** $M_n(X_n, \overline{w}_n)|_{n\in N\backslash U}$ to the operator.

Case study

Daily operation costs for 12 months in 2017 (kCNY)

	FTO	End-to-End	Ideal	Improvement
Jan.	88.517	88.115	86.808	0.454 %
Feb.	87.500	87.064	85.730	0.498 %
Mar.	86.687	86.294	85.174	0.453 %
Apr.	88.015	87.579	86.170	0.495 %
May	88.282	87.873	86.338	0.463 %
Jun.	92.834	92.442	91.786	0.423 %
Jul.	95.141	94.821	94.402	0.336 %
Aug.	93.729	93.373	92.961	0.379 %
Sep.	92.823	92.477	92.375	0.373 %
Oct.	87.576	87.301	87.197	0.314 %
Nov.	86.505	86.231	85.681	0.316 %
Dec.	85.328	85.140	84.527	0.220 %

2017 all year:
Ideal: 31012.06 kCNY
FTO: 31418.71 kCNY (101.31% ideal cost)
End-to-End: 31294.04 kCNY (100.91% ideal cost)

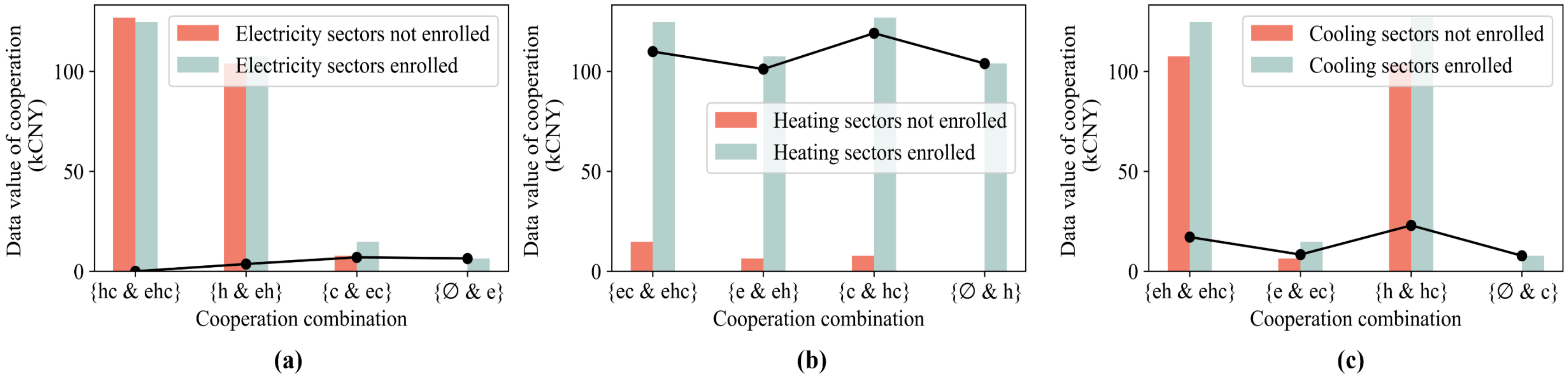


Monthly additional operation cost of FTO and end-to-end model compared to ideal cost (kCNY)

- ➢ Operation achieves a **0.40%** reduction, resulting in annual cost savings of **124.66 kCNY.**

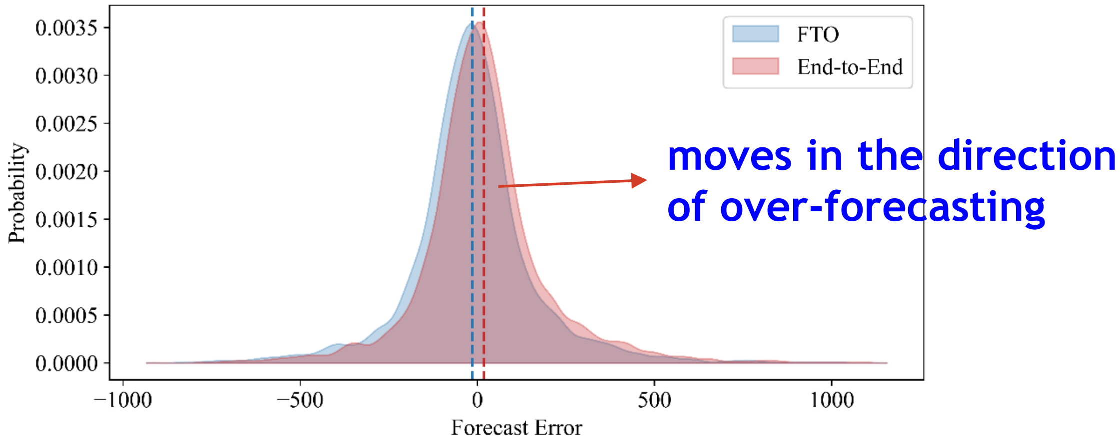
POWER AND ENERGY WEBINAR · SERIES
Powered by INESCTEC

Case study



- ➢ Electricity sector makes **little contribution.**
- • Accuracy of the electricity sector is relatively high.
- • The deviation of the electricity price in intra-day and day-ahead is relatively small.

- ➢ Heat sector cooperates with the MES operator can **markedly improve** additional profits.



The distribution of the forecasts error of heat sector

POWER AND ENERGY WEBINAR · SERIES
Powered by INESCTEC

Case study

Profit allocation result (kCNY)

Combinations		$e,\ h,\ c$	$e,\ h$	$e,\ c$	$h,\ c$	e	h	c	$\varnothing$
Operation costs $C_{(\cdot)}$		31294.04	31291.83	31311.15	31403.95	31412.30	31314.79	31410.94	31418.71
Data valuation $V(\cdot)$		124.66	126.87	107.56	14.76	6.40	103.92	7.77	0
Contributions	Electricity sector	$\Gamma(\frac{[V(e,h,c)-V(h,c)]^{+}+\frac{1}{2}[V(e,h)-V(h)]^{+}+\frac{1}{2}[V(e,c)-V(c)]^{+}+[V(e)-V(\varnothing)]^{+}}{3}) = 3.89$							
	Heat sector	$\Gamma(\frac{[V(e,h,c)-V(e,c)]^{+}+\frac{1}{2}[V(h,c)-V(c)]^{+}+\frac{1}{2}[V(e,h)-V(e)]^{+}+[V(h)-V(\varnothing)]^{+}}{3}) = 107.35$							
	Cooling sector	$\Gamma(\frac{[V(e,h,c)-V(e,h)]^{+}+\frac{1}{2}[V(e,c)-V(e)]^{+}+\frac{1}{2}[V(h,c)-V(h)]^{+}+[V(c)-V(\varnothing)]^{+}}{3}) = 13.43$							

Allocation fairness:

The heat sector's enrollment significantly reduces the operation costs → Largest share of profits.

The electricity sector has made the least contribution. → Least share of profits.

Incentive effect:

The electricity and cooling sectors: accuracy improvements & modest profits.

The heat sector: The degradation of the forecasting accuracy is compensated with economic reward.

Real-world Implementation

The signing of data sharing contract

- Before the data owners share their data, data owners must sign a **transaction agreement** with the MES operator.
- There are two important items in the sharing contract, which are the transmission of **price signals** and the **exit mechanism**.

Cost of data sharing

- These costs may include expenses related to data processing, storage, transmission, and privacy protection.
- If the cost of sharing becomes significant, the proposed framework can be **easily adjusted** to address this situation by deducting the costs from the additional profits.

Platform Support

- To support the real-world applications of the proposed data valuation method, w**ell-developed software or platforms** are necessary and achievable.
- In the software/platforms, the contents of the data-sharing agreements should be strictly enforced.
- A **centralized market** can reduce **communication requirements** and prevent the participation of **dishonest sectors**.

Contents

Conclusions

Conclusions

1. The construction of a data market is a promising solution that would allow for the well exploitation of data value to benefit data owners and break down data barriers.
2. We proposed an end-to-end data valuation approach, which is allocation fair, incentive effective and real-world implementable.
3. Conceive a specific and comprehensive data trading scenario to shorten the distance between academic research and industrial application.

Outlook

1. Current and ongoing research and applications are at a relatively preliminary stage.
2. Data value and data sharing platforms need to be cross-sectoral as well as cross-enterprise and cross-industry.

Thanks for your attention!

Yi Wang

The University o Hong Kong

May 2024