



数字化转型下的智能配用电大数据分析

第16期中国电工技术学会青年云沙龙
主题：新能源电力系统源-网-荷-储协调运行

王毅, 苏黎世联邦理工学院
yiwang@eeh.ee.ethz.ch
2021年5月16日

工作经历

2019.2- 博士后, 苏黎世联邦理工学院 (合作导师: Gabriela Hug教授)

教育经历

2014.9-2019.1 博士, 清华大学(导师: 康重庆教授)

2017.3-2018.4 访问学生, 华盛顿大学(导师: Daniel Kirschen教授)

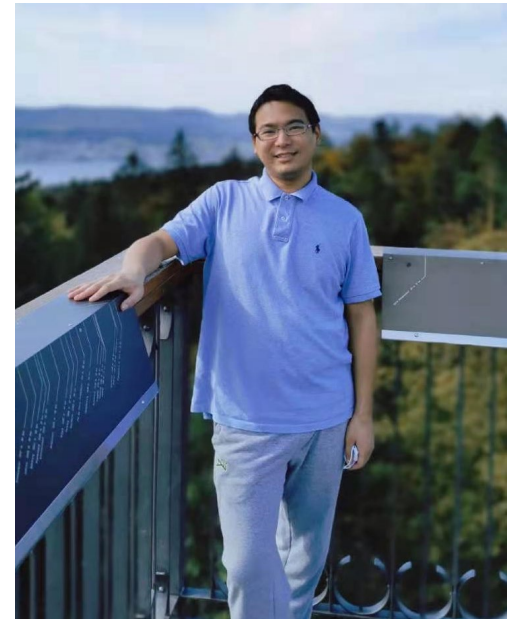
2010.9-2014.6 本科, 华中科技大学

研究方向

电力系统大数据分析

信息物理配电系统建模与分析

综合能源系统建模与优化

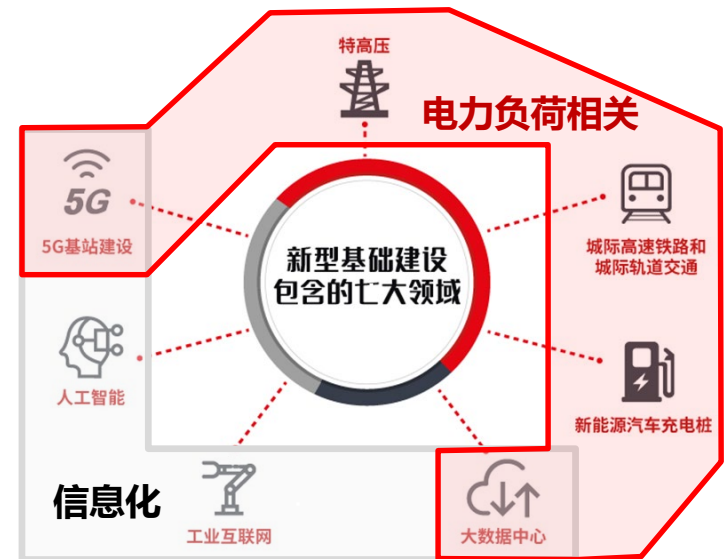


数据：重要生产要素

2020年4月9日 《关于构建更加完善的要素市场化配置体制机制的意见》

数据作为一种**新型生产要素**写入文件中，与土地、劳动力、资本、技术等传统要素并列为要素之一。加快培育数据要素市场，推进政府数据开放共享、提升社会数据资源价值、加强数据资源整合和安全保护。

新基建的设施建设和**电力行业息息相关。**



电力系统数字化转型

电力 **“数字化”** 将成为未来新一代电力系统的重要特征，将促使**发输配变用**各个环节的海量设备相连，**形成电力物联网**。

数据驱动的
分布式能源
供求平衡

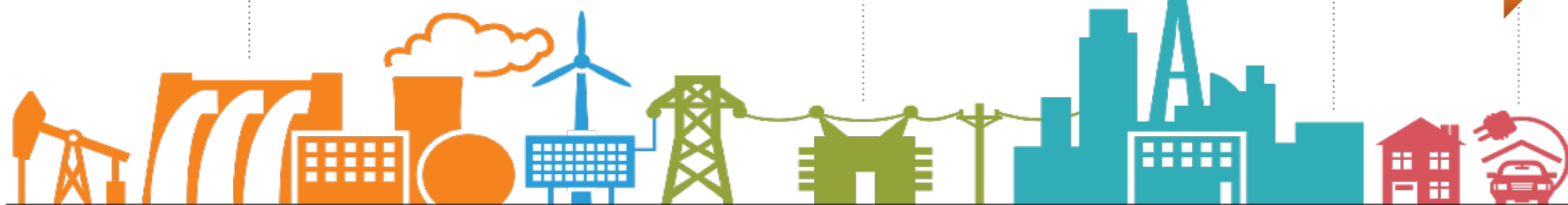
基于状态的
设备维护和
停机预测

通过自动控制
提高系统韧性、
安全性和效率

通过分析客户行为、
个性化交流促进与
客户的互动

构建平台以支持
分布式能源
及其交易市场

“源-网-荷-储” 多环节协同



云端自动化和
数据驱动决策
系统

现场工作人员可移
动访问地图、数据、
工作管理工具和实
时专业知识

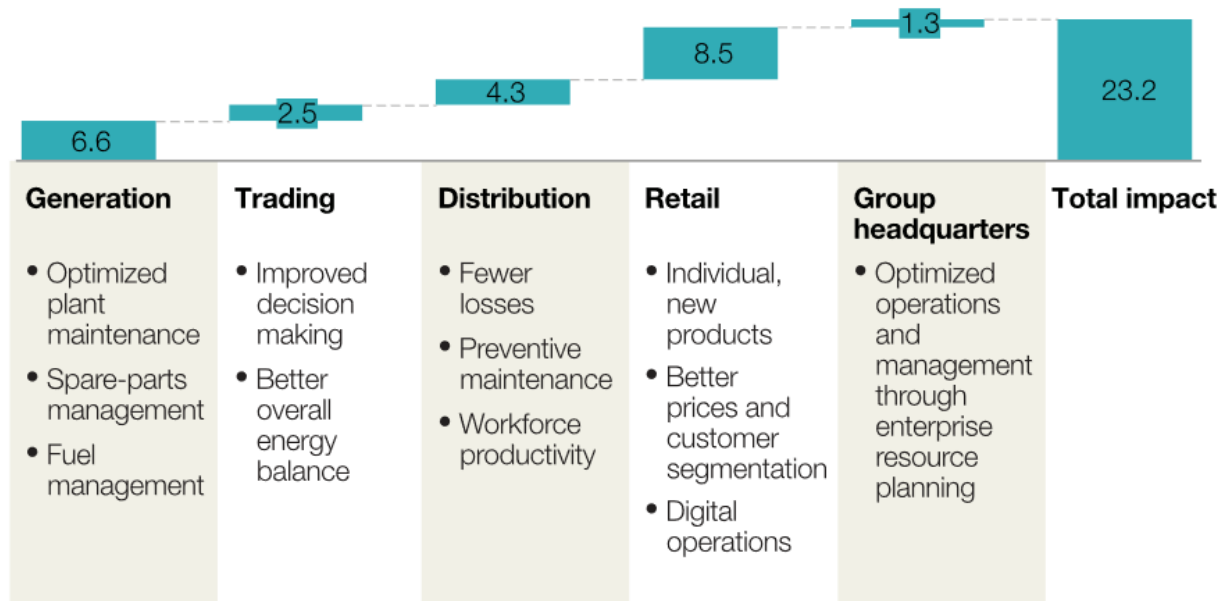
通过高水平的
态势感知能力
保障能源系统
实时平衡

电力系统数字化转型

McKinsey&Company咨询公司的一份报告对电力系统不同环节的“数字化”带来的效益进行了评估。其中，在配用电侧的效益提升一共高达14.1%！

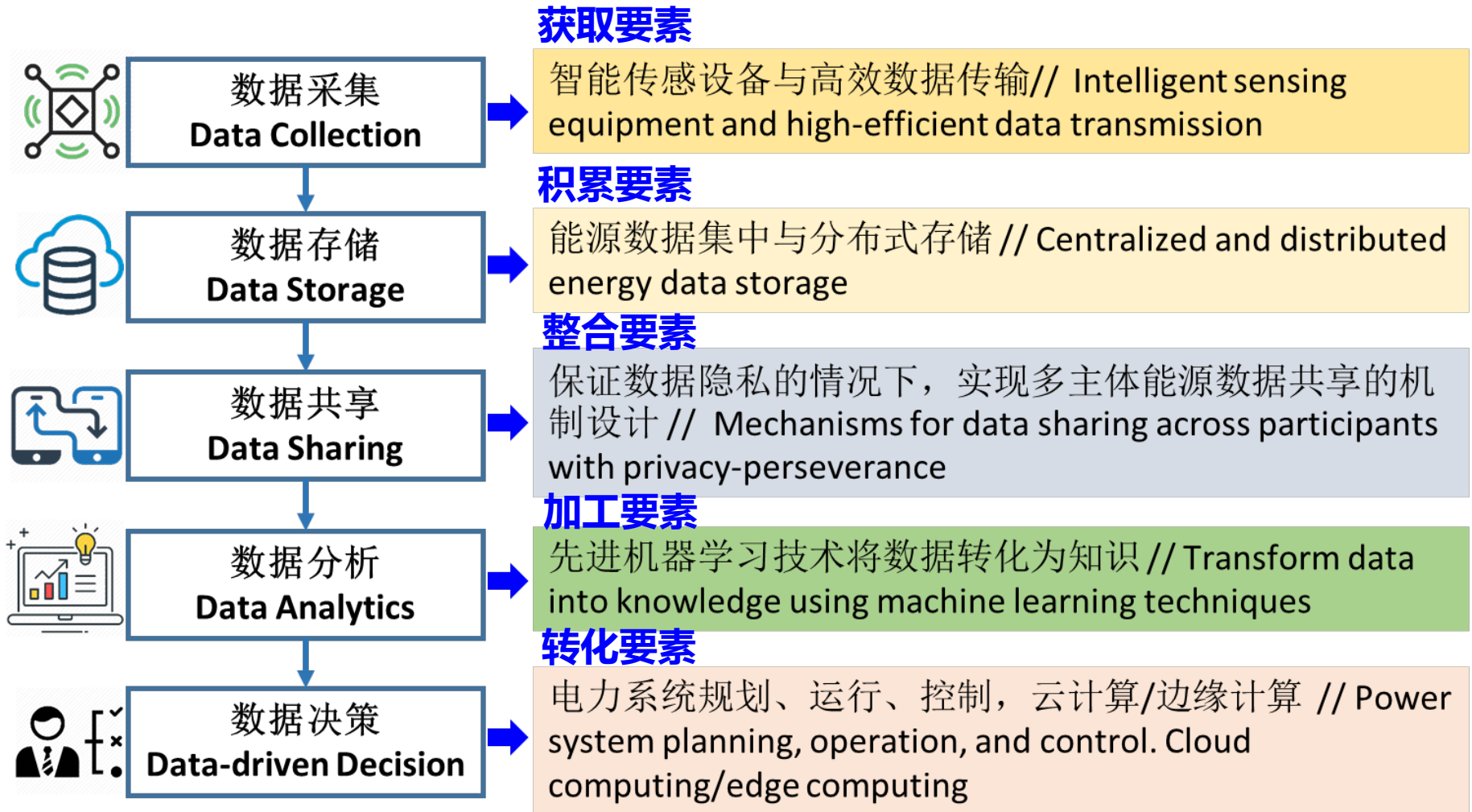
Digitization has demonstrable impact on utility earnings.

Improvement areas, case study, EBIT,¹ %



¹Earnings before interest and taxes.

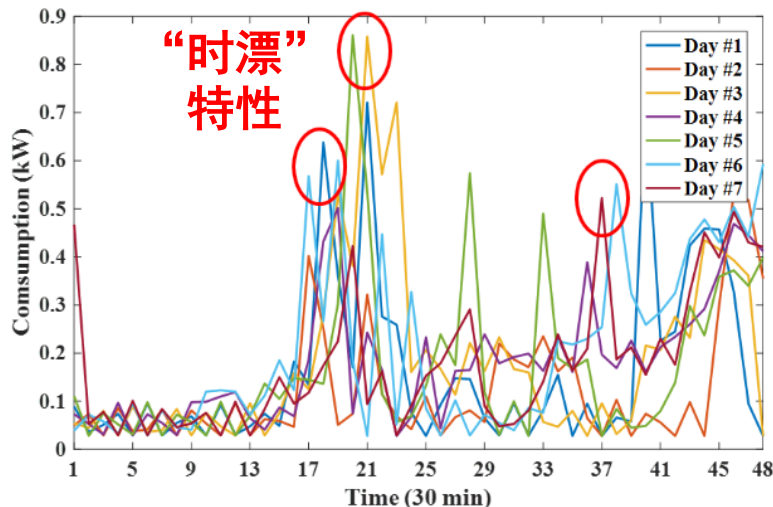
数据供应链



1. 数据分析——如何对要素进行加工？

◆ 用户肖像描绘技术

用户用电行为受到多方面因素的影响，包括**内在因素**如用户自身的社会经济信息和**外在因素**如环境温度等，其中**用户自身的社会经济信息**包括用户的职业、年龄、薪资水平、所居住房屋大小等。



研究意义：

- 个性化用电服务；
- 第三方机构推销；
- 隐私保护的反思；

1. 数据分析——如何对要素进行加工？

◆ 用户肖像描绘技术

建模方案：

将用户的性别、年龄、就业情况、社会阶层和住所等信息进行离散化，构成用户的“**标签**”，进而将用户社会经济信息辨识问题转化为分类问题。

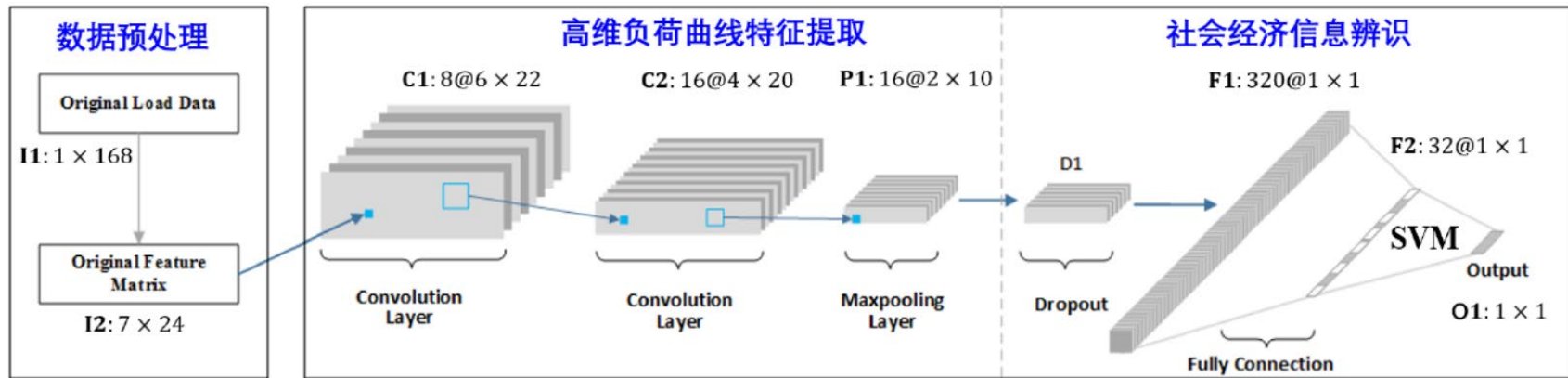
构建负荷曲线到用户信息的关系本质上是一个**高维且强非线性的分类问题**，构建分类模型前需要进行特征提取，以提升模型对非线性的描述能力。

编号	问卷编号	具体社会经济信息	选项	对应人数
1	300	主要收入者年龄	青年 (<35)	436
			中年 (35~65)	2819
			老年 (>65)	953
2	310	主要收入者是否退休	是	1285
			否	2947
3	401	主要收入者社会阶层	A 或 B	642
			C1 或 C2	1840
			D 或 E	1593
4	410	是否有小孩	是	1229
			否	3003
5	450	房屋类型	平房	2189
			楼房	1964
6	453	房屋新旧	旧 (>30)	2151
			新 (<30)	2077
7	460	房屋卧室数量	很少 (<3)	404
			少 (=3)	1884
			多 (=4)	1470
			较多 (>4)	474
8	4704	厨具类型	电厨具	1272
			非电厨具	2960
9	4905	节能灯占比	一半以下	2041
			一半以上	2191
10	6103	房屋面积	小 (<100)	232
			中等 (>100&<200)	1198
			大 (>200)	351

1. 数据分析——如何对要素进行加工？

◆ 用户肖像描绘技术

构建深度卷积学习网络，自动提取有效特征，考虑用电内在关联。



1、数据重排：问题的化归转化

2、特征提取

$$s_{i,j} = G_j(c_i, w_{1,j})$$

3、“标签”辨识

$$\hat{y}_{i,j} = F_j(s_{i,j}, w_{2,j})$$

➤ 利用数据拓展、随机遗忘、权重衰减、随机梯度下降训练等手段减少模型的过拟合风险。

层编号	层类型	超参数	待训练的参数数量
C1	卷积层	输入数据大小: $7 \times 24 \times 1$ 卷积核大小: 2×3 卷积核数量: 8	56
C2	卷积层	输入数据大小: $6 \times 22 \times 8$ 卷积核大小: 3×3 卷积核数量: 16	160
P1	最大化池化层	输入数据大小: $4 \times 20 \times 16$	无
Dr1	Dropout 层	无	无
F1	平铺层	无	无
D1	全连接层	输入数据大小: 320 神经元数量: 32	10560
D2	Softmax 层	输入数据大小: 32 神经元数量: 1	32

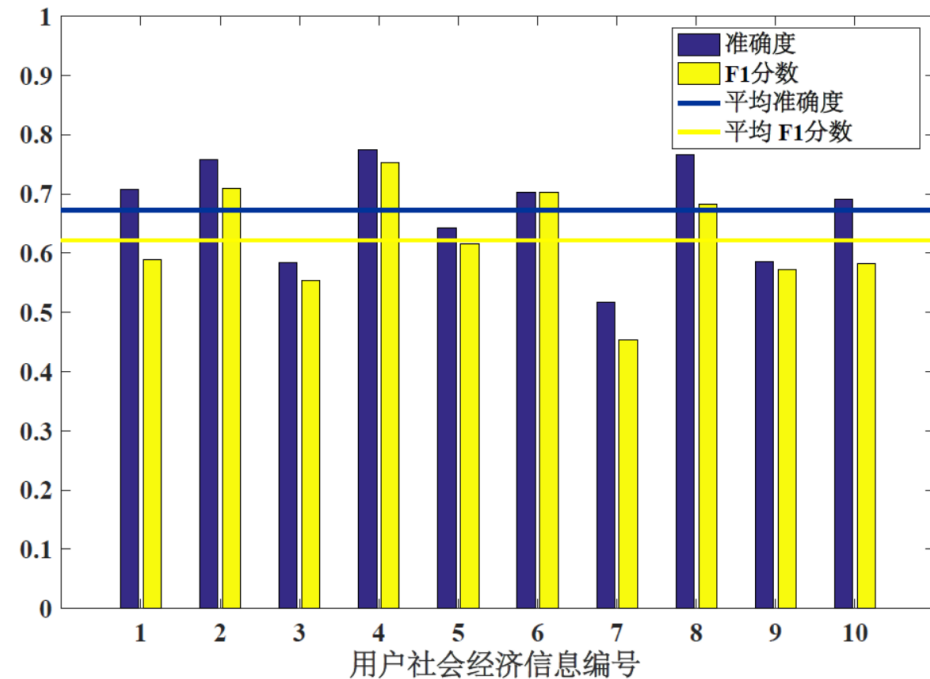
1. 数据分析——如何对要素进行加工？

◆ 用户肖像描绘技术

实验结果分析

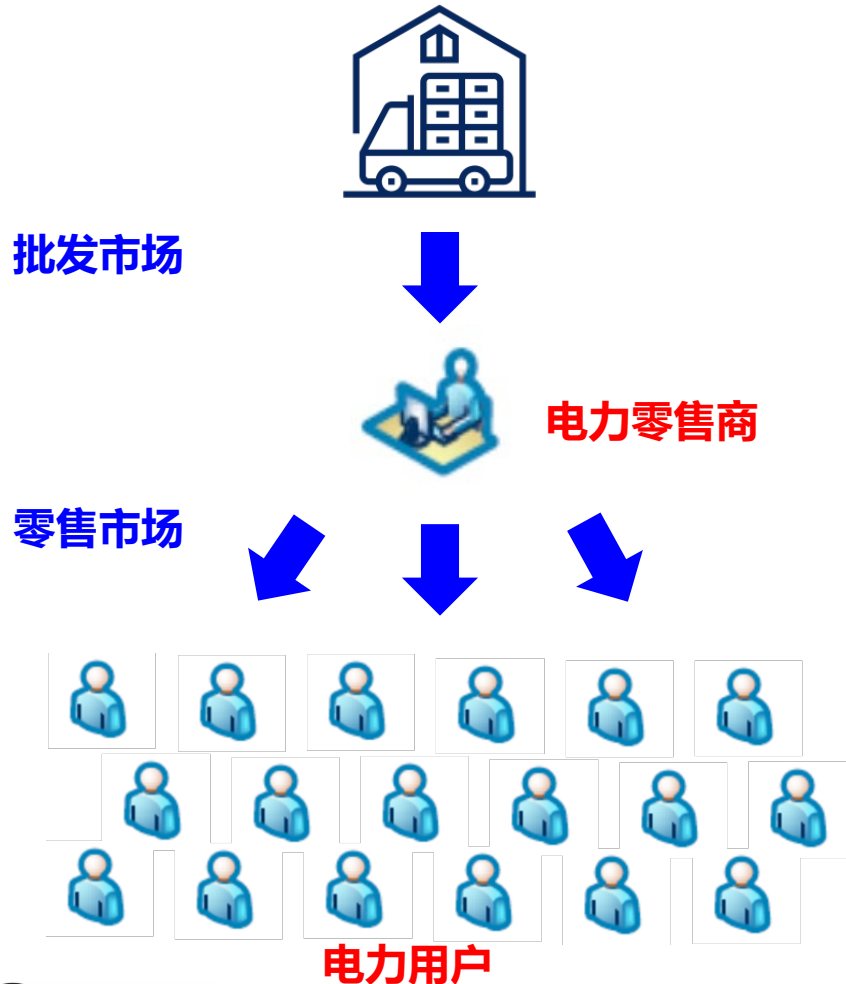
特征2（家里主心骨是否退休）、特征4（家里是否有小孩）和特征8（厨房是用电还是用气）的精度都超过75%，说明这些特征与用户用电行为具有较强的相关性；

特征7（卧室数量）和特征9（节能灯比例）的精度都低于60%，说明这些特征对用电行为的影响不明显。



2. 数据决策——如何将要素转化为生产力？

◆ 个性化电价设计



❖ 主从博弈模型

Leader——电力零售商

- 制定电价套餐
- 预测用户用电行为

Follower——电力用户

- 选择某一电价套餐
- 调整自身用电行为

2. 数据决策——如何将要素转化为生产力？

◆ 个性化电价设计

电价设计的主从博弈模型可以建模成为一个优化问题：

➤ **目标函数：** 最大化电力零售商利润

$$\max R = \sum_{r=1}^R \sum_{t=1}^T \underbrace{K_r \times p_{k,t} \times q_{k,t}}_{\text{用户支付电费}} - \sum_{t=1}^T \sum_{n=1}^{N_F} \underbrace{p_n^F \times L_n^F \times o_{n,t} \times o_n}_{\text{远期合约}} - \sum_{t=1}^T \underbrace{p_t^{D,est} \times L_t^D}_{\text{日前合约}} - \underbrace{\xi \times CVaR}_{\text{日前与实时市场风险损失}}$$

➤ **约束条件：** 负荷平衡

$$\sum_{r=1}^R \underbrace{K_r \times q_{r,t}}_{\text{用户用电负荷}} = \sum_{n=1}^{N_F} \underbrace{L_n^F \times o_{n,t} \times o_n}_{\text{远期合约}} + \underbrace{L_t^D}_{\text{日前合约}}, \forall t$$

➤ **约束条件：** 激励相容

$$U_r(\mathbf{p}_r) \geq U_r(\mathbf{p}') \quad \forall k$$

$$U_r(\mathbf{p}_r) \geq U_r(\mathbf{p}_0) \quad \forall k$$

选择电价套餐 $\mathbf{p}_r$ 是第 k 个用户的最佳策略，
较其他电价套餐，用户 k 更喜欢电价套餐 $\mathbf{p}_r$

选择电价套餐 $\mathbf{p}_r$ 是第 k 个用户的理性策略，
较传统电价模式，用户 k 更喜欢电价套餐 $\mathbf{p}_r$

* 标红的为非线性项

2. 数据决策——如何将要素转化为生产力？

◆ 个性化电价设计

优化模型——混合整数非线性优化问题

- **约束条件**：电力用户效用函数与响应行为

$$q_t = \left(\frac{p_t}{p_{t(0)}} \right)^{\frac{1}{\alpha-1}} \times q_{t(0)}$$

$$U(\mathbf{p}) = \sum_{t=1}^T \left(\frac{1}{\alpha} - 1 \right) \left[\left(\frac{p_t}{p_{t(0)}} \right)^{\frac{\alpha}{\alpha-1}} - 1 \right] \times q_{t(0)} p_{t(0)}$$

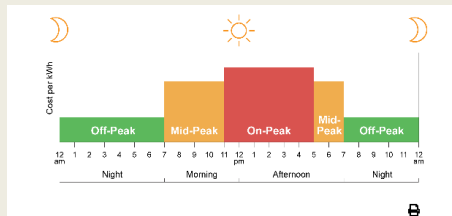
- **约束条件**：基于CVaR的风险度量

$$CVaR = \inf_{a \in R} \left\{ a + \frac{1}{(1 - \alpha^{CVaR}) \cdot N_S} \sum_{n_S=1}^{N_S} [(-\Delta R^D - \Delta \bar{R}^{RT}) - a]^+ \right\}$$

日前市场损失 实时市场损失

- **约束条件**：电价结构

Price category: CPP RTP ToU



分时电价应用较为广泛，实施效果好

$$\sum_{m=1}^M e_{r,t}^m = 1, \quad \sum_{t=1}^T e_{r,t}^m \geq D_{\min}, \quad \forall m, r$$

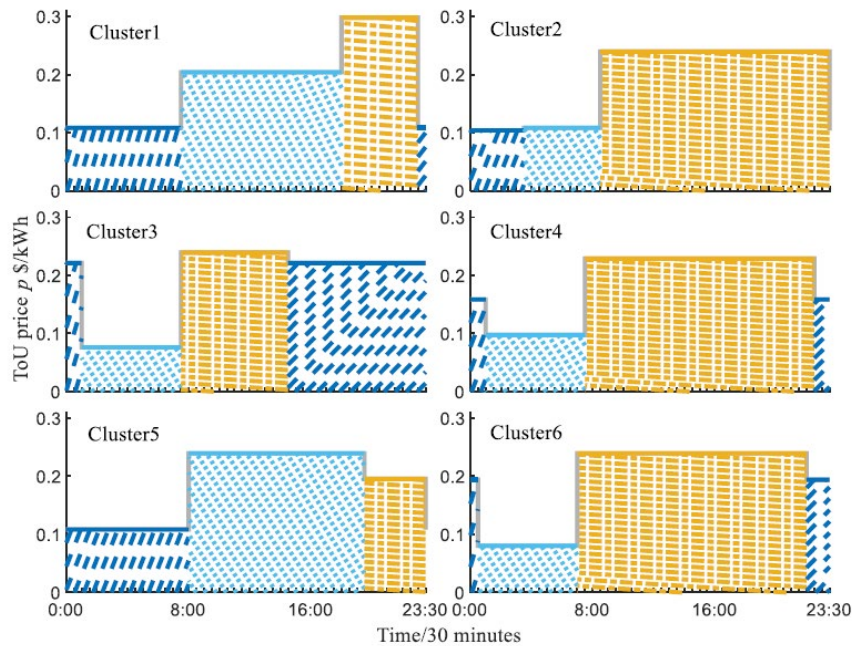
$$|e_{r,T}^m - e_{r,1}^m| + \sum_{t=2}^T |e_{r,t-1}^m - e_{r,t}^m| = 2, \quad \forall m, r$$

$$p_{r,t} = \sum_{m=1}^M e_{r,t}^m \times p_r^m, \quad \forall t, r \quad m \text{ block ToU}$$

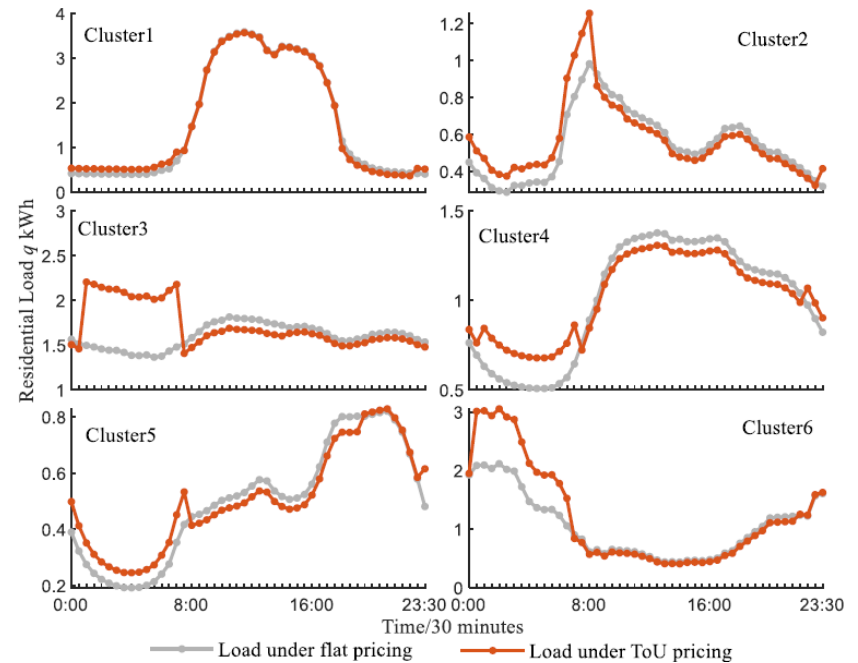
* 标红的为非线性项

2. 数据决策——如何将要素转化为生产力？

◆ 个性化电价设计



设计的个性化定价



用户的响应行为

2. 数据决策——如何将要素转化为生产力?

◆ 个性化电价设计

➤ 不同用户划分方法的结果对比

	RP	SW	AP	F/SC
Original	752.03	0	0.2	-/-
HIA-COMP	1186.01	339.72	0.1947	65%/89%
HIA-WARD	1188.70	10.01	0.1971	33%/59%
KM-PLUS	1145.68	7.01	0.1973	9%/20%
KM-SAMPLE	1137.61	4.50	0.1975	22%/48%
KM-UNIFORM	1142.61	15.76	0.1973	11%/31%
FCM(m=1.1)	1150.43	9.43	0.1970	30%/47%
FCM(m=1.2)	1176.08	18.64	0.1968	19%/35%
FCM(m=1.3)	1208.06	0.64	0.1970	8%/20%
GMEM-PLUS	1145.82	36.01	0.1965	13%/28%
GMEM-RAND	1144.85	46.60	0.1967	10%/24%

电力零售商能够获取多少利润?

- **RP**=Retailing Profit(\$)

用户能够得到多大的社会福利?

- **SW**=Social Welfare

- **AP**=Average Price(\$/kWh)

不同用户划分方法效果如何?

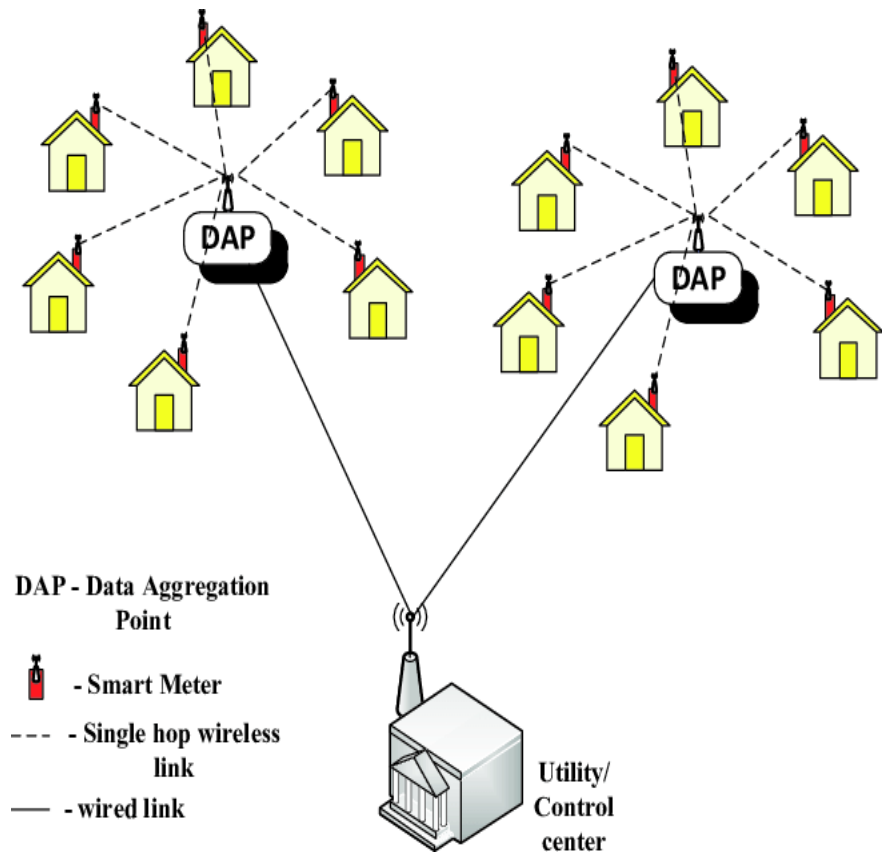
- **F/SC**=First/Second Choice

➤ 对用户行为的建模越准确，电力零售商的利润就越大、用户的平均用电成本也越低，实现共赢。

3. 数据采集——如何实现要素的高效获取？

◆ 数据聚合点 (Data Aggregation Point) 优化配置

- 工作目标：优化从智能电表到数据中心的通信网络，以实现经济高效且可靠的数据传输。
- 数据聚合点是关键组件，扮演桥梁角色。
- 最佳的DAP放置至关重要，以确保使用最少数量的DAP进行高质量的通信。
- 传统通信模式是单跳（single-hop）模式，而多跳（multi-hop）模式可能让整个无线通信网络更经济更鲁棒。



3. 数据采集——如何实现要素的高效获取？

◆ 数据聚合点优化配置

- 目标函数：安装成本+传输成本+时延成本(1)

$$\min_{x_j, y_{ij}^{sp}, z^{sp}} c_{inst} + c_{trans} + c_{dly} \quad (1)$$

$$\text{s.t.} \quad z^{sp} \leq x_p \quad \forall s, p \quad (2)$$

- 数据聚合点需安装在指定的地方(2)

$$\sum_{p=1}^{N_{poles}} z^{sp} = 1, \quad \forall s \quad (3)$$

$$y_{ij}^{sp} \leq z^{sp} \quad \forall s, p, i, j \quad (4)$$

- 智能电表的最大传输距离(8)

$$\sum_{i=1}^{N_{SM}} y_{ip}^{sp} = z^{sp} \quad \forall s, p, \quad (5)$$

- 每个数据聚合点连接智能电表数量限制(9)

$$\sum_{i=1}^{N_{poles}+N_{SM}} y_{si}^{sp} = z^{sp} \quad \forall s, p \quad (6)$$

$$\sum_{j=1}^{N_{poles}+N_{SM}} y_{ij}^{sp} = \sum_{j=1}^{N_{SM}} y_{j(i+N_{poles})}^{sp}, \quad \forall s, \forall p, \forall i \text{ with } i \neq s \quad (7)$$

$$d_{ij} y_{ij}^{sp} \leq d_{max}, \quad \forall s, p, i, j \quad (8)$$

- 每个智能电表连接其他智能电表数量限制(10)

$$\sum_{s=1}^{N_{SM}} z^{sp} \leq A_1, \quad \forall p \quad (9)$$

$$\sum_{s=1}^{N_{SM}} \sum_{p=1}^{N_{poles}} \sum_{i=1}^{N_{SM}} y_{ij}^{sp} \leq A_2, \quad \forall j = N_{poles} + 1, \dots, N_{SM} \quad (10)$$

- 网络信息流相关约束(3), (5) (6), (7)

$$x_j \in \{0, 1\} \quad \forall j \quad (11)$$

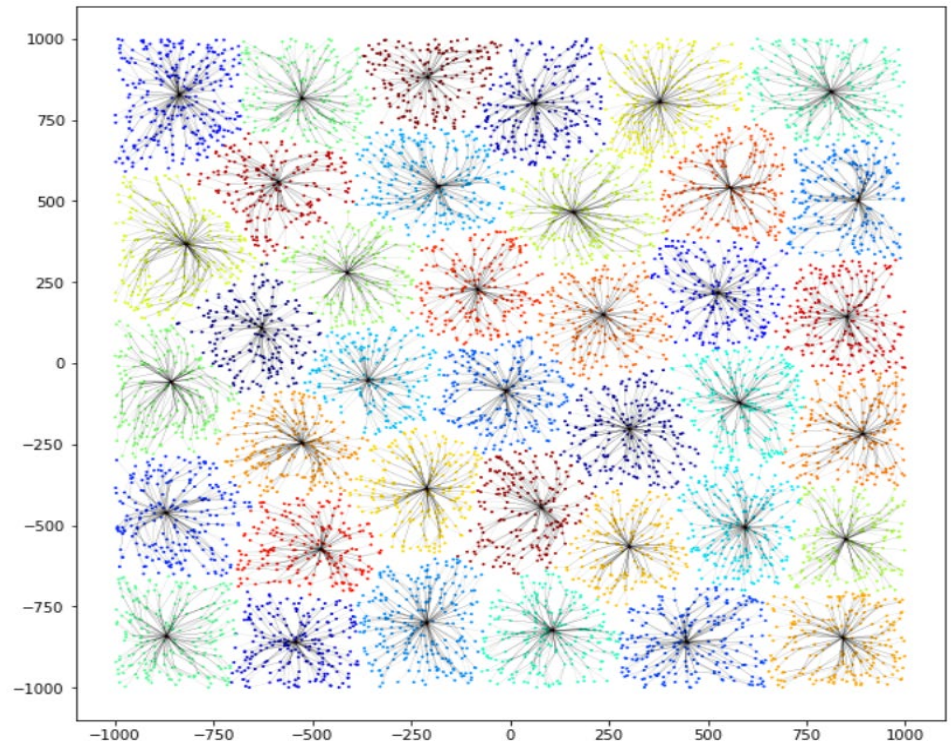
$$y_{ij}^{sp} \in \{0, 1\} \quad \forall s, p, i, j \quad (12)$$

$$z^{sp} \in \{0, 1\} \quad \forall s, p \quad (13)$$

3. 数据采集——如何实现要素的高效获取？

◆ 数据聚合点优化配置

- 算例设置
- 在一个4km²区域，8000个智能电表、2500个杆塔待选
- 最大传输距离: 100m
- 每个数据聚合点最多连接2000个智能电表
- 每个智能电表最多连接3个其他智能电表
- 最终一共配置38个数据聚合点

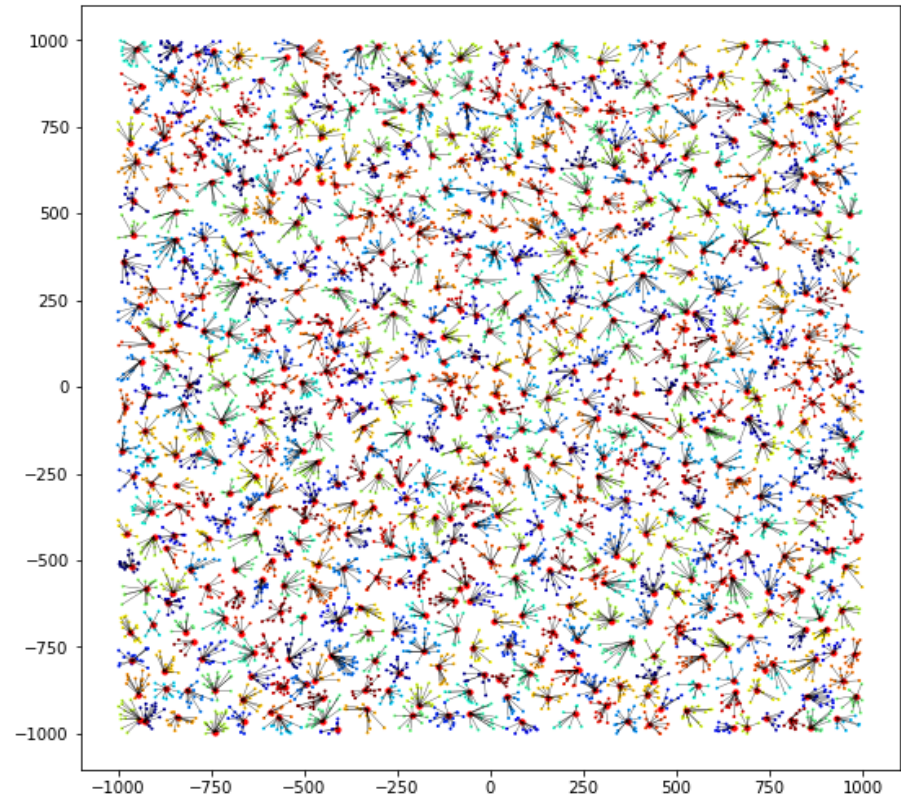


```
number of poles selected 38
installation cost: 76000
transmission cost: 19927.397125449977
delay cost: 10393.0
total cost 106320.39712544998
max number of SM per DAP 272
min number of SM per DAP 181
average number of SM per DAP 210.52631578947367
direct connections to DAP 2090
```

3. 数据采集——如何实现要素的高效获取？

◆ 数据聚合点优化配置

- 与单跳模型对比
- 一共配置599个数据聚合点
- 单跳总成本\$1'200'706（时延成本较低）
- 多跳总成本\$106'320（没有考虑智能电表升级带来的成本）
- 智能电表传输距离的限制对优化结果影响最大

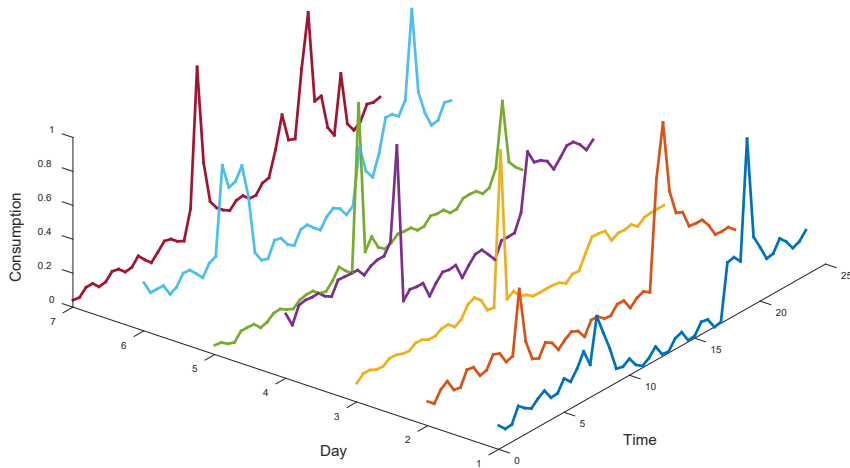


```
installation cost 1198000
transmission cost 2706.81435304163
total cost 1200706.8143530416
number of poles selected 599
max number of SM per DAP 27
min number of SM per DAP 5
average number of SM per DAP 13.35559265442404
```

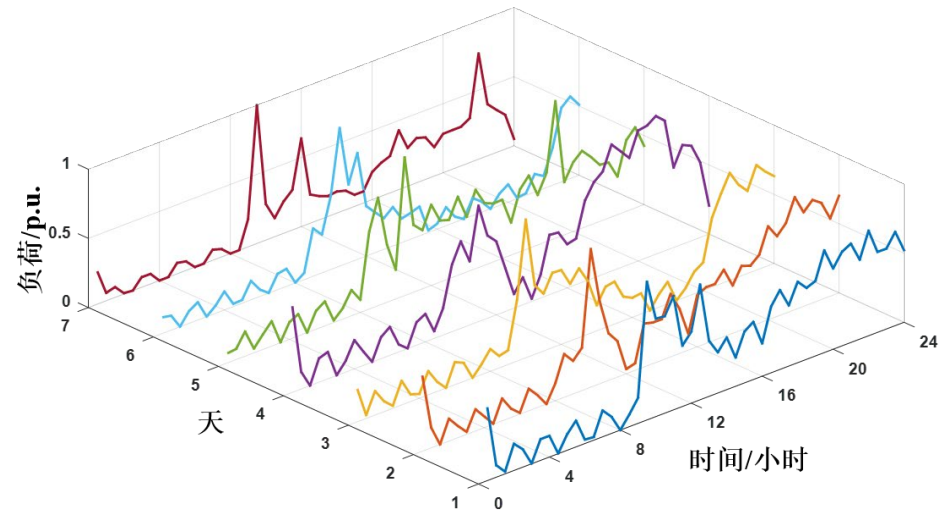

4. 数据存储——如何经济高效地积累要素？

◆ 智能用电数据压缩感知

用户用电曲线的初步认知：



某居民用户一周负荷曲线



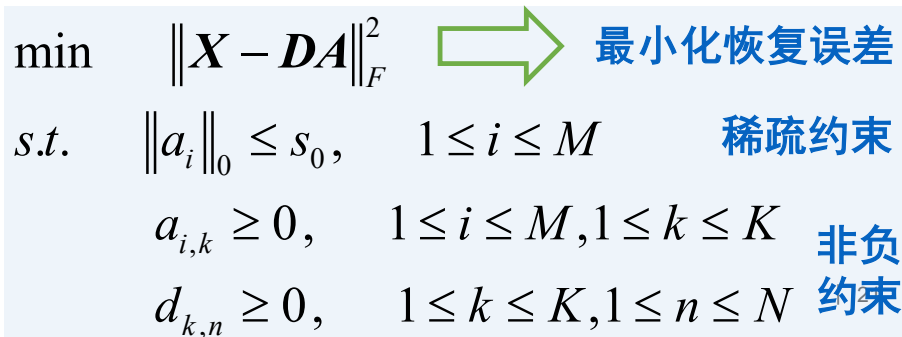
某工商业用户一周负荷曲线

用户用电特性

- 1) **稀疏性**：很多时段负荷值均接近于基荷，能“大致”看出一些固定的用电模式；
- 2) **多样性**：不同用户用电曲线差异大，甚至同一用户不同天也是多变的。

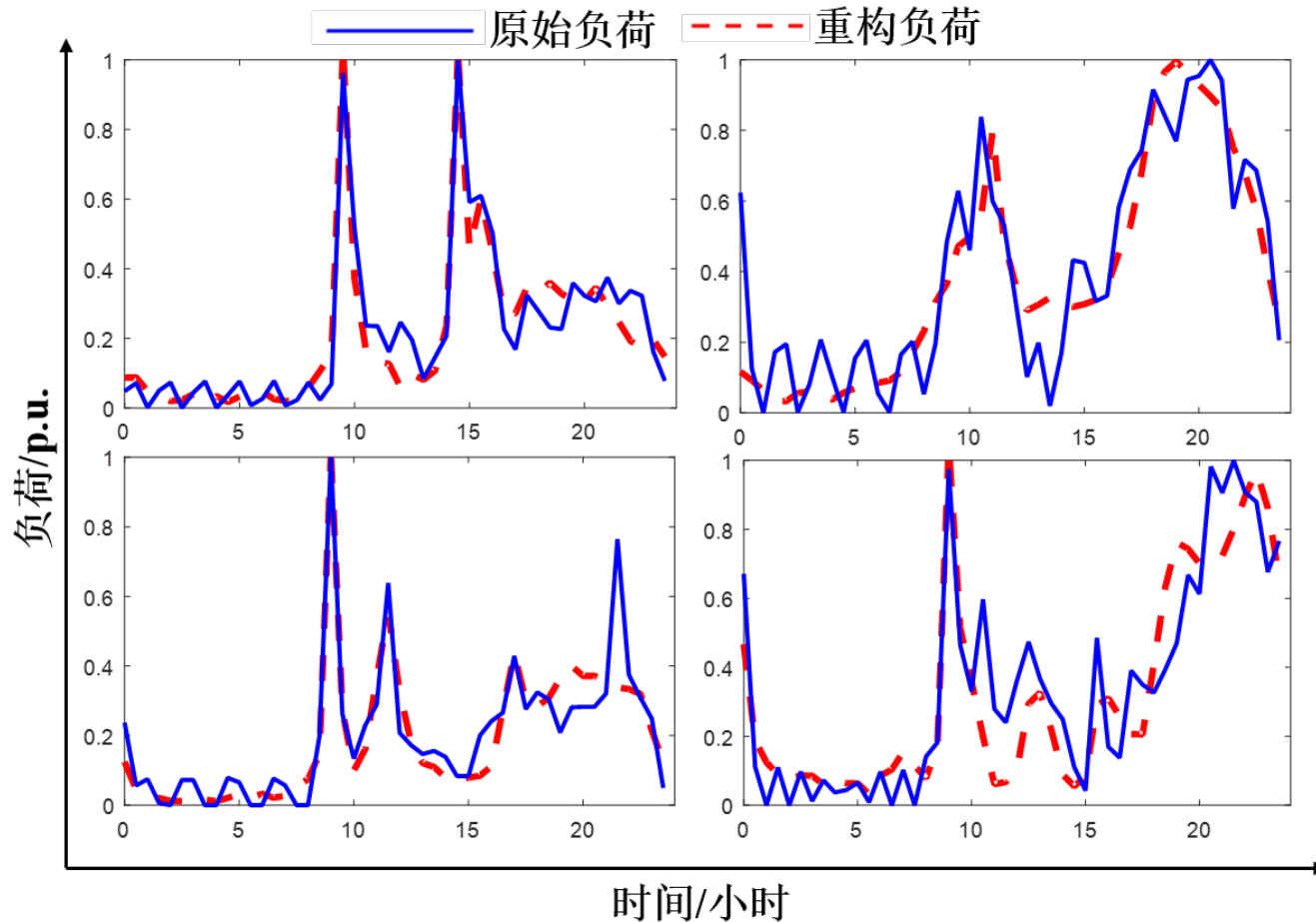
将用户用电曲线看作若干用电“子模式”的线性叠加，采用稀疏编码技术对用电曲线进行分解。

$$x_i = \sum_{k=1}^K a_{ik} d_k$$



4. 数据存储——如何经济高效地积累要素？

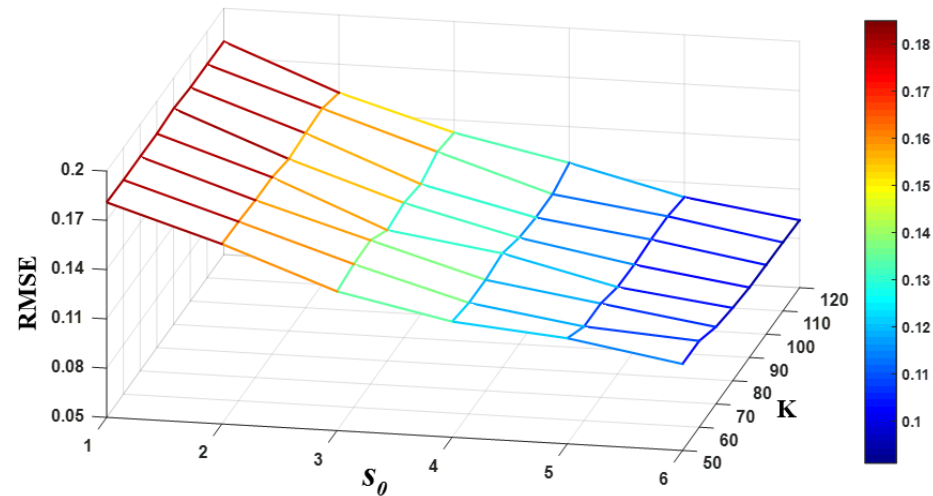
◆ 智能用电数据压缩感知



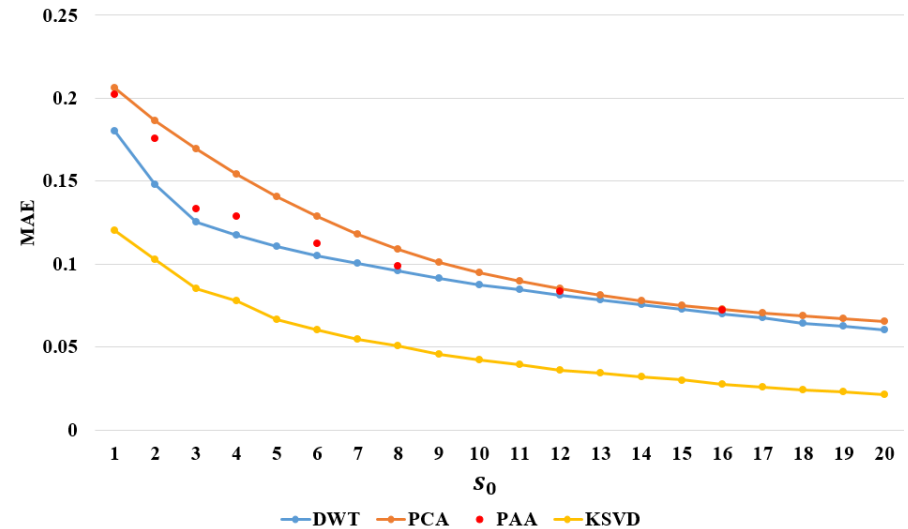
用电曲线恢复效果：仅用3-4个用电模式就能够很好地重构负荷曲线；
意味着：用户用电曲线本质上是几种简单用电行为的叠加！

4. 数据存储——如何经济高效地积累要素？

◆ 智能用电数据压缩感知



重构误差随参数变化情况

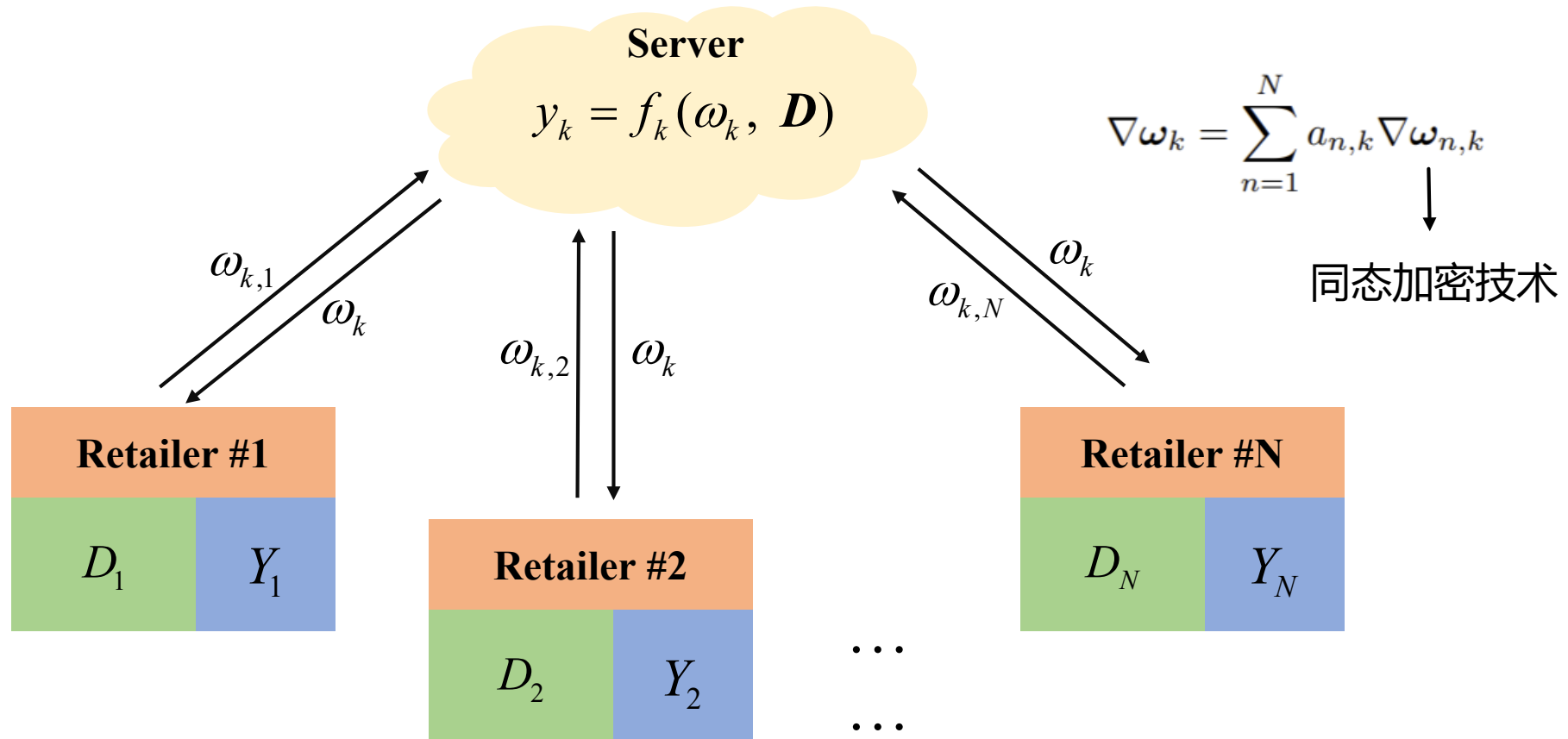


各方法重构误差变化情况

5. 数据共享——如何实现要素整合与交易?

◆ 基于联邦学习的肖像描绘

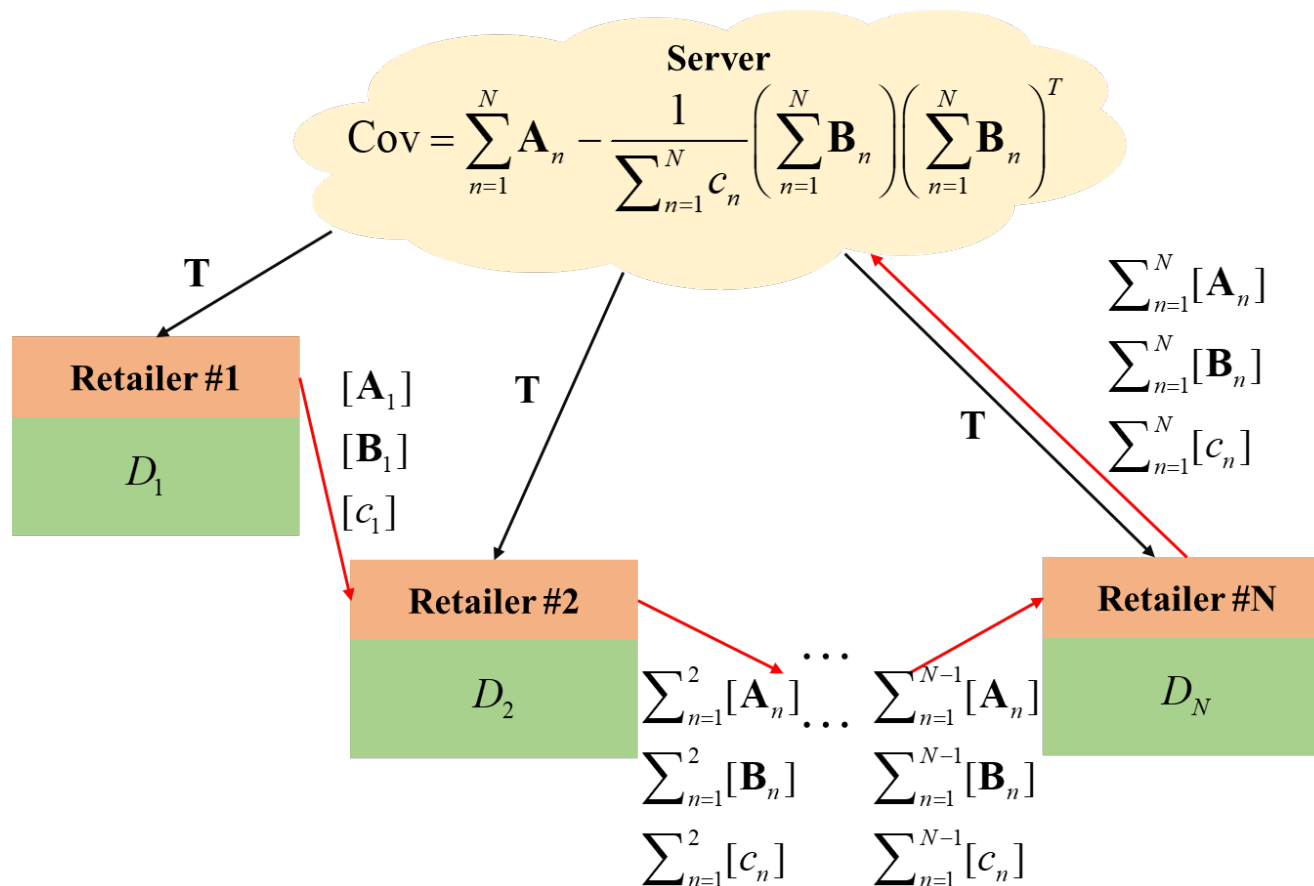
如何用于肖像描绘的数据隶属于不同的售电商，他们如何共享数据，且保护隐私？



5. 数据共享——如何实现要素整合与交易?

◆ 基于联邦学习的肖像描绘

基于联邦学习的PCA特征提取:



5. 数据共享——如何实现要素整合与交易？

◆ 基于联邦学习的肖像描绘

基于Shapely Value的利益分配与数据交易？

三种不同的联邦学习策略

Characteristics	Accuracy				
	Centralized	C-PCA	LA	LS	Normal
#1 Employment status	0.7190	0.7332	0.7367	0.7432	0.7420
#2 Retirement status	0.7591	0.7698	0.7577	0.7594	0.7651
#3 Have children	0.7686	0.7769	0.7727	0.7739	0.7739
#4 Live alone	0.8442	0.8288	0.8412	0.8453	0.8465
#5 Number of residents	0.7485	0.7580	0.7503	0.7500	0.7527
#6 Number of bedroom	0.5302	0.5882	0.5911	0.5920	0.5941
#7 Age	0.6770	0.6805	0.6728	0.6728	0.6743
#8 Number home appliances	0.6423	0.6572	0.6610	0.6665	0.6706
#9 Light bulb proportion	0.6021	0.5762	0.6308	0.6322	0.6284
#10 Type of home	0.6017	0.6173	0.6191	0.6191	0.6176
#11 Social class	0.5583	0.5656	0.5644	0.5660	0.5647
#12 Number of entertainment appliances	0.5396	0.5714	0.5517	0.5466	0.5487
#13 Cooking	0.7155	0.7332	0.7267	0.7264	0.7311
#14 Age of house	0.4782	0.4947	0.5068	0.5162	0.5121
#15 Rent or owned	0.9265	0.9289	0.9295	0.9301	0.9298
Average	0.6740	0.6853	0.6875	0.6893	0.6901

联邦学习与集中式学习的效果相当！

小结

- 用户侧管理，在数字化背景下大有可为！
- 能否用**数据思维**重新审视传统问题？
- 数据作为一种要素，能否实现**数据供应链管理**？
- 数据（信息）与能量的相互支撑关系如何？

做个广告.....

Yi Wang · Qixin Chen ·
Chongqing Kang

Smart Meter Data Analytics

Electricity Consumer Behavior
Modeling, Aggregation,
and Forecasting

Science Press
Beijing

Springer

Foreword

viii

Foreword

alone. There are definitely broader qualitative understanding that can be gained from massive data collected in the realm of generation, transmission, distribution, and end use of the smart grid.

September 2019

Prof. Saifur Rahman
Joseph Loring Professor and Founding Director
Advanced Research Institute at Virginia Tech
Arlington, VA, USA

President of the IEEE Power and Energy Society
New York, NY, USA

Smart grid is a cyber-physical-social system where the power flow, data flow, and business flow are deeply coupled. Enlightened consumers facilitated by smart meters form the foundation of a smart grid. Countries around the world are in the midst of massive smart meter installations for consumers on the pathway towards grid digitalization and modernization. It enables the collection of extensive fine-grained smart meter data, which could be processed by data analytical techniques, especially now widely available machine learning techniques. Big data and machine learning terms are widely used nowadays. People from different industries try to apply advanced machine learning techniques to solve their own practical issues. The power and energy industry is no exception. Smart meter data analytics can be conducted to fully explore the value behind these data to improve the understanding of consumer behavior and enhance electric services such as demand response and energy management.

This book explores and discusses the applications of data analytical techniques to smart meter data. The contents of the book are divided into three parts. The first part (Chaps. 1–2) provides a comprehensive review of recent developments of smart meter data analytics and proposes the concept of “electricity consumer behavior model”. The second part (Chaps. 3–5) studies the data analytical techniques for smart meter data management, such as data compression, bad data detection, data generation, etc. The third part (Chaps. 6–12) conducts application-oriented research to depict electricity consumer behavior model. This part includes electrical consumption pattern recognition, personalized tariff design for retailers, socio-demographic information identification, consumer aggregation, electrical load forecasting, etc. The prospects of future smart meter data analytics (Chap. 13) are also provided at the end of the book. The authors offer model formulations, novel algorithms, in-depth discussions, and detailed case studies in various chapters of this book.

One author of this book, Prof. Chongqing Kang, is a professional colleague. He is a distinguished scholar and pioneer in the power and energy area. He has done extensive work in the field of data analytics and load forecasting. This is a book worth reading; one will see how much insight can be gained from smart meter data

This is a book worth reading; one will see how much insight can be gained from smart meter data alone.



Prof. Saifur Rahman
IEEE Fellow
President of the IEEE Power and Energy Society

Yi Wang, Qixin Chen, Chongqing Kang, “Smart Meter Data Analytics,” Springer, 2020.

谢谢!

Yi Wang | yiwan@eeh.ee.ethz.ch | www.eeyiwan.com

