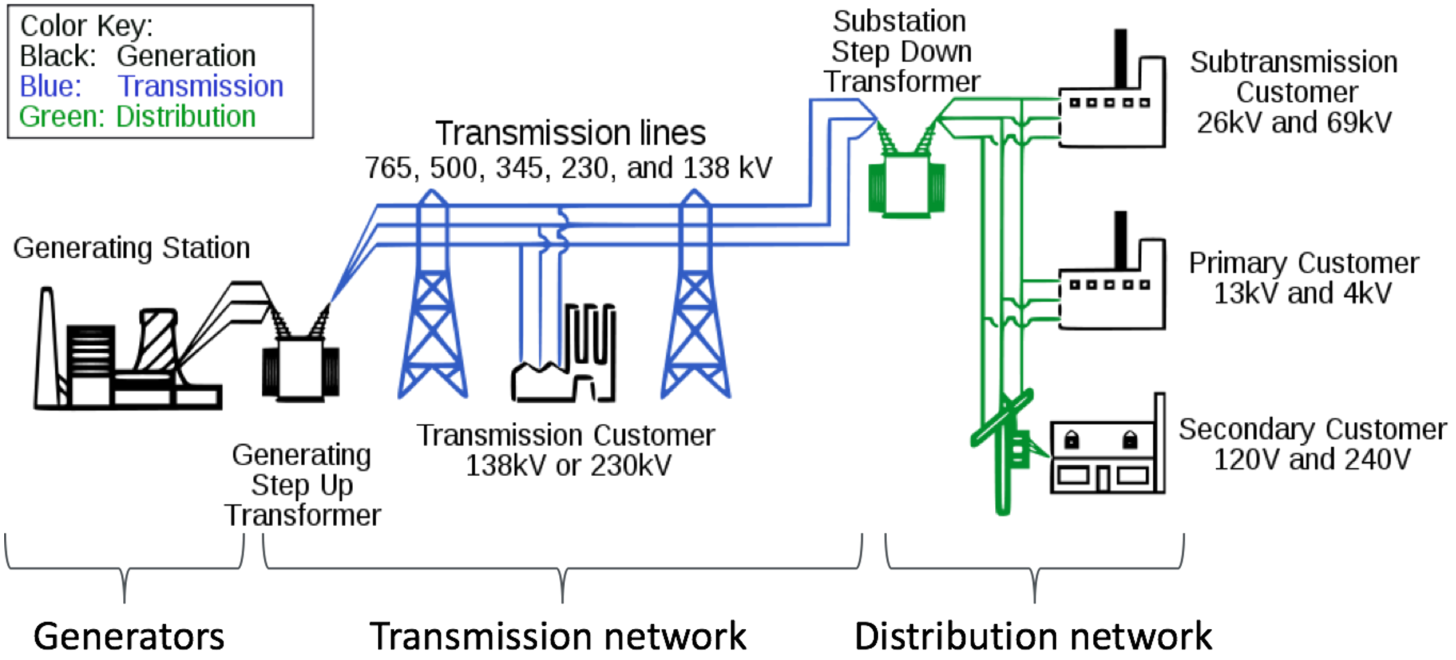
# Human-centric demand side management in electricity:

## lifestyles, privacy, and fairness

Xiao (Mark) Chen, CEE
advised by Prof. Ram Rajagopal

# Outline

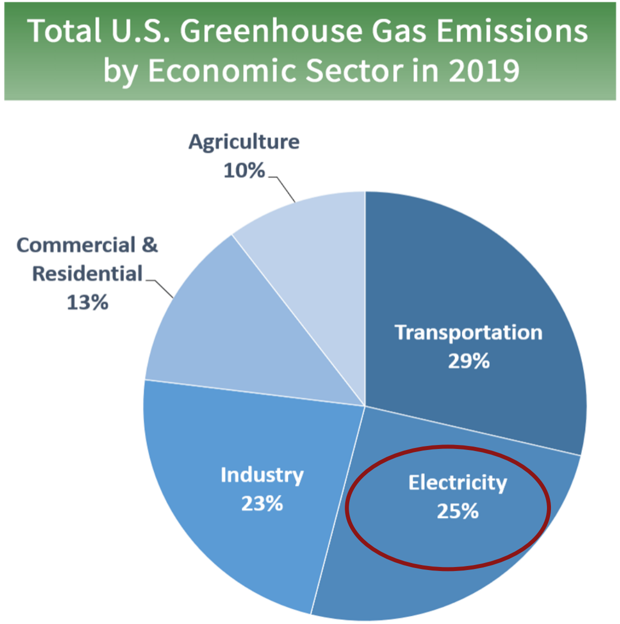
# Traditional electricity grid



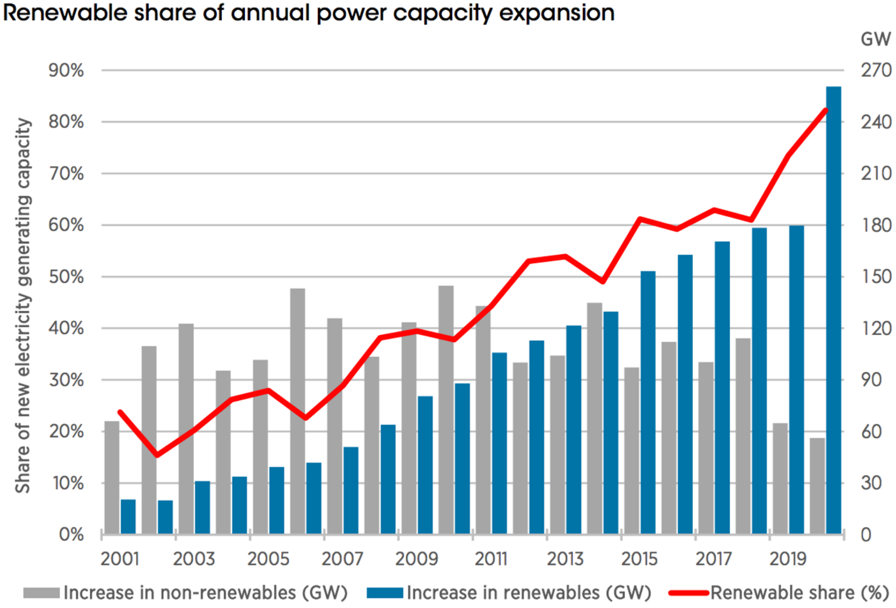

[Source: Wikipedia, https://en.wikipedia.org/wiki/Electrical_grid]

# Issues in a traditional grid

- Carbon emissions




- Renewables increase 261GW (+10.3%) in 2020
- 91% shares are wind + solar

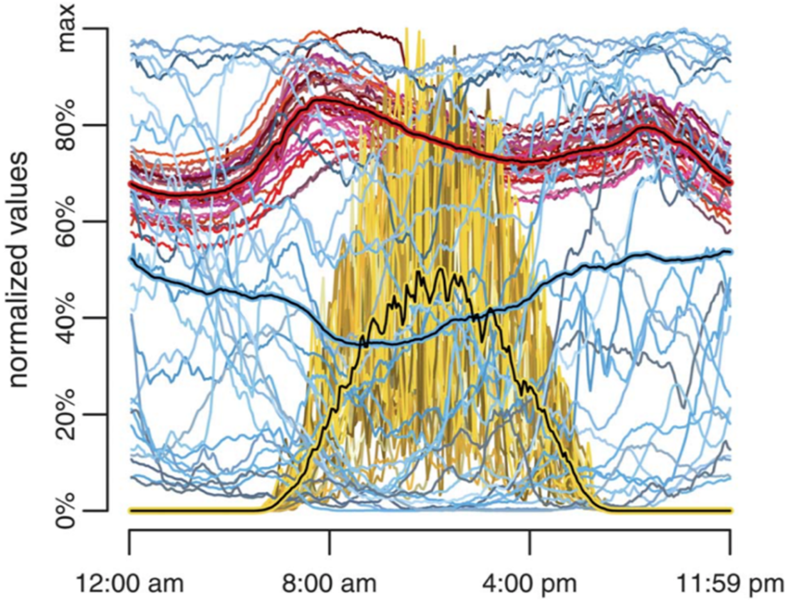



[Source: epa, sources-greenhouse-gas-emissions; renewable capacity statistics, https://www.irena.org]

# Challenges: demand correlates poorly with renewables

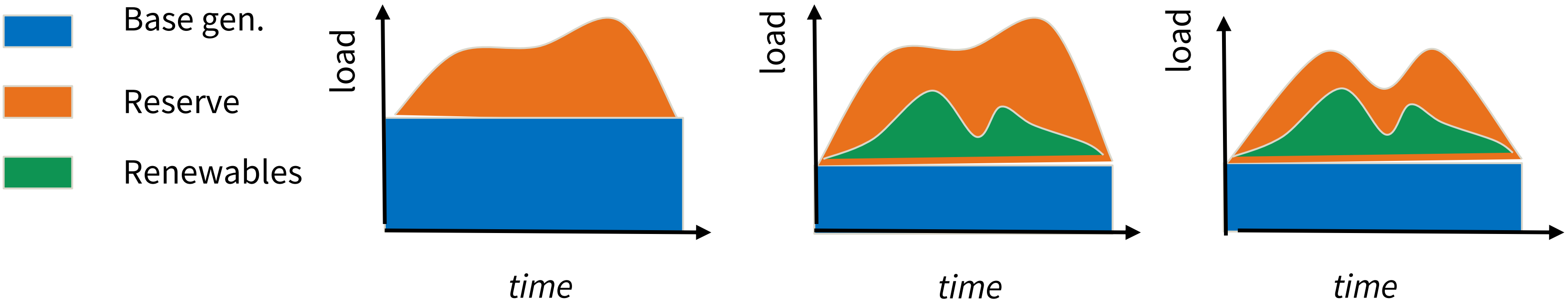
30 days of data from Bonneville Power Administration (OR, WA, ID)

- Red: demand
- Yellow: solar
- Blue: wind




[Charles J. Barnhart, Michael Dale, Adam Brandt, Sally M. Benson, "The energetic implications of curtailing versus storing solar- and wind-generated electricity," Energ. Environ. Sci. 2013.]

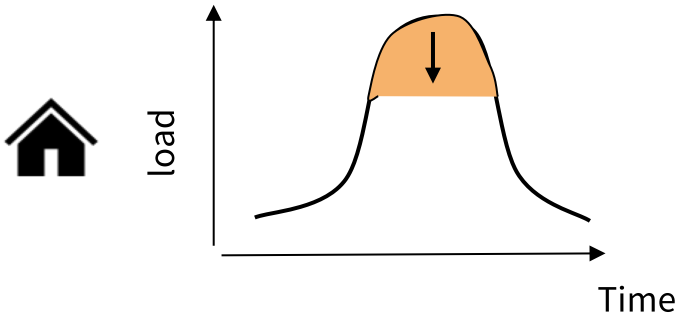
# Demand Side Management (DSM) addresses the challenges

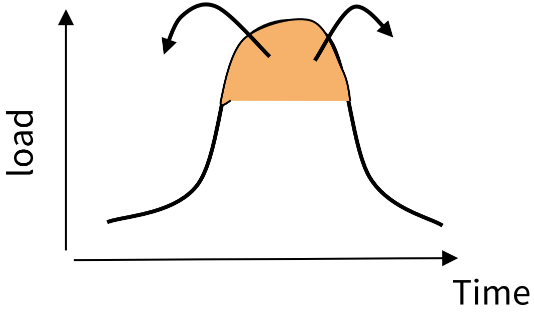



- Renewables replace some base generation but increase the (expensive) reserve
- Demand side management can reduce the reserves, aligning with decarbonization

[Figure credit: Stanford class EE100]

# The concept of DSM




Reducing the electricity consumption (at peak hours)

Shifting the electricity demand (from peak to off-peak hours)

# Example of DSM: SmartHours in OG&E

Time-of-Use(TOU) Price: Peak price during 2pm-7pm weekdays


Off-peak: 5 ¢

Peak hours:
Standard: 10 ¢
High: 27 ¢

Critical-peak: 47 ¢

OG&E®

Enrolled over 70k customers

Effects for a customer:

- $191 bill savings annually

Effects for the utility:

- reduced 70 MWh load
- deferred two peak plants (170 MW) ~$363 million

# Issues of DSM planning and deployment

Identifying right user to participate in the program

Mass promotion to target users uniformly

Full knowledge of (controllable) loads

Exposure to the private information of households

The heterogeneity of the users is viewed by the volume of consumption only.

Social demographics, vulnerable groups

# Outline for research work presented today

Understand user behavior

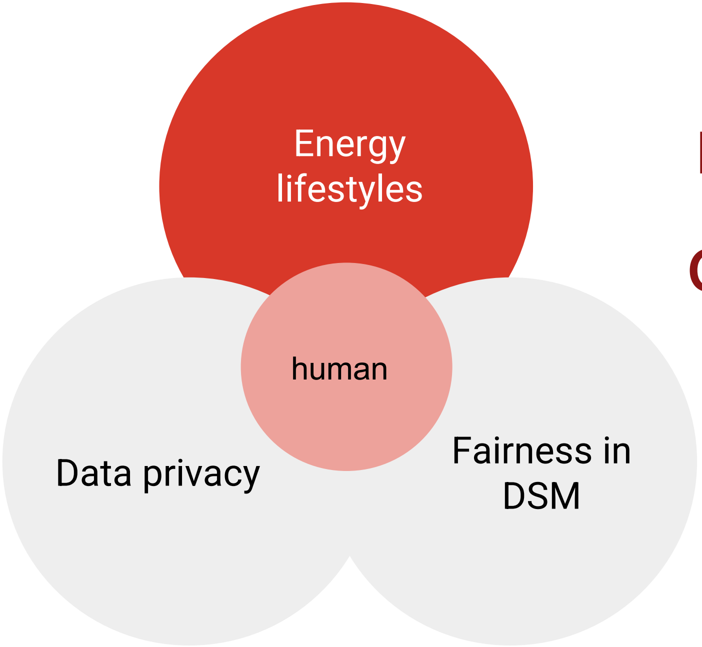
**Construct energy lifestyles**

Protect user privacy

**Generate privacy preserved data**

Ensure fairness among users

**Fairness-aware demand response**




# Part I

## Constructing Energy Lifestyles

using Latent Dirichlet Allocation

# Energy lifestyles: background


Smart meter installation:

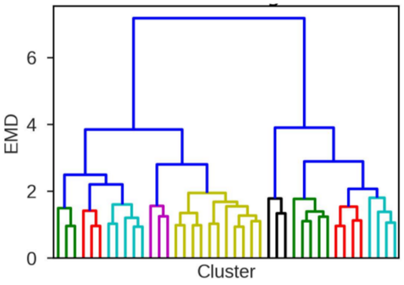
- 102 million in 2020 (in US); 88% in residential
- 100 thousand meters leads to 17.5 TB/yr;

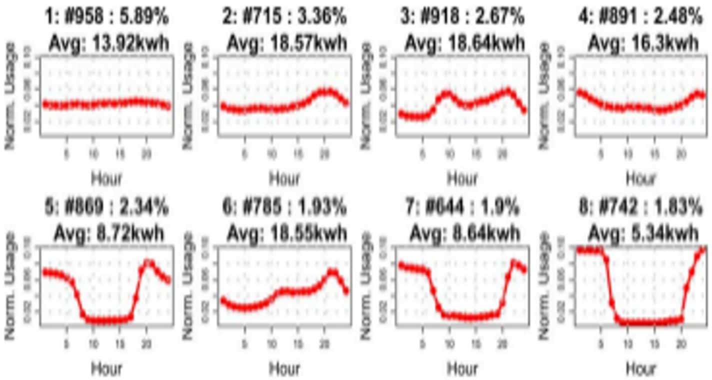
Data support planning of DSM:

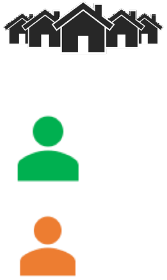
- Pricing (Time-of-use; Critical-peak pricing, etc.)
- Distributed generation (solar; battery storage, etc)

[Cooper & Shuster, Electric Company Smart Meter Deployments: Foundation for a Smart Grid (2021); EIA 2020;]

# Current gaps in characterizing energy use patterns



Methods

- K-means, K-median
- Hierarchical clustering
- DBSCAN

Create loadshape dictionary:

- Representative of daily shapes

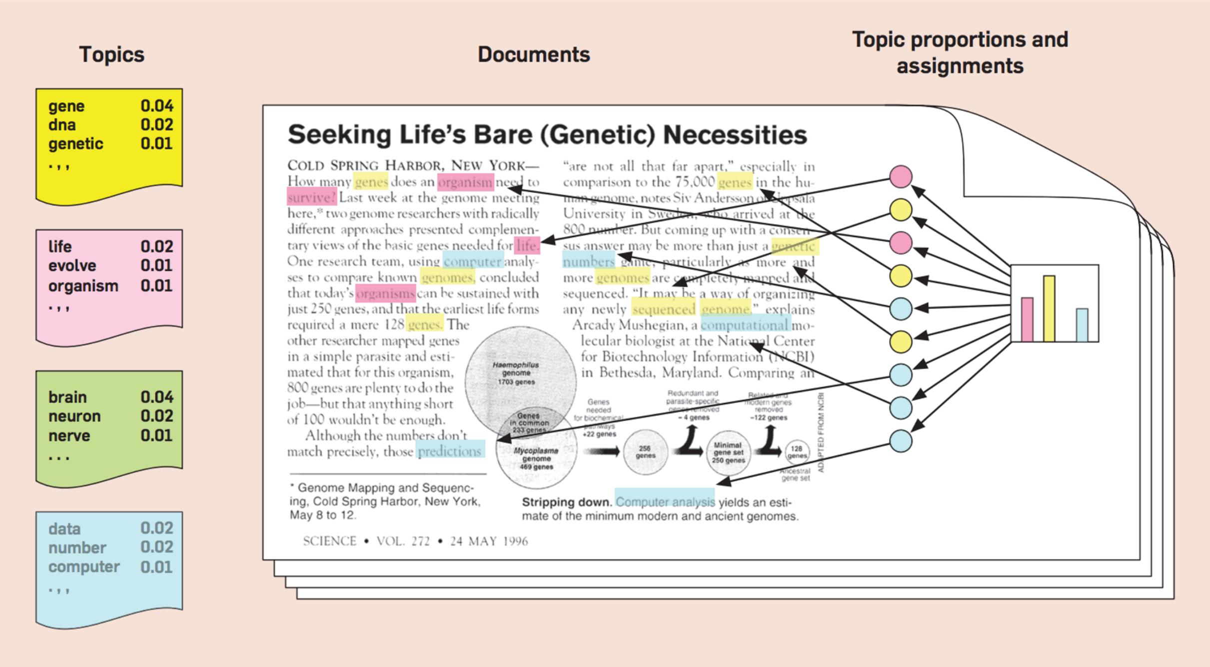
From dictionary to comprehensive views (dynamically)?

[Albert & Rajagopal 2013; Kwac et al. 2014; Barbour & Gonzalez 2018; Haben et al. 2016; Kwac et al. 2016; Teichgraeber & Brandt 2018;]

# Example: topic modeling

Word dictionary:

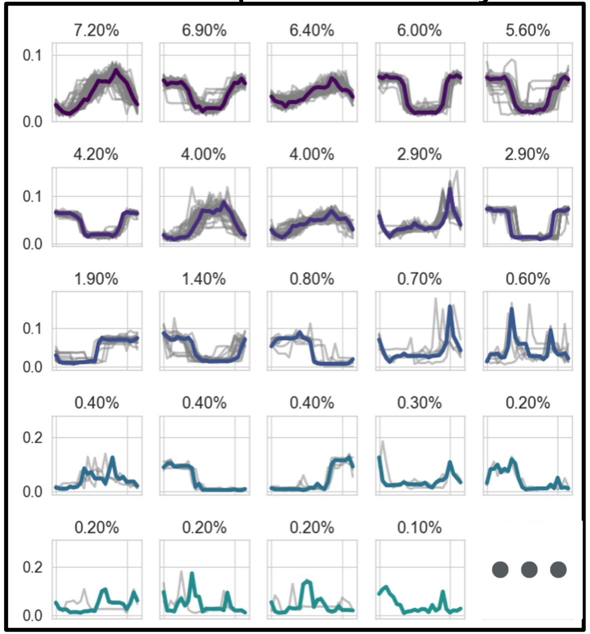
a
abandon
ability
able
about
above
abroad
absolute
...



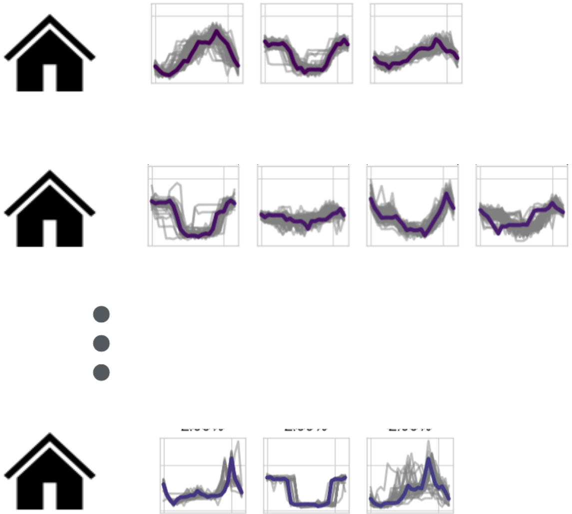

[Figure credit: Blei, David M. "Probabilistic topic models." Communications of the ACM 55.4 (2012): 77-84.]

# Motivation of using Latent Dirichlet Allocation

Loadshape dictionary:




Households:

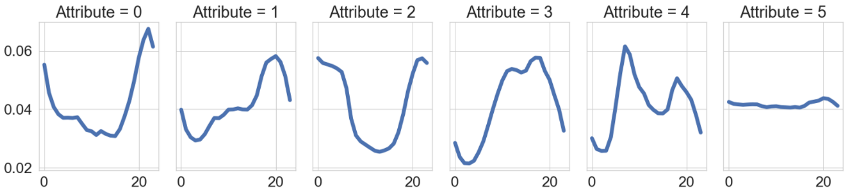


Consumption Topics?

# What are energy lifestyles?

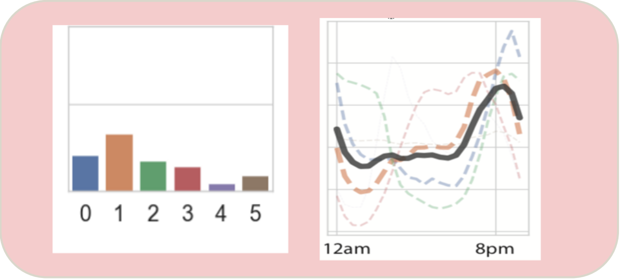
Household

Energy attribute (or topic)




*T (24 hr)*

Energy lifestyle



# Energy lifestyles: Research questions

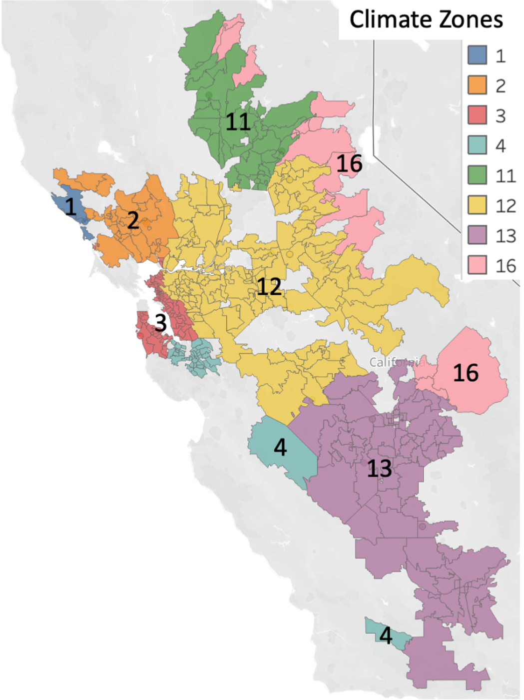
**How can we gain behavioral insights about energy use for residential customers?**

1. How does energy use relate to (daily) lifestyles? (constructing lifestyles)
2. What are the patterns of these lifestyles across time, in seasons?

# Energy lifestyles: PG&E Dataset

PG&E residential meter data:

- hourly time series for 1 year (2010 - 2011)
- 60k households
- 436 zip codes
- Covering the 7 climate zones

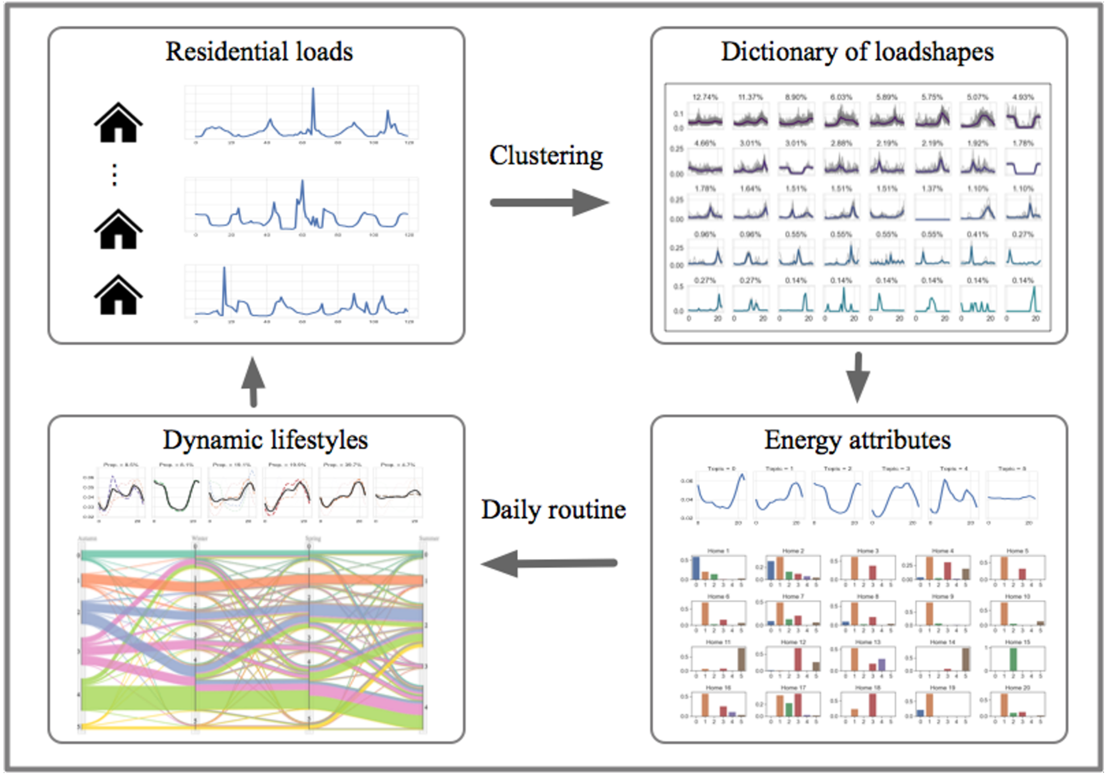


# The workflow of constructing energy lifestyles



Step 1

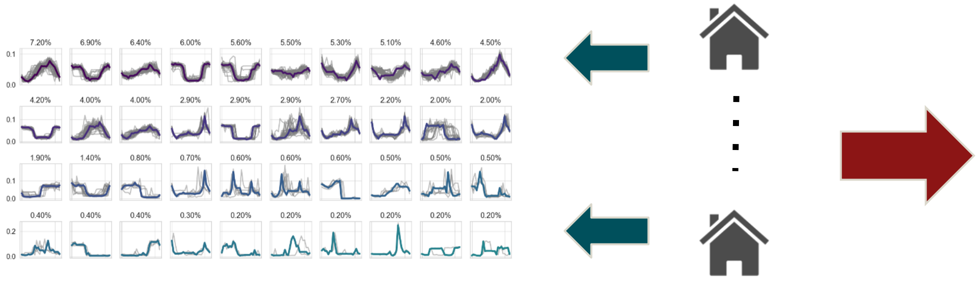
Step 2

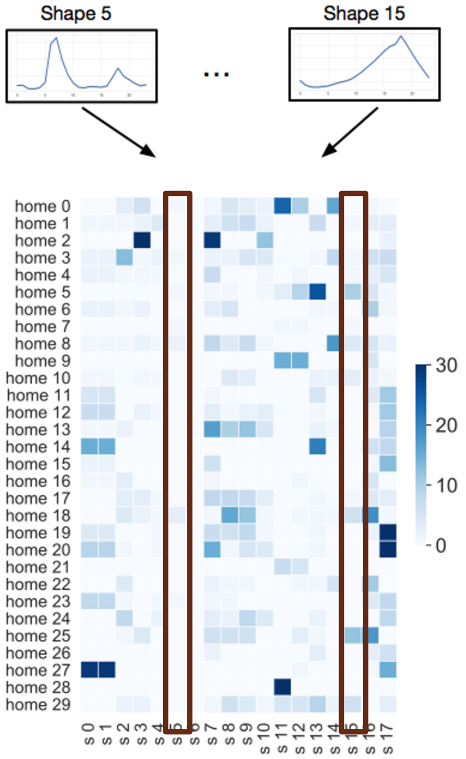



Step 4

Step 3

# Frequency count of load shapes



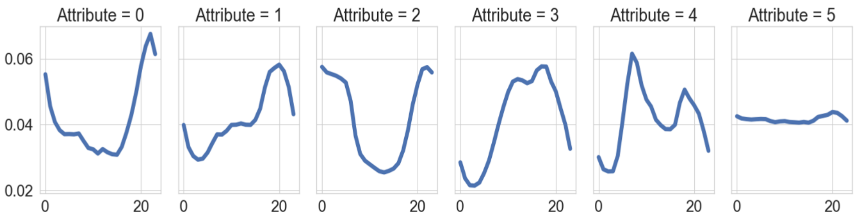

Intuitions:

- Consider cluster centers as representative shapes
- Construct a frequency count table (similar to word frequency counts) via Nearest Neighbor

# Generating energy attributes and lifestyles

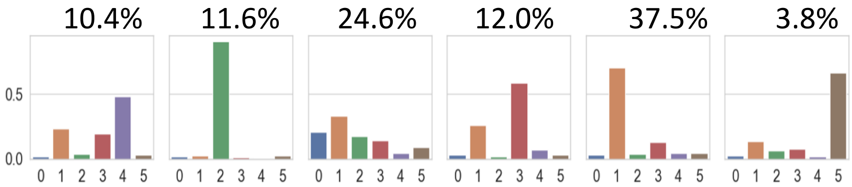
We apply LDA to uncover six latent energy attributes (topics)

Late night peak | evening peak | midnight peak | afternoon peak | dual peak | flat

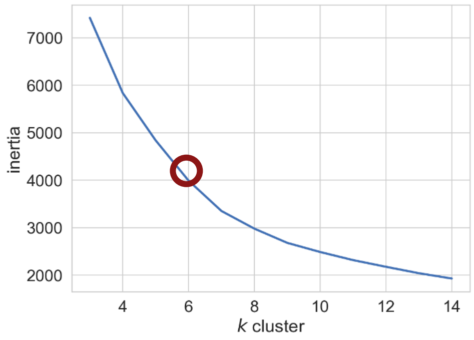


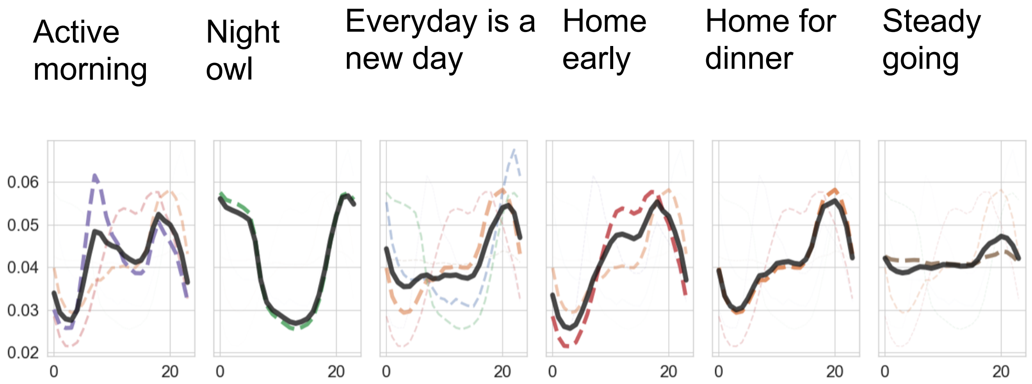

We apply k-means to cluster the homes (6-dim vectors) to:

- create labels of lifestyles
- reveal transition dynamics

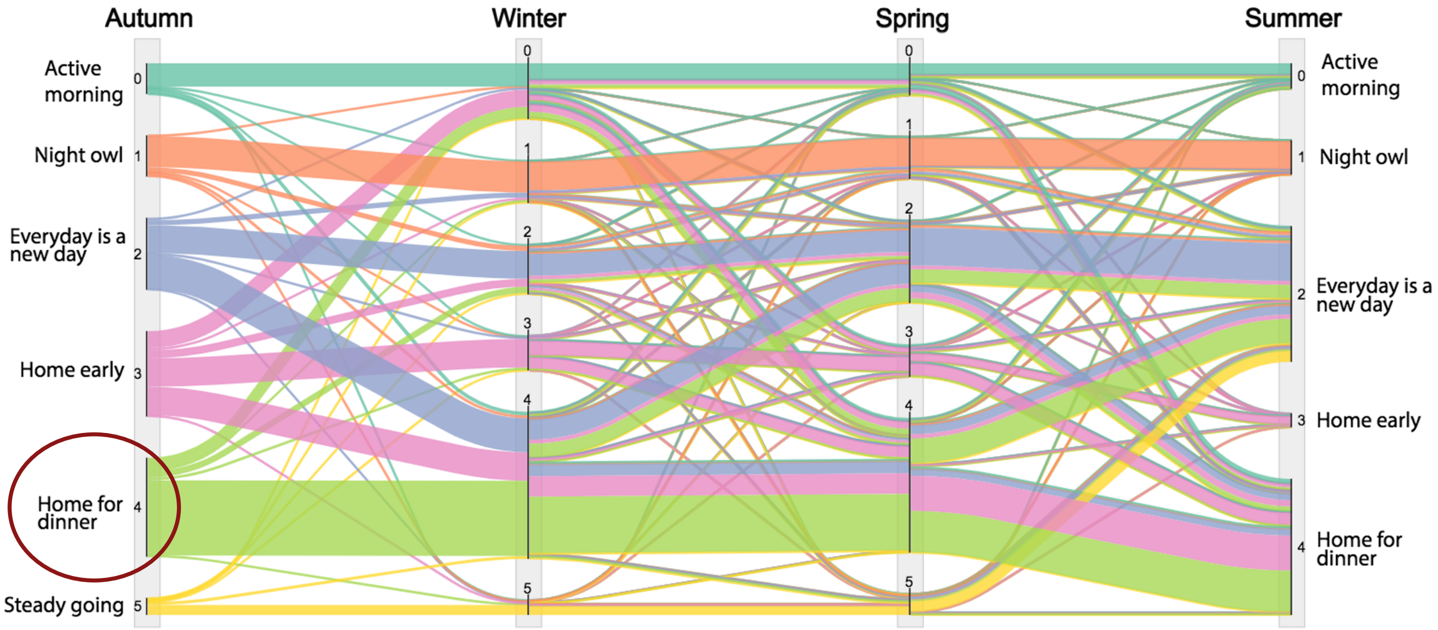


# Energy lifestyles: user behavior of daily usage



# Energy lifestyle transitions across seasons

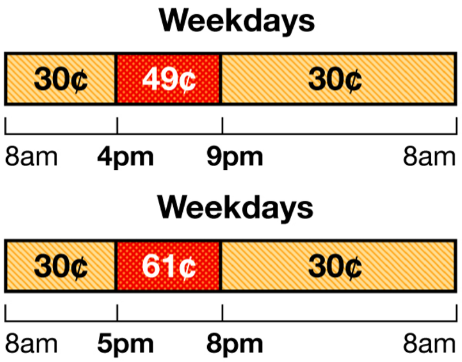


# Application of lifestyles: DSM targeting

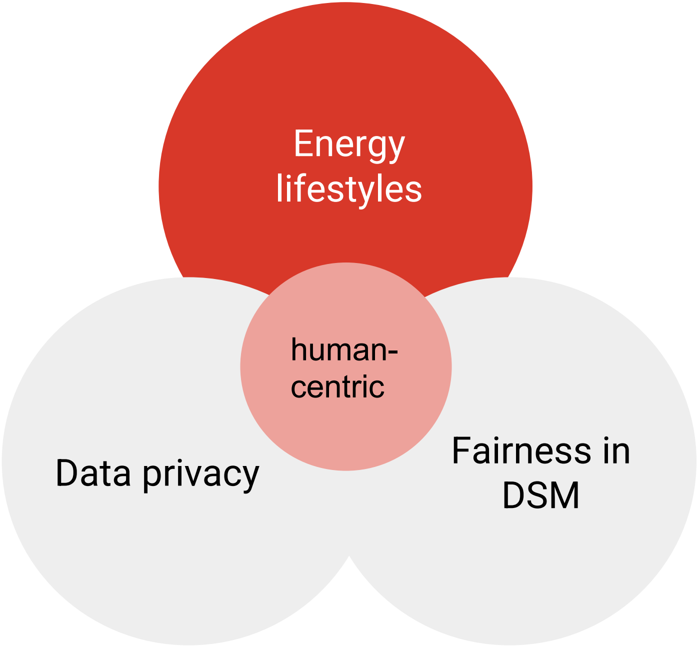
Few potential strategies:

- **Home for dinner** for demand response programs or critical peak pricing.
- **Home for dinner** for distributed energy devices, e.g., solar and/or storage.
- **Everyday is a new day** for smart AC, energy efficiency program

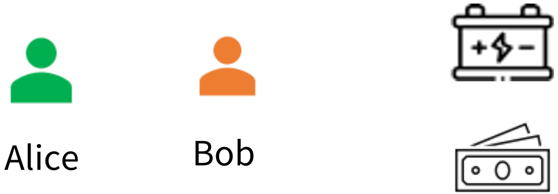
Example:

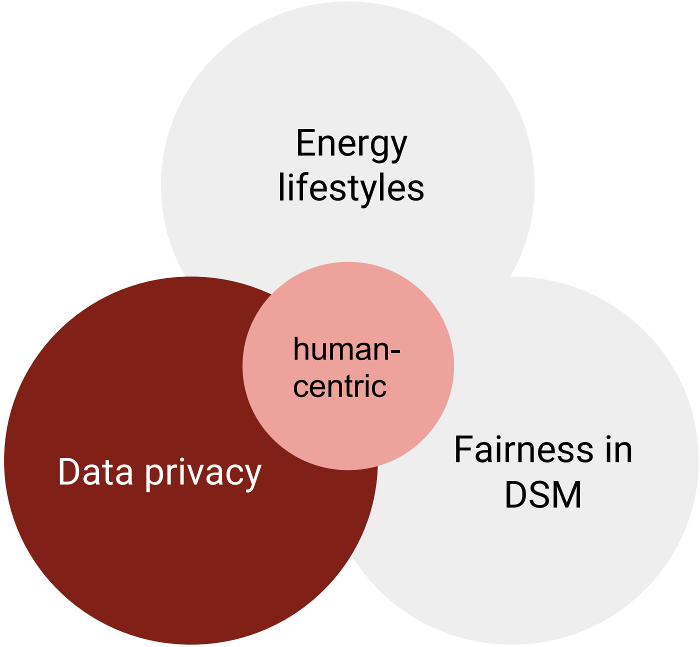


# Conclusions about lifestyles




- We discovered six concise lifestyles from the PG&E dataset with a novel application of LDA
- We show seasonal dynamics of energy lifestyles that some households maintain a single lifestyle, others multiple lifestyles.
- There are practical applications of lifestyles for DSM program targeting (e.g., batteries for "home for dinner")






# Part II

## Distributed energy resource (DER) control via private data

# Trade-off between running DSM and preserving privacy

There is an inherent tradeoff:

- A service provider (e.g., a utility company, DER aggregator) needs privacy compliance for users to exchange the required data.
- DER operation requires user consumption data that may reveal sensitive information.

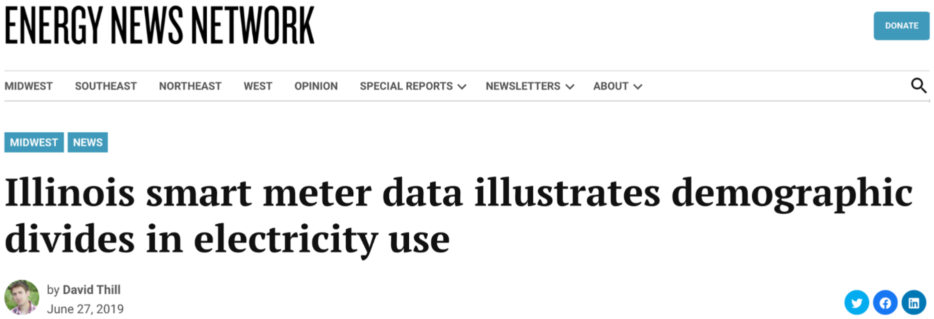

Forbes

Jun 1, 2014, 01:15pm EDT

Smart Meters: Between Economic Benefits And Privacy Concerns

Federico Guerrini Contributor

Tech


ENERGY NEWS NETWORK

DONATE

MIDWEST SOUTHEAST NORTHEAST WEST OPINION SPECIAL REPORTS NEWSLETTERS ABOUT

MIDWEST NEWS

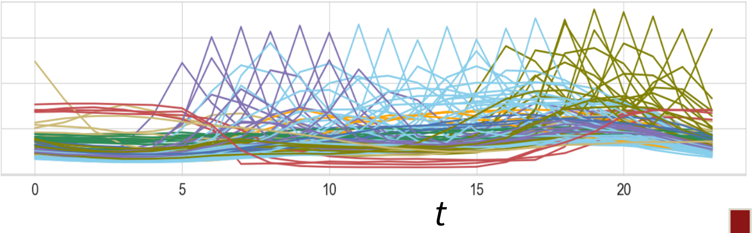
Illinois smart meter data illustrates demographic divides in electricity use

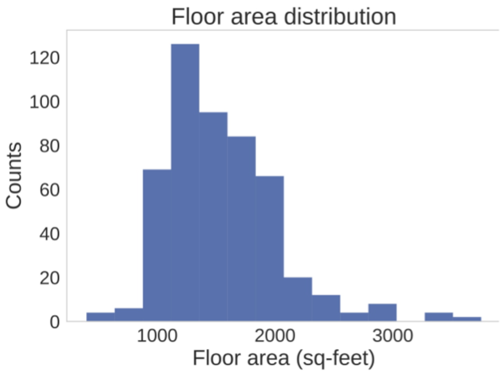
by David Thill

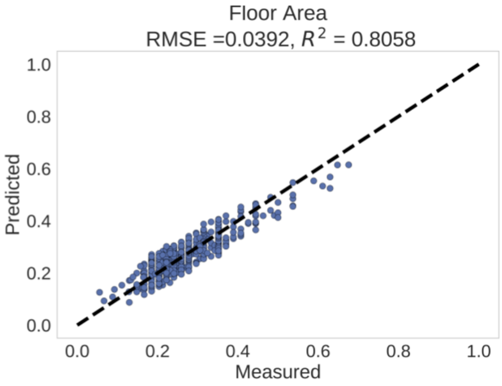
June 27, 2019

# Privacy risk in smart meter

Smart meter data



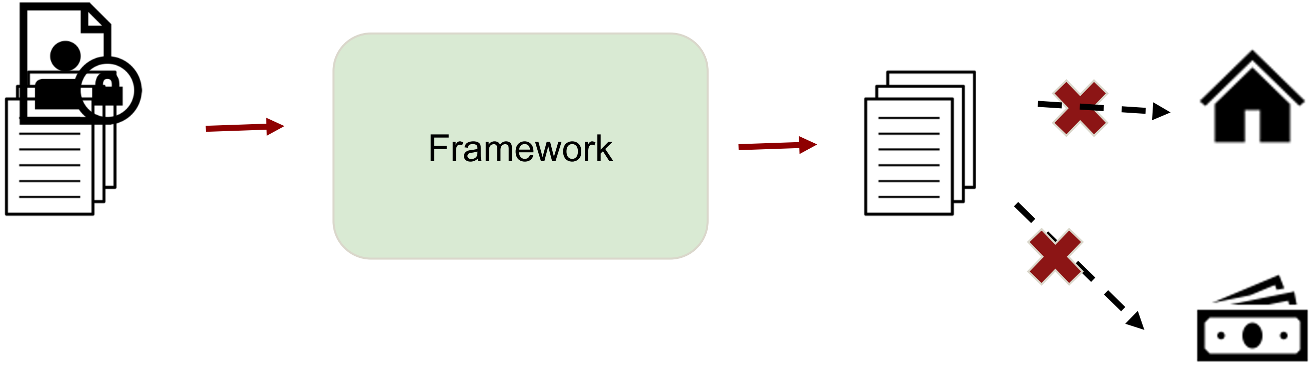

Protecting private attributes is important:

- Sq-ft of houses
- Age
- Income

[Chen, Kairouz, and Rajagopal, CDC 2018]

# Privacy: goal

Find a mechanism or framework to protect private attributes of users while maintaining data utility for certain tasks.



# Privacy: research questions

**How can we protect user private information while maintaining data utility from smart meters for designated tasks?**

1. How to quantify the privacy in the context of DER management?
2. How to generate the privacy preserved load data?
3. How to run DER control based on privatized data?

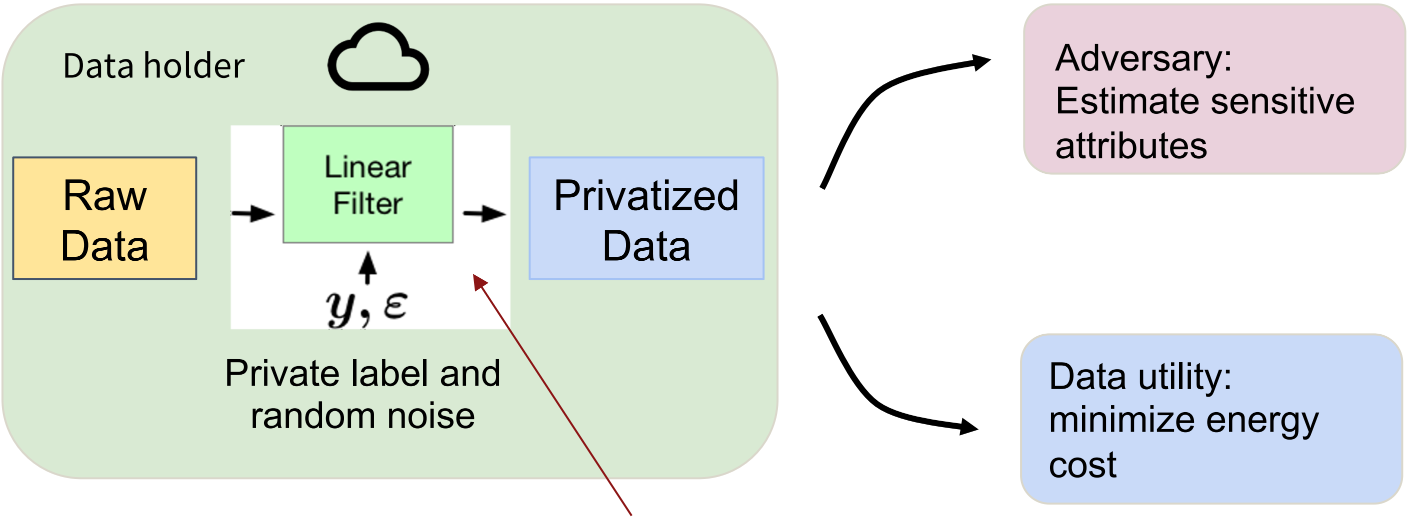
# Quantifying data privacy and utility

- **Privacy**: Classification accuracy of a private attribute.
  - detecting a private attribute from smart meter data.
  - Less accuracy indicates more privacy

- **Data utility**: Performance of battery optimal control. (e.g. minimum of energy cost)

**Case study:** storage control for energy cost minimization with TOU price

# The workflow of preserving privacy




$$\tilde{d} = d + G\begin{bmatrix}\varepsilon \\ y\end{bmatrix}$$

# Method: formulating a minmax optimization

$y$ private label

$\psi$ adversarial classifier

$L_a$ cross-entropy loss

Adversary minimizes classification error:

$$\min_{\psi} \mathcal{L}_a\left(f_\psi\left(\tilde{d}\right), y\right)$$

System planner minimizes energy cost (demand is stochastic):

$$\min \mathcal{L}_u(\mathbf{x},\ d) = \min_{x \in \mathcal{X}} \mathbf{E}_{d \sim P_{data}}\left[p^\top (x+d)_+\right] + \phi(x)$$

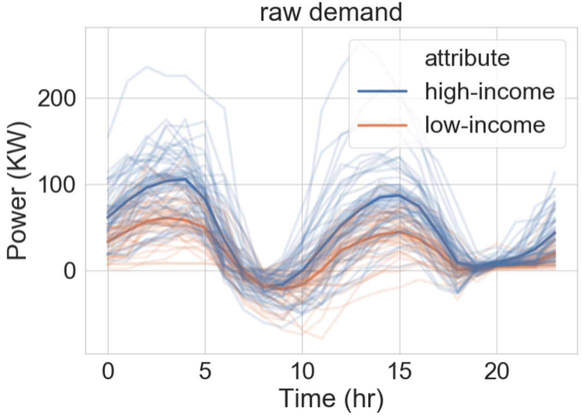
and maximize adversarial loss

$$\lambda \max_{G} \min_{\psi} \mathcal{L}_a\left(f_\psi\left(\tilde{d}\right), y\right)$$

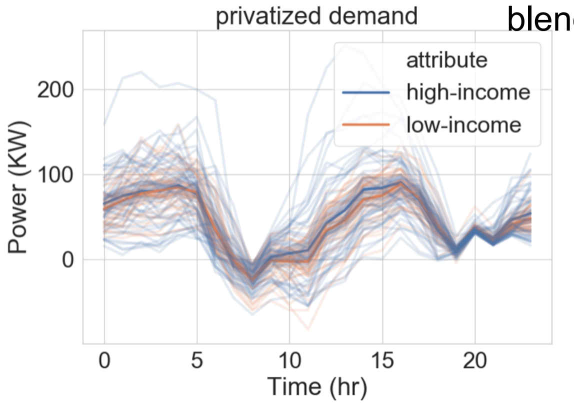
where $\lambda$ is a weighing factor between utility and privacy

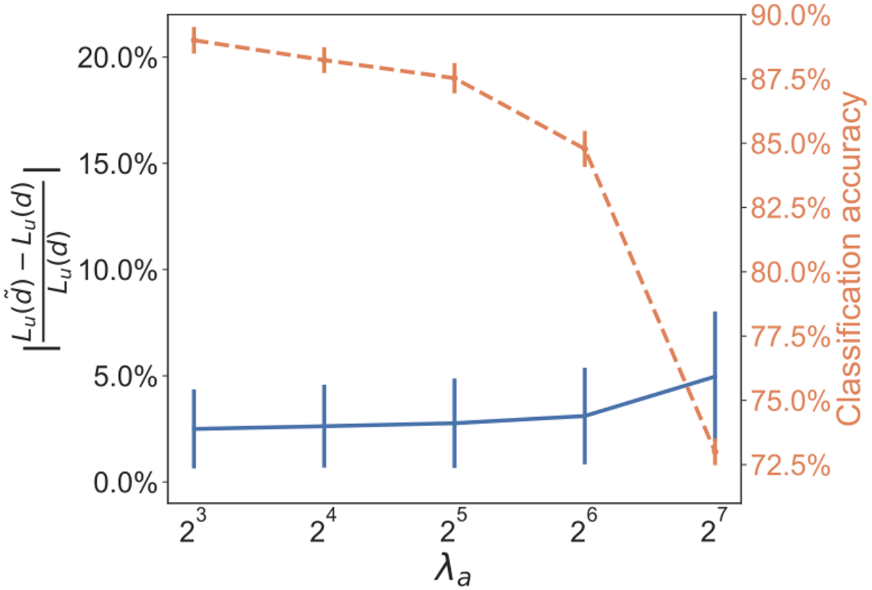
# Case study: aggregated homes

clear separation of groups



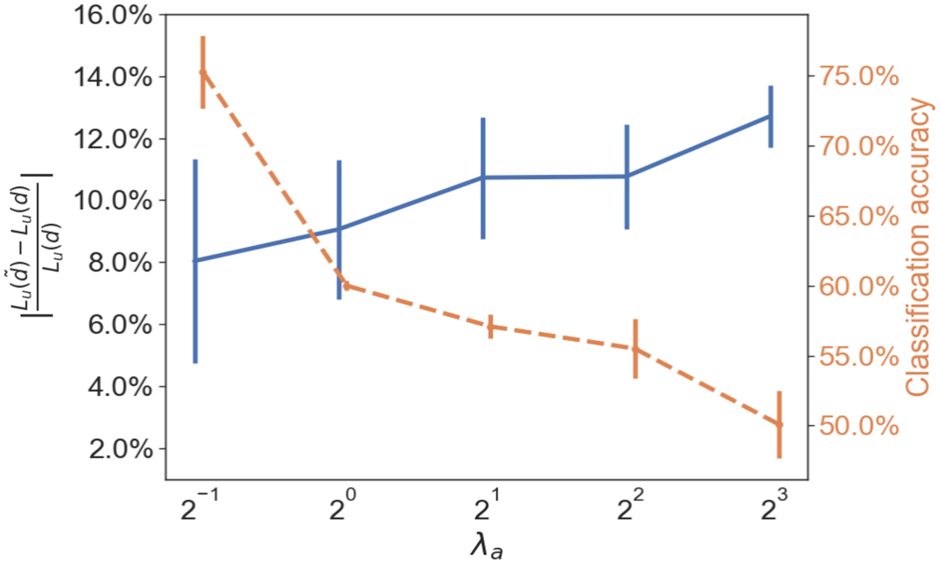

blends two groups




- Tradeoff between private label classification accuracy and cost savings.
- Achieve 17% decrease in classification accuracy with 5% decrease in cost savings.

# Case study: Commission for Energy Regulation (CER) data




| **train/test(%)** | baseline | 65/35 | 70/30 | 75/25 | 80/20 | 85/15 |
|---|---|---|---|---|---|---|
| **acc. (%)**† | 77.5 | 63.3 | 61.1 | 57.6 | 58.5 | 56.9 |
| **cost gap (%)*** | 0 | 11.7 | 9.4 | 12.2 | 10.0 | 10.9 |

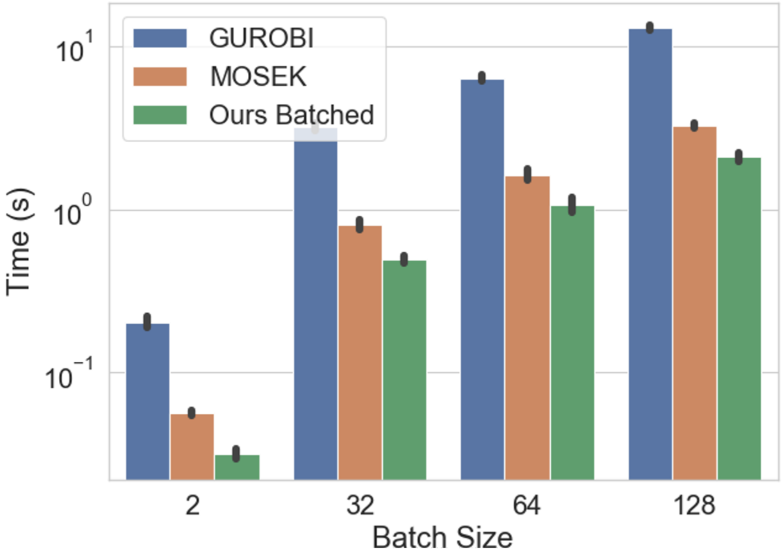
Generative filter is robust to training size of the data ($\lambda$ = 2)

Tradeoff between private label classification accuracy and cost savings.

We can reduce classification accuracy to that of random guessing with 12% reduced cost savings.

[Commission for Energy Regulation (CER). (2012). CER Smart Metering Project - Electricity Customer Behaviour Trial, 2009-2010 (dataset) ]

# Implementation of inbatch parallelism

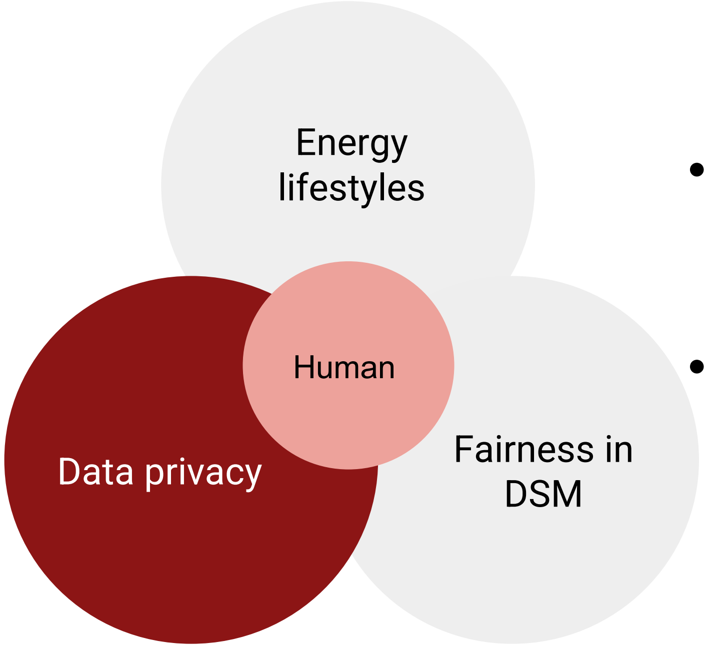



Our approach is

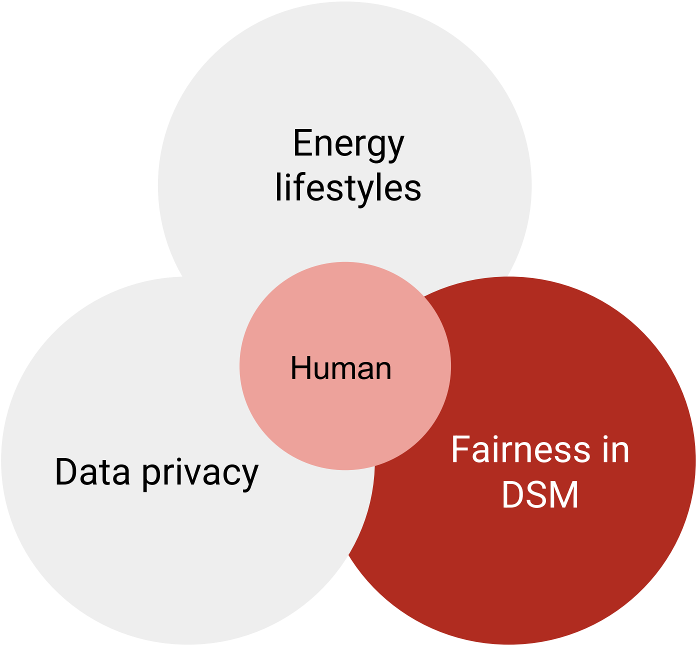
- Based on pytorch
- Automatic differentiation
- In-batch parallelism
- Faster than SOTA solvers

[open source: https://github.com/markcx/DER_ControlPrivateTimeSeries]

# Conclusions about data privacy




- We developed a minmax method to privatize data through noise injection that balances the tradeoff between data privacy and data utility
- We showed limited performance loss in terms of minimizing the energy cost while protecting private attributes of homes
- We injected noise through a simple linear filter that is lightweight to run on small local devices after training




# Part III

# Fairness-aware demand response

# Background: running decision-making fairly



Algorithmic Accountability Act of 2019

116TH CONGRESS
1ST SESSION

**H. R. 2231**

IN THE HOUSE OF REPRESENTATIVES

April 10, 2019

Ms. Clarke of New York introduced the following bill; which was referred to the Committee on Energy and Commerce

**A BILL**

To direct the Federal Trade Commission to require entities that use, store, or share personal information to conduct automated decision system impact assessments and data protection impact assessments.

**Section 1. Short title**

Article | Published: 16 December 2019

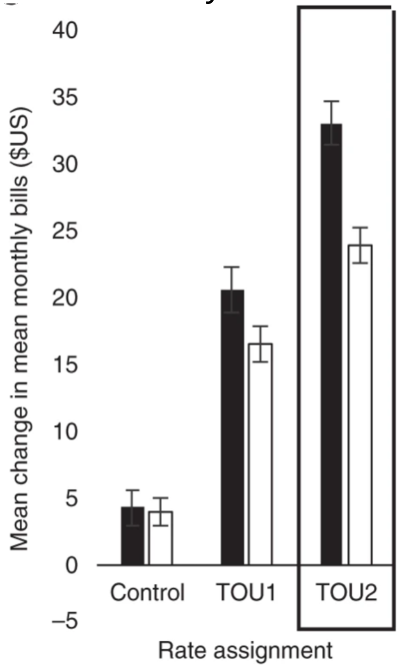
**Health and financial impacts of demand-side response measures differ across sociodemographic groups**

Lee V. White & Nicole D. Sintov

*Nature Energy* **5**, 50–60 (2020) | Cite this article

# Fairness: Motivations

Elderly



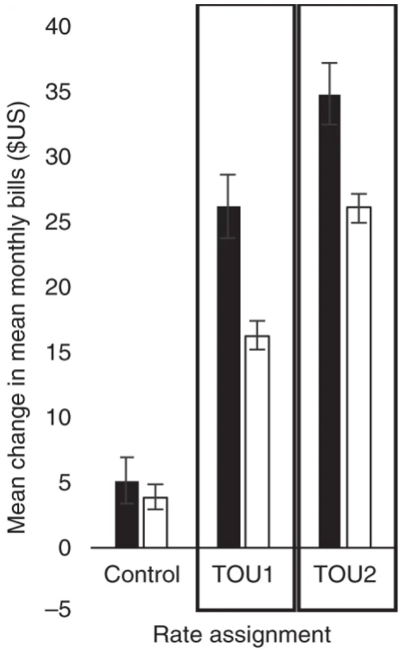

Vulnerable

(low income, elderly, disabled, etc.)

Non-vulnerable

Monthly bill increases disproportionately in the vulnerable group

Disabled



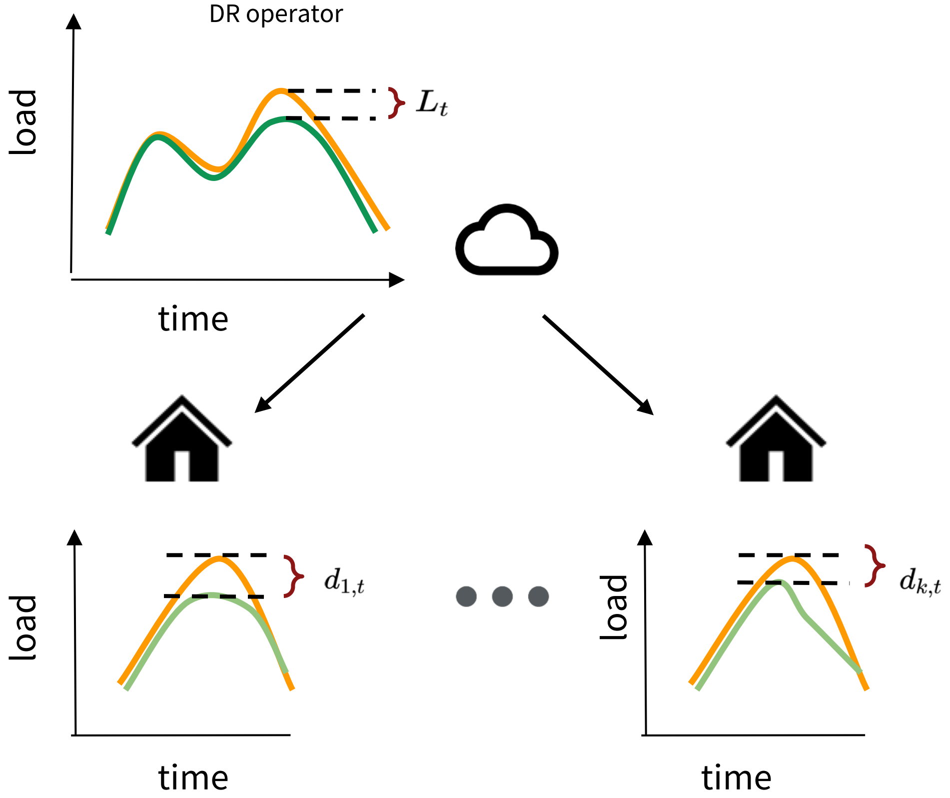

White, L.V., Sintov, N.D. Health and financial impacts of demand-side response measures differ across sociodemographic groups. Nature Energy 5, 50–60 (2020)

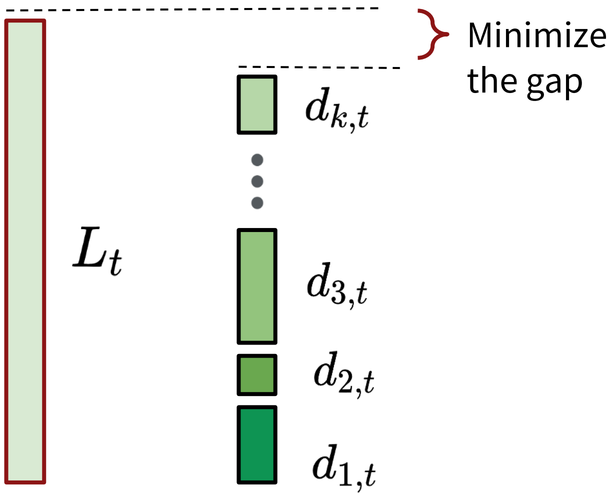
# Fairness: Research questions

**How can we prevent discrimination against certain groups throughout the demand response program?**

1. What fairness metric shall we consider? (formulate problem)
2. How do we fairly select users in demand response programs? (develop alg. )
3. What is the performance of our proposed method? (eval. performance)

# Settings of a demand response (DR) program




Assumptions:

- DR operator has direct controls
- *K* homes are enrolled
- Each home can reduce demand at some level

# DR optimization problem (without fairness)

$$\min_{x_t} \overbrace{\left(L_t - d_t^\top x_t\right)_+}^{f_{d,t}(x_t)}$$

$$c_t^\top x_t \leq \tilde{b}$$

$$x_t \in \{0,\, 1\}^K$$

At time $t$, we have

| | |
|---|---|
| $L_t$ | Load curtailment target scalar |
| $d_t$ | Demand reduction vector |
| $\tilde{b}$ | Payment budget scalar |
| $c_t$ | Payment cost vector |
| $x_t$ | Binary vector of decision variable indicating selection of users |
| $f_{d,t}(x_t)$ | Demand response objective |

# Quantifying group fairness

equal opportunity: *P(X=1 | vulnerable ) = P(X=1 | non-vulnerable)*

$$\gamma \leq \frac{P(X=1 \mid vulnerable)}{P(X=1 \mid non\text{-}vulnerable)} \leq 1$$

Vulnerable

*J*

Non-Vulnerable

*K-J*

$$\frac{1}{TJ}\sum_{t=1}^{T}\sum_{j=1}^{J} x_{j,t} - \frac{1}{T(K-J)}\sum_{t=1}^{T}\sum_{j=J+1}^{K} x_{j,t} \leq 0$$

Expand the fairness constraints out as two linear constraints

# Formulate a general online optimization

$$\min_{x_t} \sum_{t=1}^{T} f_{d,t}(x_t) \quad + \phi\left(\frac{1}{T}\sum_{t=1}^{T} C_t x_t\right)$$

$$\sum_{t=1}^{T} C_t^\top x_t \leq \mathbf{b}T$$

$$x_t \in [0,1]^K$$

$C_t$ and $\mathbf{b}$

- Budget constraint
- Fairness constraints

$\phi(\cdot)$

- Regularization

Offline: parameters *d, C* are known from step *1* to *T* (future)
Online: parameters *d, C* are known up to *t-1*
Regret: the gap between online and offline

# Fully vs partially observed demand reductions

1. When *d* of all customers are observed each round:
   - Primal-Dual Gradient updates with regularization and fairness (PDG-RF)

**Theorem 1**: PDG-RF achieves regret of $O\left(\sqrt{T}\right)$

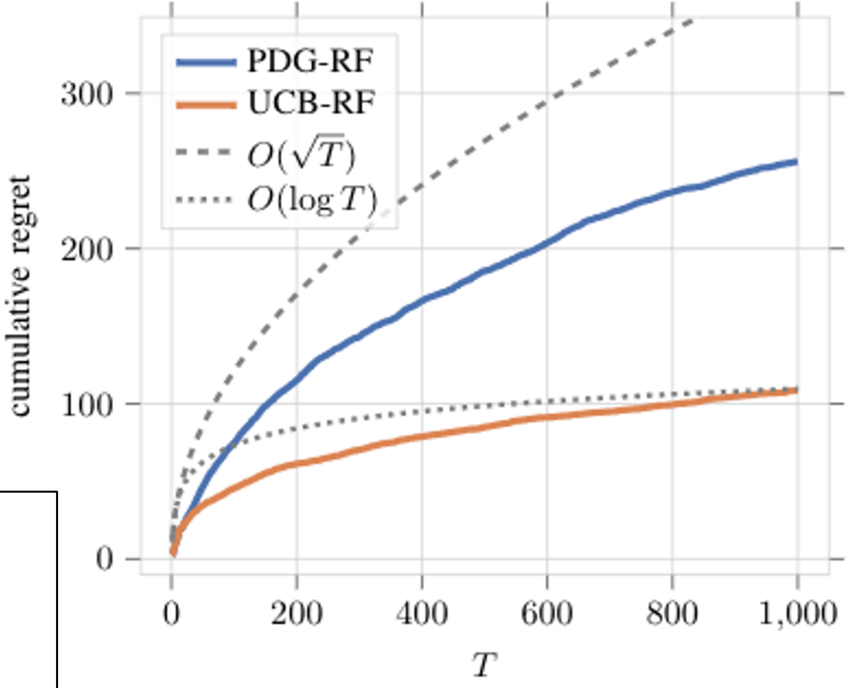
2. When *d* of some customers are observed each round:
   - Upper-confidence bound with regularization and fairness (UCB-RF)

**Theorem 2**: UCB-RF achieves regret of $O(\log T)$

# Case1: simulating DR in a small group

Simulation setup:

1. 50 homes:
   - 50% in vulnerable group
   - $d_t$ (vul)~ Unif[0.1, 0.5]
   - $d_t$ (non-vul) ~ Unif[0.45, 1]
2. Budget
   - $\tilde{b}$ ~ Unif[1, 5]
3. Payment Cost
   - c ~ Unif[0.1, 1]



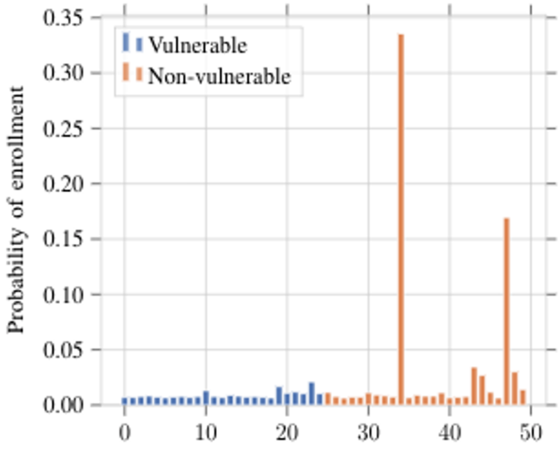

**Findings:**

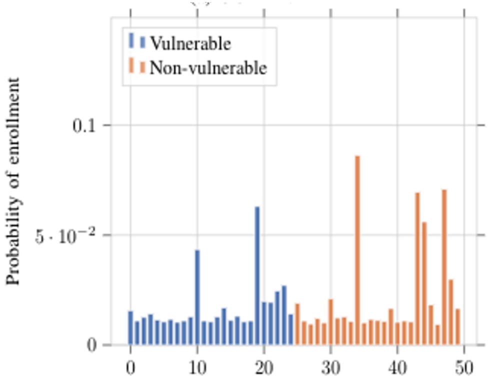
- Both methods match with theoretical regret bounds
- UCB-RF converges faster than PDG-RF

Introduction Part I Part II Part III Conclusion

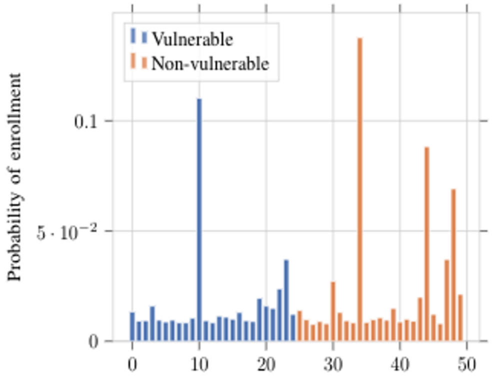
# Case1: Selection probability in DR events (small group)

No fairness



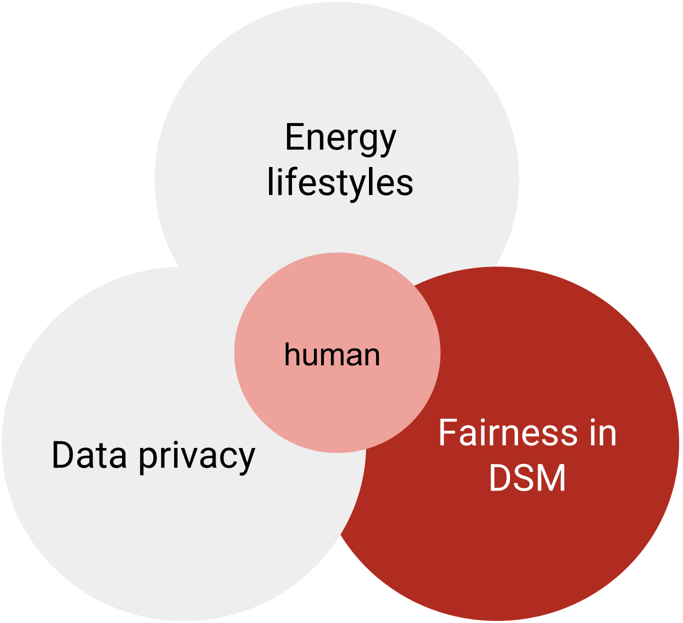

PDG-RF




UCB-RF

The selection probability are more dispersed when considering fairness

# Conclusions about fairness-aware DR




We proposed a simple constraint, a variant of equal opportunity, to capture the group fairness

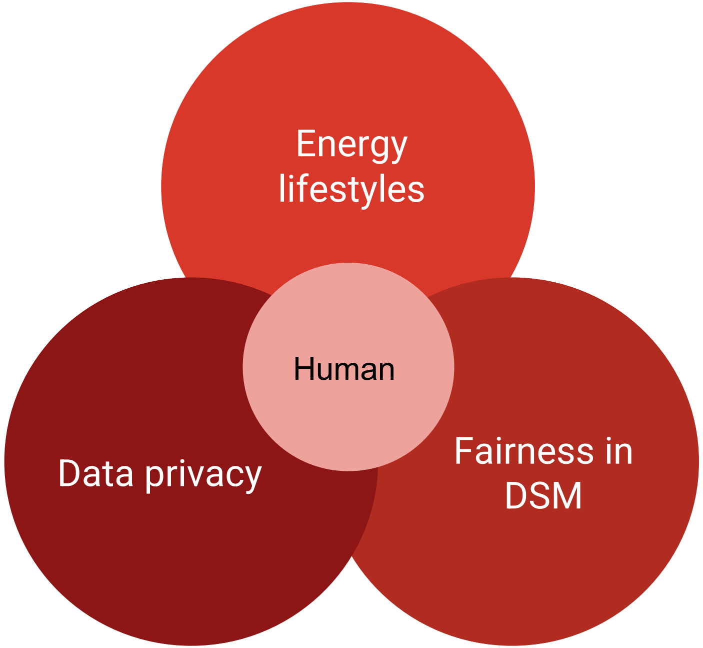
We introduced two algorithms to solve the fairness-aware demand response problem with sublinear regrets

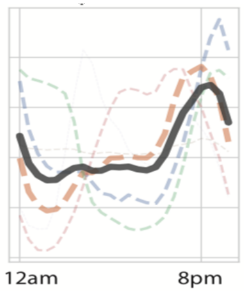
Enrollment probability of the vulnerable group is more dispersed when fairness is introduced

# Concluding remarks about human-centric DSM

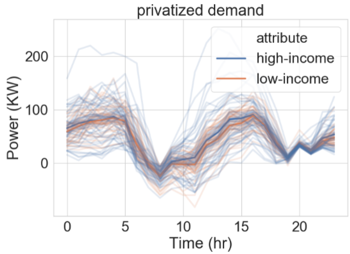
1. Energy lifestyles
   a. We construct 6 typical lifestyles that can be used for DSM
   b. We reveal dynamic changes of lifestyles
2. Data privacy
   a. We preserve privacy by reducing the corr. between raw data and sensitive info.
   b. Our minimax framework can be applied to many privacy protection tasks
3. Fairness in DSM
   a. We introduce fairness metrics in the context of DR
   b. We develop two algorithms with sublinear regret (under two settings)

# Concluding remarks: overview

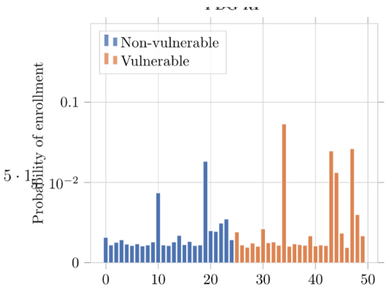



Sparse representation of energy use




Less correlation with sensitive attributes



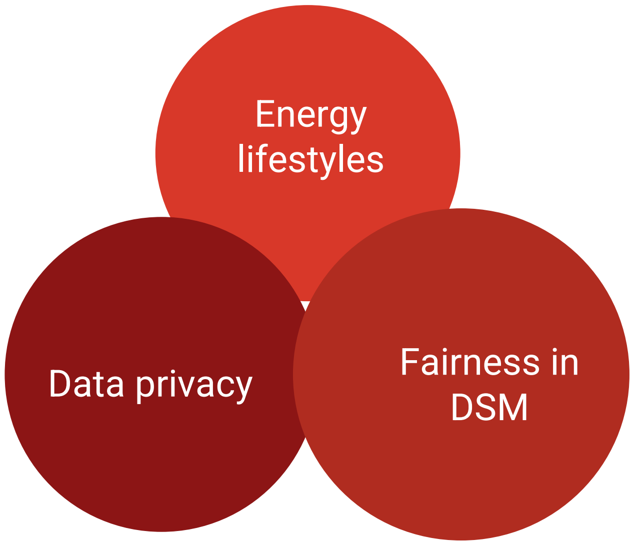

Enrollment probability

# Future work

Lifestyles:
- Adding socio-demographic info.
- Estimating the dynamic transition pattern and causal study on changes

Privacy:
- Differential privacy on lifestyles (e.g., counts)
- Relaxed Differential privacy on time series

Fairness:
- Contextual bandit
- Distributionally Robust Optimization



# Acknowledgements

# Acknowledgements


## S3L members

Gustavo Cezar
Aaron Goldin
Jose Bolorinos
Ryan Triolo
Justin Luke
Lily Buechler
Siobhan Powell
Oskar Triebe
Tao Sun
Sonia Martin
Anthony Degleris
Tianyuan Huang
Zhecheng Wang
Others…

## Former members

Raffi Sevlian
Jungsuk kwac
Junjie Qin
Jiafan Yu
Yang Yu
Yizheng Liao
Sid Petal
Sam Borgeson
Michaelangelo Tabhone
Baosen Zhang
Yang Weng
Wenyuan Tang
Yuting Ji
Nikolay Laptev
Mohammad Rasouli
Heidi von korff

# Acknowledgements

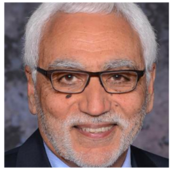
Collaborators:

Thomas Navidi
Chad Zanocco
Hao Wang
Peter Kairouz
Sila Kiliccote
Emre Can Kara
Chong Huang
Lalitha Sankar
Stefano Ermon
David Lowbell
Marshall Burke
Hao Sheng
Andrew Ng
June Flora

Committee:

Chin-Woo Tan
Dorsa Sadigh
Martin Fischer
Abbas El Gamal
Ram Rajagopal




# Publications

*[TRR 2016] A Parking sensing and information system: Sensors, deployment, and evaluation

*[Entropy 2017] Context-aware generative adversarial privacy

[IEEE CDC 2018] Understanding compressive adversarial privacy

[IEEE PESGM 2019] Electric vehicle charging during the day or at night: a perspective on carbon emissions

[ICLR 2019] Distributed Generation of private data with user customization

*[PSCC 2020 & EPSR 2020] ***<u>Energy resource control via privacy preserving data</u>***

[KDD 2018] Infrastructure quality assessment in Africa using satellite imagery and deep learning

[CVPR 2020] Effective data fusion with GVI: land cover segmentation in agriculture

[IISE 2020] Using Satellite Imagery to Automate Building Damage Assessment (best student paper)

*[Applied Energy, 2nd review] ***<u>Constructing dynamic residential energy lifestyles using LDA</u>***

*[IEEE Trans. Smart Grid, working] ***<u>Fairness-Aware DR: Online Learning Approach and Regret Analysis</u>***

* Journal article ***<u>Covered in this talk</u>***

# Thank you

QUESTIONS ?

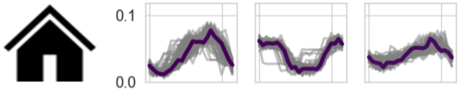
# Human-centric demand side management in electricity

## LIFESTYLES, PRIVACY, FAIRNESS

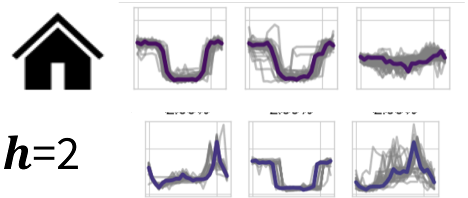
Questions?

# Appendix

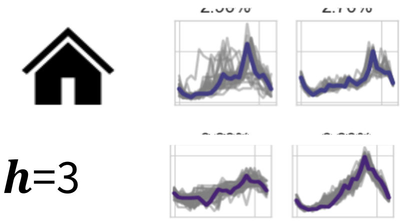
# Intuition of Latent Dirichlet Allocation




$\boldsymbol{h}$=1



$\boldsymbol{h}$=2



$\boldsymbol{h}$=3

$\theta_1 = [0, 0.1, 0.4, 0.5]$

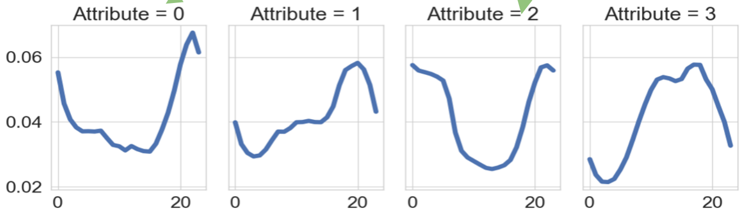
$\theta_2 = [0.3, 0, 0.7, 0]$

$\theta_3 = [0, 1, 0, 0]$

$\boldsymbol{\theta_h} \sim \text{Dir}(\boldsymbol{\alpha})$

$\boldsymbol{\beta}_k \sim \text{Dir}(\gamma)$

$k$=1:$K$




$$\max \arg_{\theta,\,\beta} P\left(\{s\}_{n=1,h=1}^{N,H} \,\middle|\, \beta_{1:K}, \theta_{1:H}\right)$$

$\boldsymbol{\beta}_1$=[0.01, 0, 0.008, …]

$\boldsymbol{\beta}_2$=[0, 0.02, 0.003, …]

…, $\boldsymbol{\beta}_K$

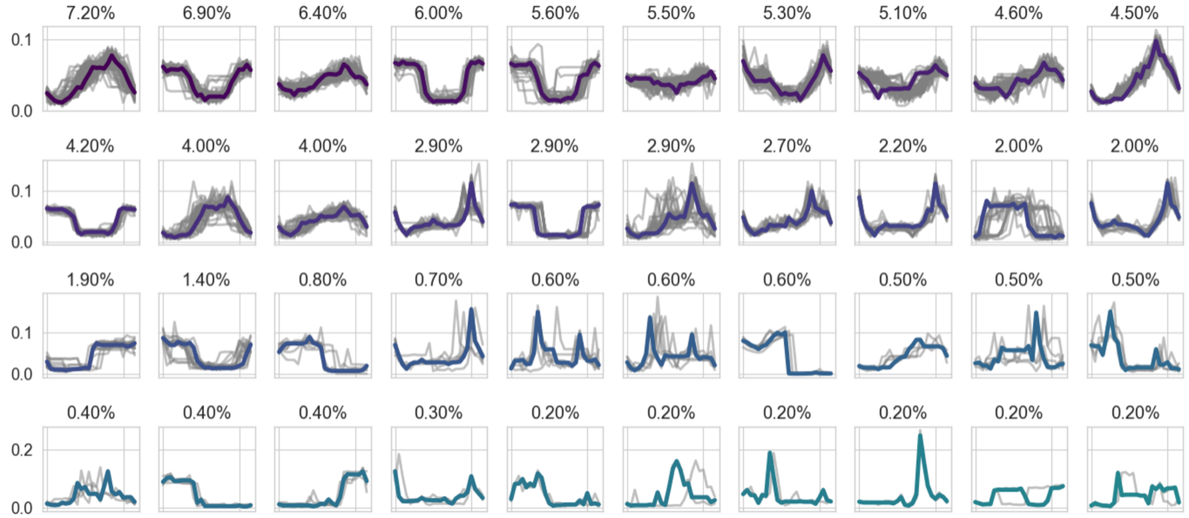
# LDA maximum likelihood

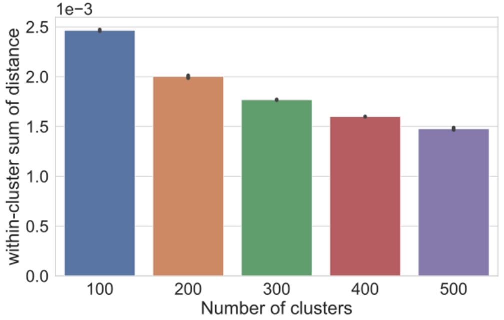
$$\hat{\beta}_{1:K}, \hat{\theta}_{1:D} = \arg\max_{\beta,\, \theta} P\left(\{s\}_{n=1,h=1}^{N,H} \,\middle|\, \beta_{1:K}, \theta_{1:H}\right)$$

$$= \arg\max_{\beta,\theta} \Pi_{n=1}^{N} \Pi_{h=1}^{H} P(s_{nh} \mid \beta_{1:K}, \theta_h)$$

$$= \arg\max_{\beta,\, \theta} \Pi_{n=1}^{N} \Pi_{h=1}^{H} \sum_{z=1}^{K} \underbrace{P(s_{nh} \mid z)}_{\text{Elements of } \beta_z} \underbrace{P(z \mid \theta_h)}_{z\text{-th element of } \theta_h}$$

# The Dictionary of load shapes




We pick the dict size of 200

# Energy lifestyles: Load shape dictionary

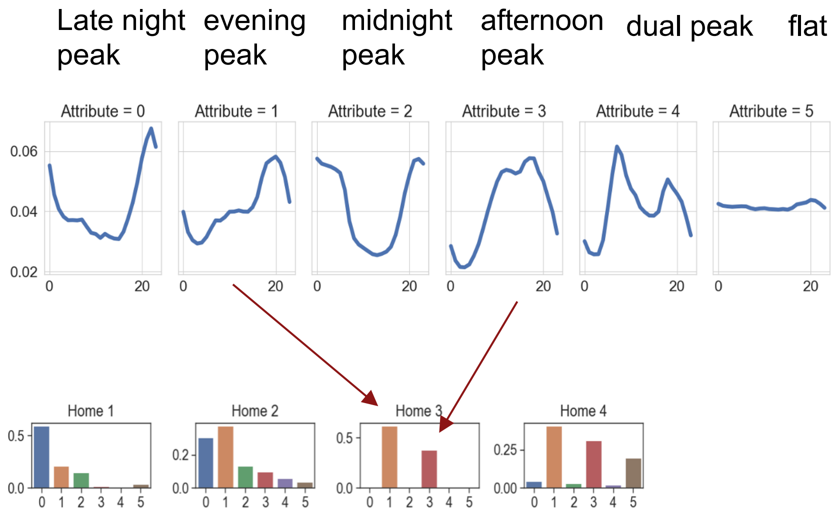
Compare several distance metrics:
- Euclidean
- L1
- Cosine
- Dynamic Time Warping
- Hybrid (w*Euclidean + (1-w)*DTW )

Evaluate the separation of clusters:
- Calinski-Harabasz Index (CHI)
- Davies-Bouldin Index (DBI)

| method | distance | $CHI \uparrow$ | $DBI \downarrow$ |
|---|---|---|---|
| $k$-means | Euclidean | 107.42 | 4.53 |
| | cosine | 102.31 | 3.87 |
| | $\ell_1$ | 99.51 | 4.14 |
| | DTW | 113.93 | 3.89 |
| | $d_{hybrid}(\gamma = 0.5)$ | 116.76 | 3.67 |
| **$k$-median** | Euclidean | 109.53 | 4.50 |
| | cosine | 108.11 | 4.05 |
| | $\ell_1$ | 102.40 | 4.19 |
| | DTW | 115.84 | 3.82 |
| | $\boldsymbol{d_{hybrid}(\gamma = 0.5)}$ | **118.31** | **3.54** |
| Hierarchical (Ward) | Euclidean | 93.21 | 4.99 |
| | cosine | 92.18 | 4.81 |
| | $\ell_1$ | 90.53 | 5.16 |
| | DTW | 98.65 | 4.87 |
| | $d_{hybrid}(\gamma = 0.5)$ | 101.32 | 4.58 |
| DBSCAN ($\epsilon = 0.1$) | Euclidean | 82.44 | 5.17 |
| | cosine | 85.37 | 5.29 |
| | $\ell_1$ | 80.15 | 5.18 |
| | DTW | 88.03 | 5.25 |
| | $d_{hybrid}(\gamma = 0.5)$ | 89.75 | 5.07 |

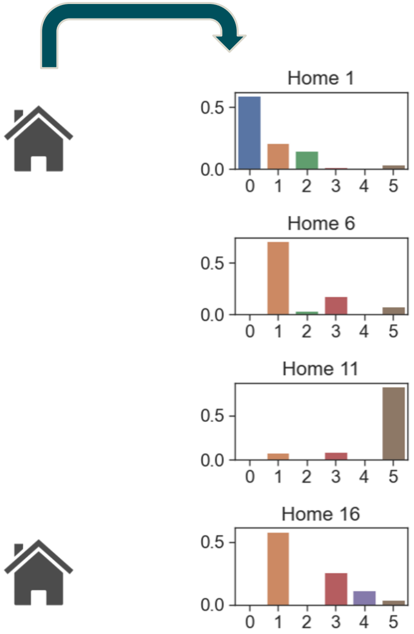
# Generating energy attributes and lifestyles

We apply Latent Dirichlet Allocation to generate six latent energy attributes

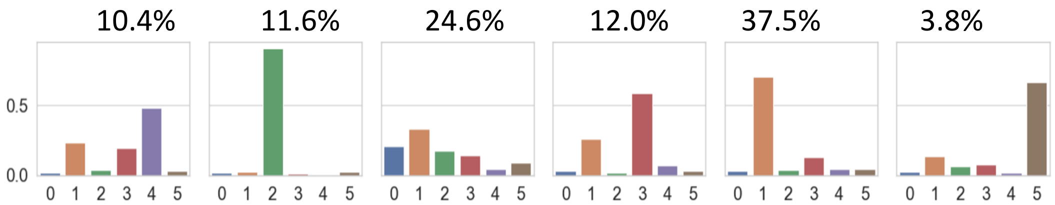


# Constructing energy lifestyles

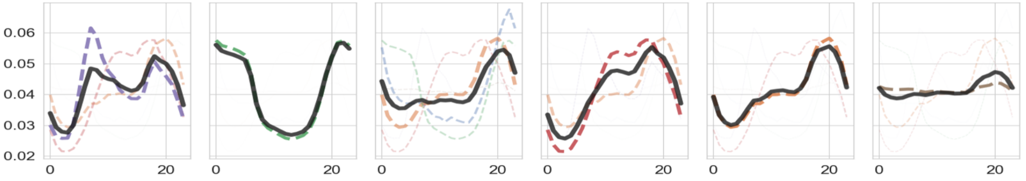
Load shapes → LDA → Lifestyles




We apply k-mean to cluster the homes which are characterized by 6-dim vectors



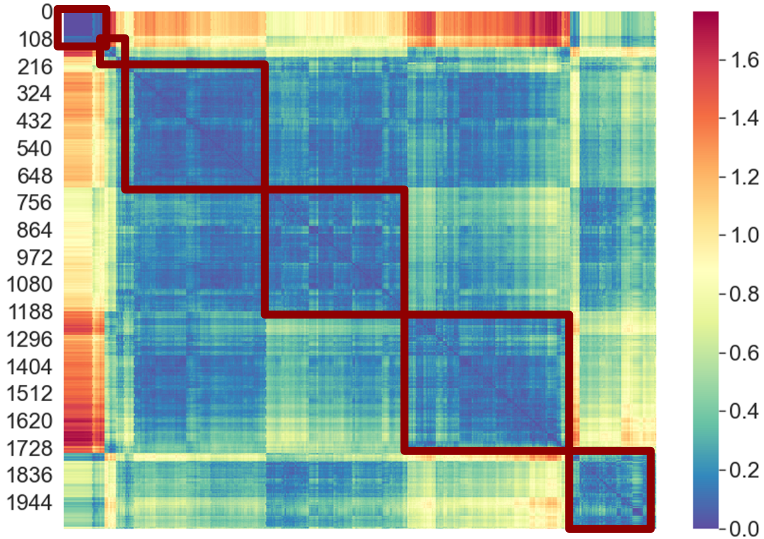

Active morning | Night owl | Everyday is a new day | Home early | Home for dinner | Steady going



# The energy lifestyles

Load shapes | LDA | *Lifestyles*

Why six topics ?




Pairwise distance matrix

# Temporal characteristics and features

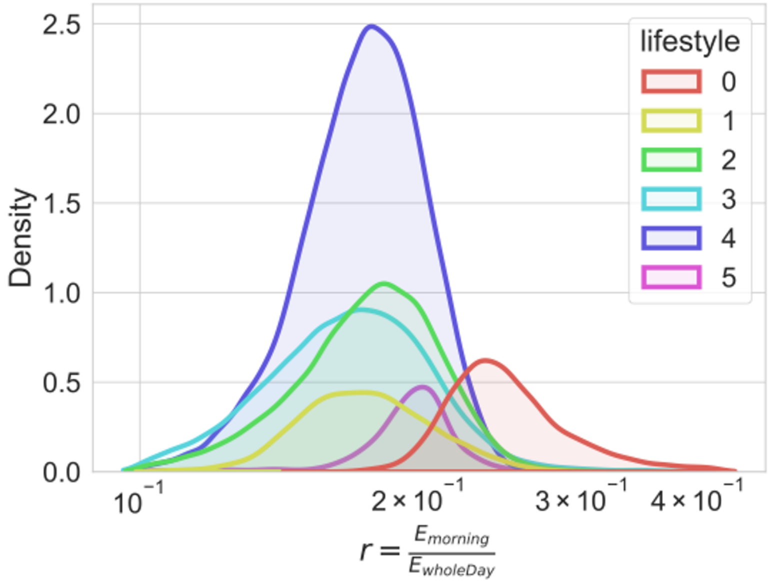
| feature | description |
|---|---|
| $E_{day}$ | mean of daily energy use |
| $E_{hour}$ | mean of hourly energy use |
| $E_{peak}$ | mean energy use of peak hour in a day, equivalent to $E_{max}$ |
| $E_{base}$ | mean base energy use of a day |
| $E_{min}$ | mean of min energy use of a day |
| $E_{morning}$ | morning energy use between 6am to 10am |
| $E_{noon}$ | morning energy use between 10am to 2pm |
| $E_{evening}$ | evening energy use between 6pm to 10pm |
| $E_{night}$ | night energy use between 10pm to 2am |
| $E_{wholeday}$ | 24 hour energy use |
| $r_{base}$ | base load ratio, i.e. mean of $\frac{E_{base}}{E_{day}}$ |
| $r_{min2max}$ | mean ratio of min hourly load divide by max hourly load, i.e. mean of $\frac{E_{min}}{E_{max}}$ |
| $r_{m2w}$ | mean of morning energy use divide by whole day energy use, i.e. mean of $\frac{E_{morning}}{E_{wholeday}}$ |
| $r_{n2w}$ | mean of noon energy use divide by whole day energy use, i.e. mean of $\frac{E_{noon}}{E_{wholeday}}$ |
| $r_{e2w}$ | mean of evening energy use divide by whole day energy use, i.e. mean of $\frac{E_{evening}}{E_{wholeday}}$ |
| $r_{ni2w}$ | mean of night energy use divide by whole day energy use, i.e. mean of $\frac{E_{night}}{E_{wholeday}}$ |
| $\pi_j$ | multinomial distribution over 24 hours showing the normalized frequency of peak hour occurrence. The $j$ takes value from 0, 1, ..., 23, indicating $j$-th peak hour in a day |

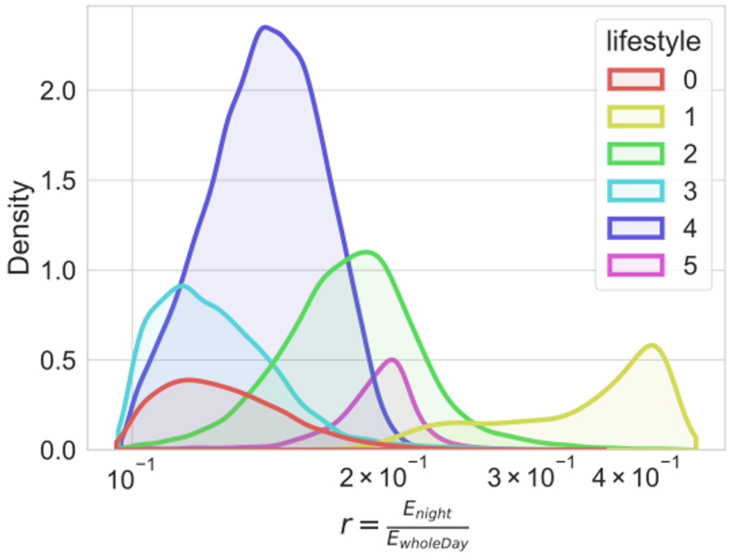
We use 44 different features :

- to assess the distinctions of different lifestyles;
- to evaluate distribution differences between changer and nochanger in every lifestyle group.

Once we determine the lifestyle categories of homes, we check the distribution of features based on raw meter data.

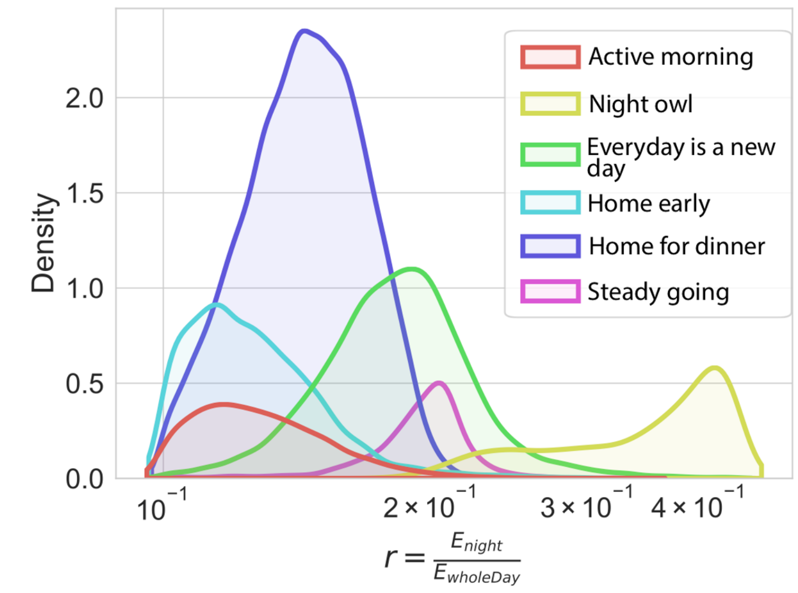
# Temporal characteristics and features



# Identifying lifestyles and detecting change of lifestyles

Once we determine the lifestyle categories, we check the distribution of features based on raw meter data.




Input:
Electricity features

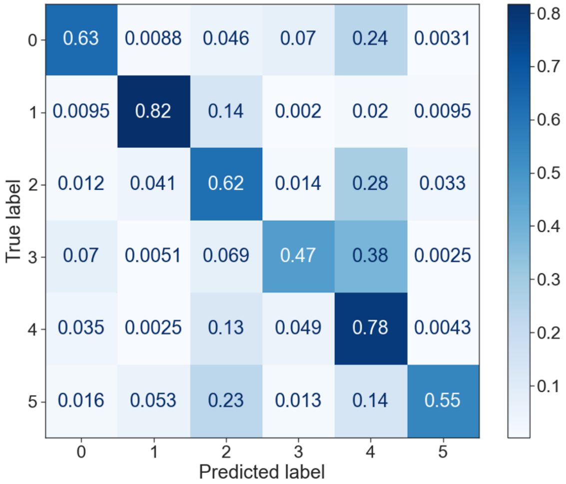
- mean daily consumption;
- ratio of morning use over whole-day use;
- peak hour frequency;
- (...44 features in total)

Output:
Lifestyles

Remark:

- Identifying lifestyles from few weeks observations
- Classification problem

# Results of classifying lifestyles




mance

| style index | lifestyle | precision | recall | F1 score |
|---|---|---|---|---|
| 0 | active morning | 0.7164 | 0.6315 | 0.6713 |
| 1 | night owl | 0.8657 | 0.8176 | 0.8409 |
| 2 | everyday is a new day | 0.6411 | 0.6191 | 0.6299 |
| 3 | home early | 0.6551 | 0.4685 | 0.5463 |
| 4 | home for dinner | 0.6652 | 0.7841 | 0.7198 |
| 5 | steady going | 0.6417 | 0.5493 | 0.5919 |
| | | | | $\boxed{\text{average acc} = 0.685}$ |

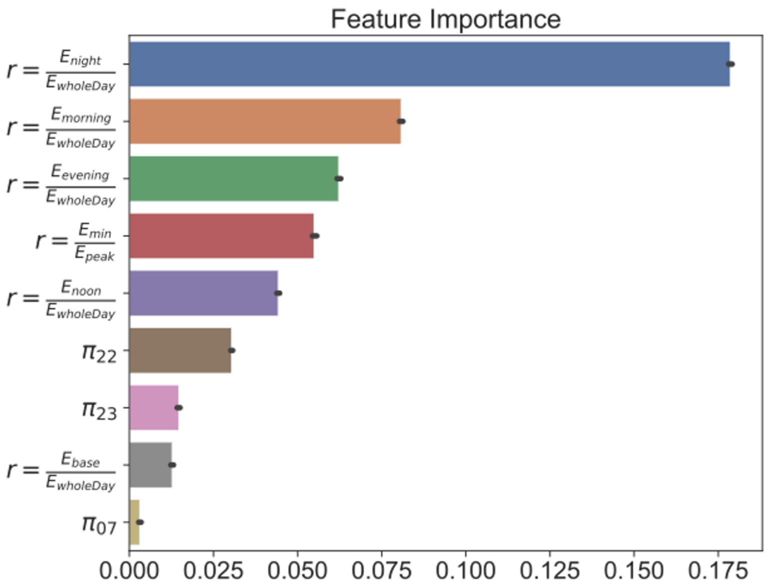
Input: energy consumption (features)
Output: lifestyles

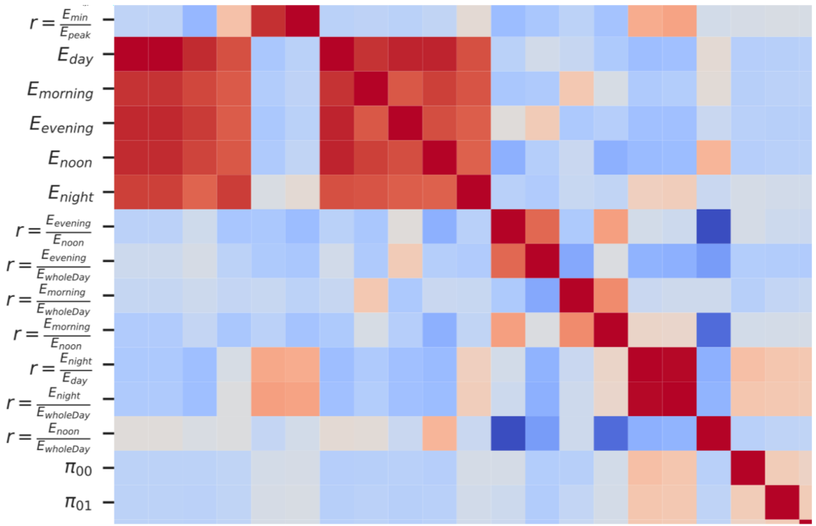
- We use random forest classifier
- We split data in train/val/test as 80/10/10%

Remark:

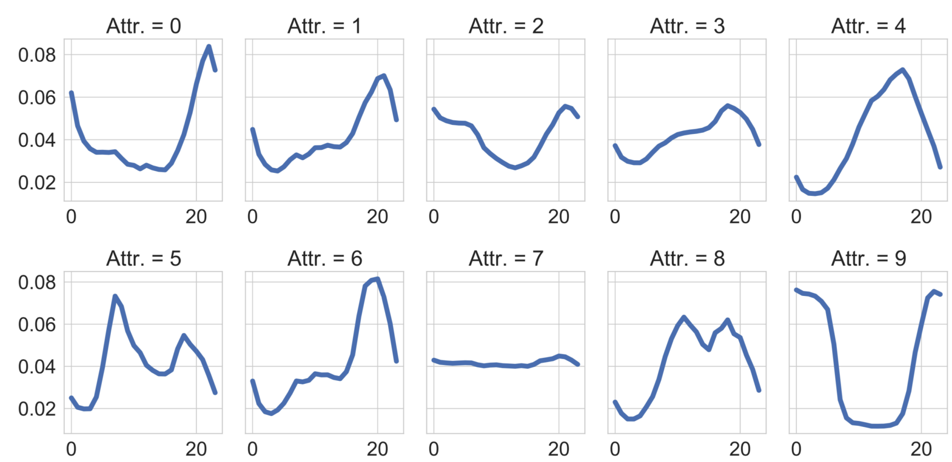
- ***Night owl*** is easy to classify
- ***Home early*** is difficult to classify as it is similar to ***home for dinner***

# Classifying lifestyles

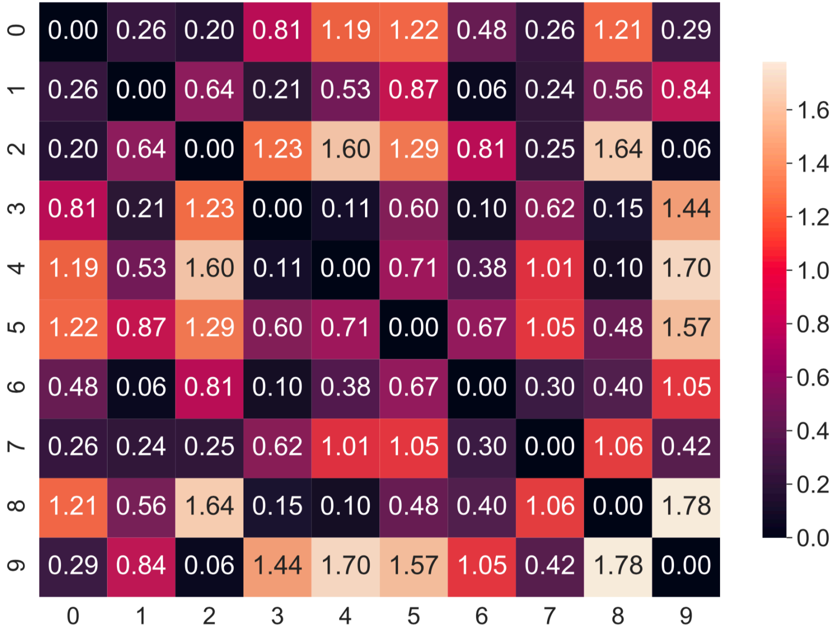


# Initial ten topics



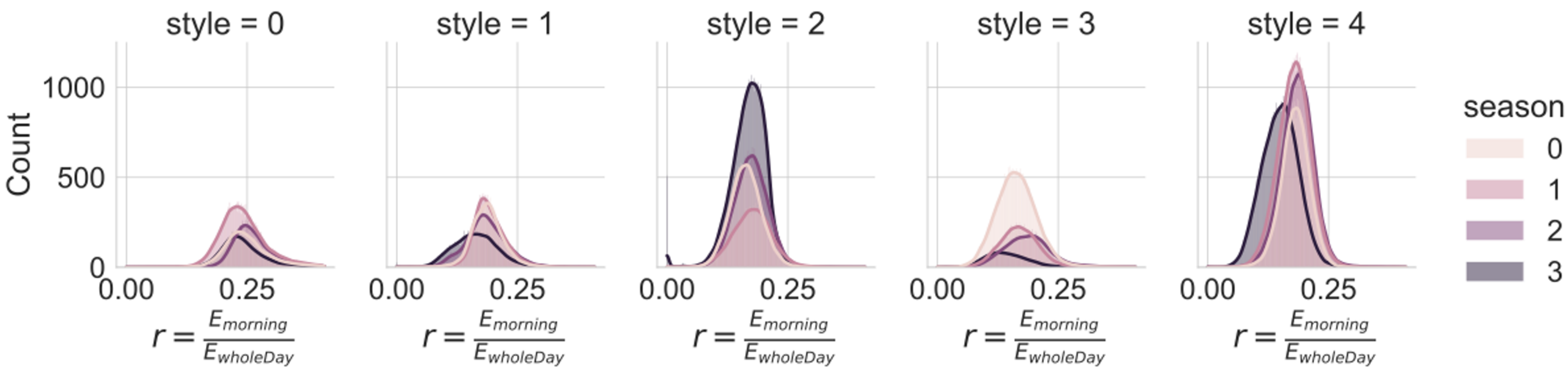

Calculate correlation distance, then apply projection



| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.00 | 0.26 | 0.20 | 0.81 | 1.19 | 1.22 | 0.48 | 0.26 | 1.21 | 0.29 |
| 1 | 0.26 | 0.00 | 0.64 | 0.21 | 0.53 | 0.87 | 0.06 | 0.24 | 0.56 | 0.84 |
| 2 | 0.20 | 0.64 | 0.00 | 1.23 | 1.60 | 1.29 | 0.81 | 0.25 | 1.64 | 0.06 |
| 3 | 0.81 | 0.21 | 1.23 | 0.00 | 0.11 | 0.60 | 0.10 | 0.62 | 0.15 | 1.44 |
| 4 | 1.19 | 0.53 | 1.60 | 0.11 | 0.00 | 0.71 | 0.38 | 1.01 | 0.10 | 1.70 |
| 5 | 1.22 | 0.87 | 1.29 | 0.60 | 0.71 | 0.00 | 0.67 | 1.05 | 0.48 | 1.57 |
| 6 | 0.48 | 0.06 | 0.81 | 0.10 | 0.38 | 0.67 | 0.00 | 0.30 | 0.40 | 1.05 |
| 7 | 0.26 | 0.24 | 0.25 | 0.62 | 1.01 | 1.05 | 0.30 | 0.00 | 1.06 | 0.42 |
| 8 | 1.21 | 0.56 | 1.64 | 0.15 | 0.10 | 0.48 | 0.40 | 1.06 | 0.00 | 1.78 |
| 9 | 0.29 | 0.84 | 0.06 | 1.44 | 1.70 | 1.57 | 1.05 | 0.42 | 1.78 | 0.00 |

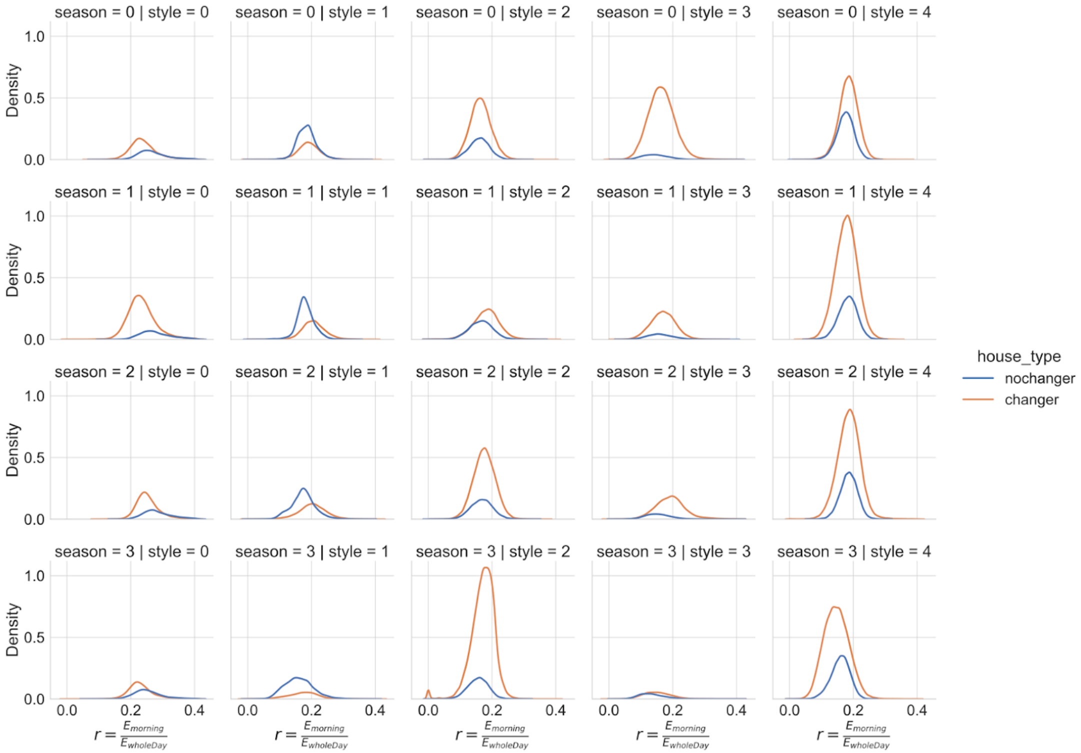
# Temporal characteristics – No Changer

Check the stability of nochangers



# Distributions in Changer and No Changer

We use random forest to classify changer vs nochanger



# Classify changers

Example results of changer classification on night owl lifestyle and home for dinner lifestyle

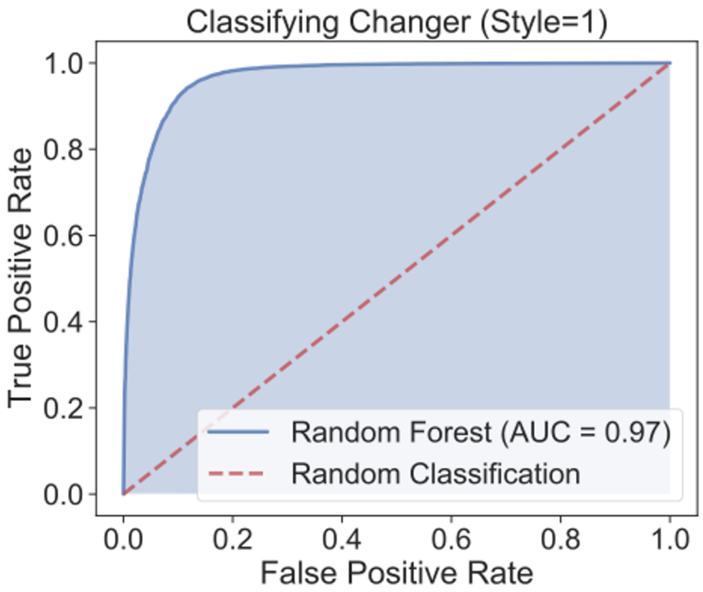
Table 7: night owl lifestyle

| label | precision | recall | F1 score |
|---:|---|---|---|
| nochanger (0) | 0.9301 | 0.8706 | 0.8994 |
| changer (1) | 0.9073 | 0.9508 | 0.9285 |
| | | | average acc = 0.917 |

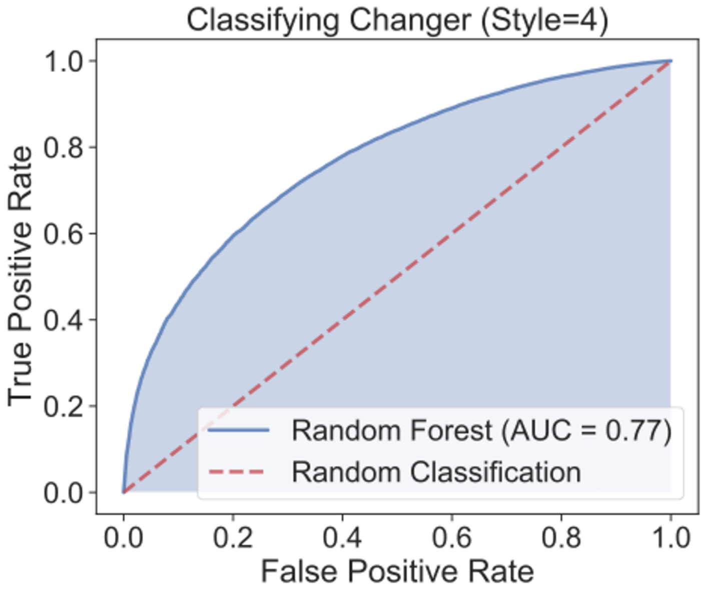
Table 10: home for dinner lifestyle

| label | precision | recall | F1 score |
|---:|---|---|---|
| nochanger (0) | 0.5919 | 0.1734 | 0.2682 |
| changer (1) | 0.9657 | 0.9897 | 0.9775 |
| | | | average acc = 0.856 |

# Classify Changers




Night owl

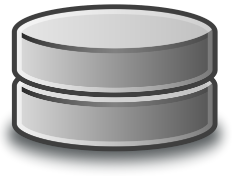



Home for dinner

# [Appendix] Findings of energy lifestyles

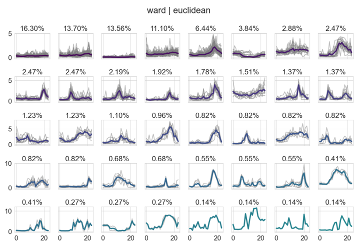
1. **Home for dinner:**
    - Most common lifestyle category
    - Higher proportion in winter, lower in summer,  high daily use within time-of-use price window
    - Commonly exchange with **everyday is a new day**
2. **Night owl:**
    - Occurs in ~9% of samples across all seasons
3. **Steady going:**
    - More commonly exchange with **everyday is a new day**
4. **Everyday is a new day**:
    - Commonly exchange with **home for dinner**

# Architecture (optional map-reduce)

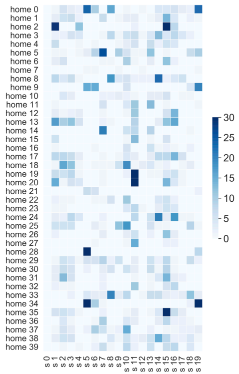


Data Loader/
Sampler

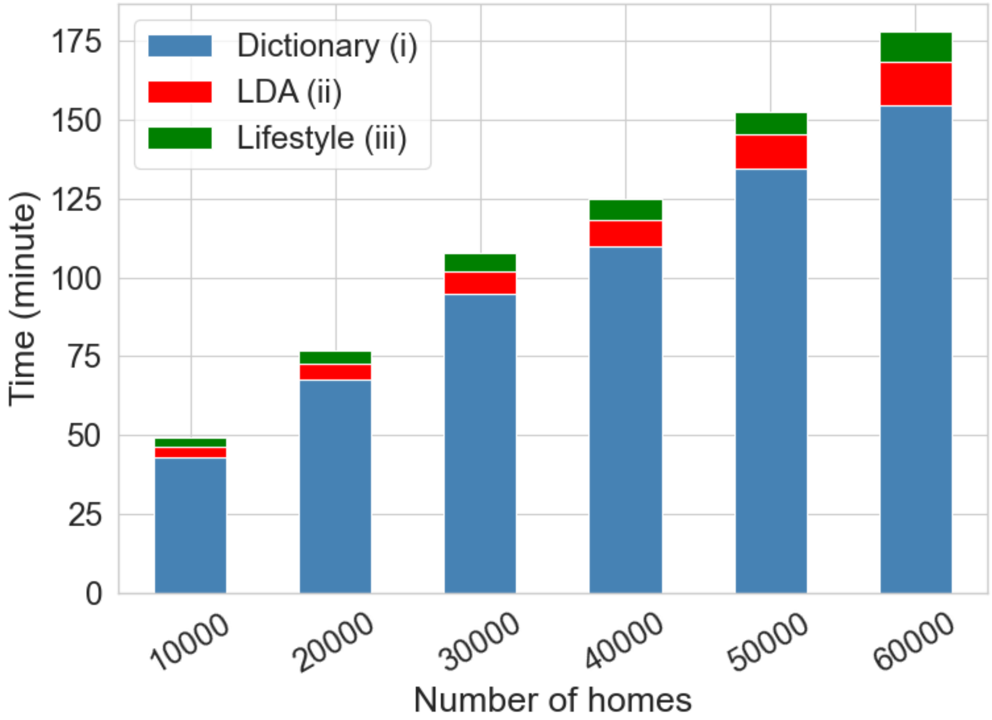
Clustering Models: Generate a dictionary of load shapes




Counting (Map-reduce, e.g., Hadoop): Frequency of shape occurrence



# Scalability




<= 3 hours

# Frequent Questions

Applying other datasets?

Flexible to new / different data source, possible tweaks include building dictionary or frequency counts, and then down streaming analysis

Applying to limited observations?

If some good/reliable priors are obtained (e.g., from historical observations), we can tag the lifestyle via classification (assuming classifier is robust)

Harnessing with socio-demographic information?

Very promising extension and a great augmentation to customer targeting

# Break

# [Appendix] Privacy: perturbation based approach

## Preprocessing/Aggregation

- have correlation between sensitive labels and data
- aggregation alters the utility of the data
- demand shaping is limited by device constraints

## Noise injection

- Differential privacy (DP) → too strict and excessively alter the utility of the data
- Generative adversarial privacy (GAP) → balances privacy and data utility

Dwork 2006; Sweeney 2002; Sankar et al 2013; Duchi et al. 2013;

# [Appendix] Traditional approaches of preserving privacy (in smart grid)

- Anonymity-based [Aitzhan, et.al 2018, Li, et. al 2020]
  - vulnerable to linkage attacks
- Cryptosystem-based
  - high computation and communication overhead [Hu et.al 2016]
- Perturbation-based
  - Low computation and communication cost
  - Statistical privacy (e.g. differential privacy, mutual information,...)
    - Preprocessing, Aggregation, noise perturbation

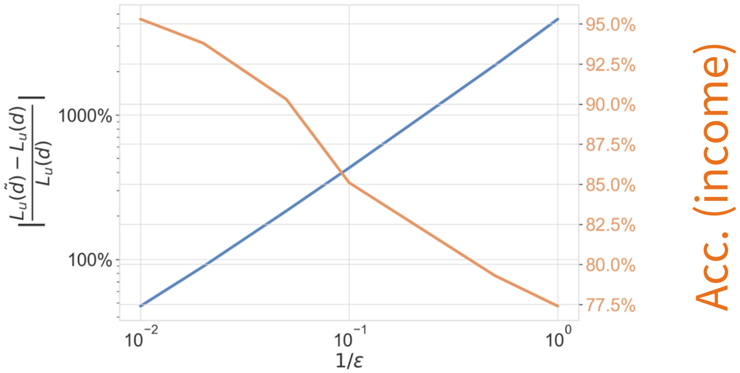
      [Dwork 2006; Sweeney 2002; Sankar et al 2013; Duchi et al. 2013]

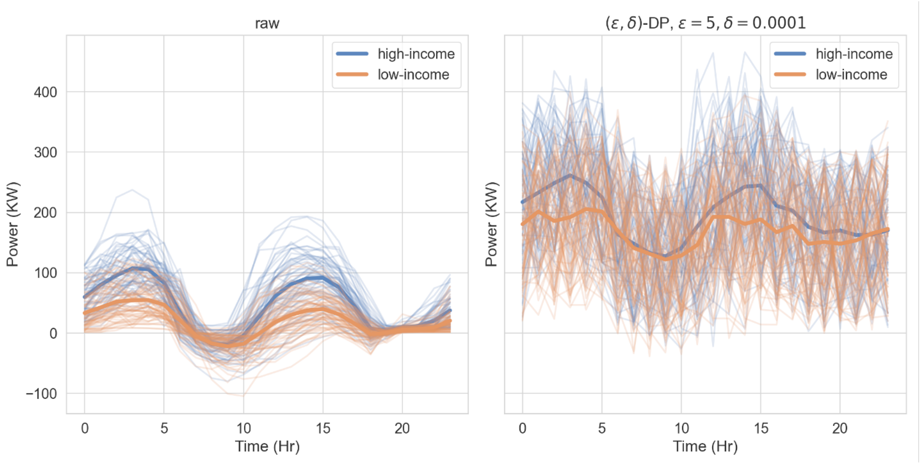
Traditional high computation cost,
Perturbation based method, privacy and performance

# Degraded performance in (standard) DP

The Gaussian mechanism: $(\varepsilon,\,\delta) - DP$

$$\tilde{d} = f(d) + Z \quad \text{where} \quad Z \sim \mathcal{N}\left(0,\, \frac{2\log(1.25/\delta)L_f^2}{\varepsilon^2}\right)$$

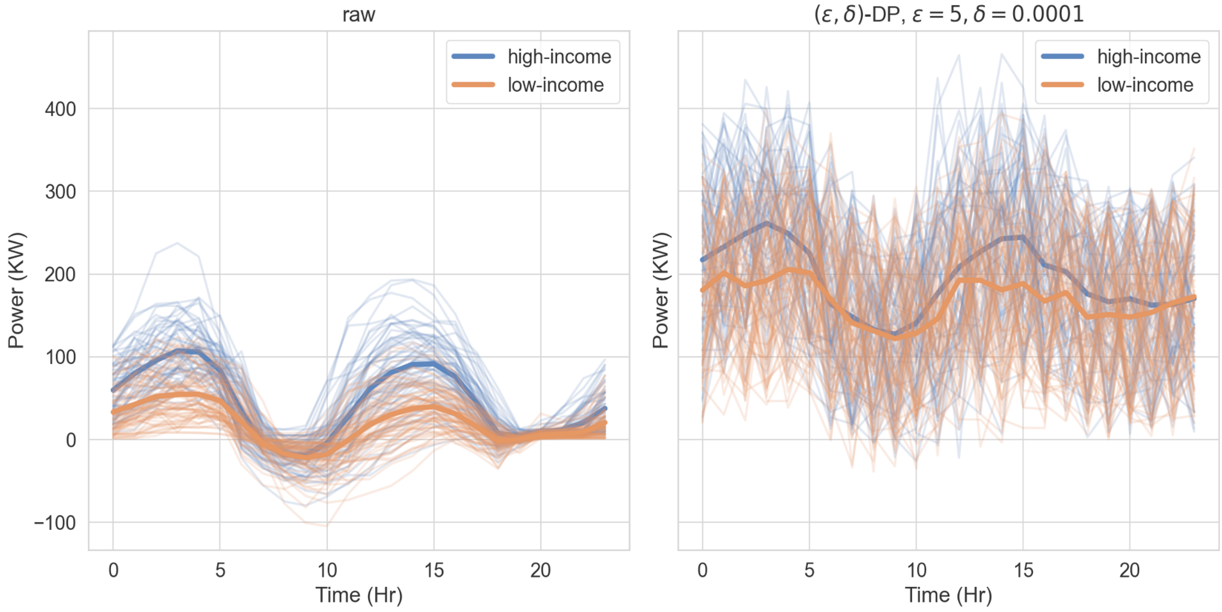



Delete (keep motivation into 2 slides)

[Dwork & Roth 2013]

# Example of using differential privacy (DP)




**Drawbacks:**

1. shift the raw scale of the data;
2. perturb almost every coordinate;
3. deteriorate the value of the data to run DER control;

# Battery control (TBD in appendix)

1. Deterministic battery charging problem:

$$\min_{\mathbf{x}} \underbrace{\mathbf{p}^\top(\mathbf{x}_{in} - \mathbf{x}_{out} + \mathbf{d})_+}_{\text{Cost of electricity}} + \underbrace{\beta_1||\mathbf{x}_{in}||_2^2 + \beta_2||\mathbf{x}_{out}||_2^2 + \beta_3||\mathbf{x}_s - \alpha B||_2^2}_{\text{Battery Operating Cost}}$$

$$\boxed{\begin{aligned} \mathbf{x}_s(j+1) &= \mathbf{x}_s(j) - \frac{1}{\eta_{out}}\mathbf{x}_{out}(j) + \eta_{in}\mathbf{x}_{in}(j) \\ \mathbf{x}_s(1) &= B_{init} \\ 0 \leq \mathbf{x}_{in} &\leq c_{in} \\ 0 \leq \mathbf{x}_{out} &\leq c_{out} \\ 0 \leq \mathbf{x}_s &\leq B \end{aligned}} \quad \forall j \in [H]$$

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_{in} \\ \mathbf{x}_{out} \\ \mathbf{x}_s \end{bmatrix}$$

Storage Operating State

$$\mathbf{x} \in \mathcal{X}$$

Storage Constraints

# Battery control

System planner (demand is stochastic):

$$\min \mathcal{L}_u(\mathbf{x}, \mathbf{d}) := \min_{\mathbf{x}} \mathbf{E}_{\mathbf{d}\sim P}\left[\mathbf{p}^\top(\mathbf{x}_{in} - \mathbf{x}_{out} + \mathbf{d})_+\right] + \beta_1||\mathbf{x}_{in}||_2^2$$
$$+ \beta_2||\mathbf{x}_{out}||_2^2 + \beta_3||\mathbf{x}_s - \alpha B||_2^2$$
$$\mathbf{x} \in \mathcal{X}$$

Adversary:

$$\min_{\psi} \mathcal{L}_a\left(f_\psi(\mathbf{d}), y\right) := \min_{\psi}\left\{-y\log(f_\psi(\mathbf{d})) - (1-y)\log\left(1 - f_\psi(\mathbf{d})\right)\right\}$$

Where y is the private attributes such as households income or sq-ft.
$\psi$ is adversarial classifier.

# Minimax optimization

$$\min_{\mathbf{G}} \underbrace{\mathcal{L}_u\left(\tilde{\mathbf{x}}^*(\tilde{\mathbf{d}}), \mathbf{d}\right)}_{\text{Utility loss}} + \lambda_a \max_{\mathbf{G}} \min_{\psi} \underbrace{\mathcal{L}_a(f_\psi(\tilde{\mathbf{d}}), \mathbf{y})}_{\text{Adversarial Loss}}$$

$$\left.\text{s.t. } \tilde{\mathbf{d}} = \mathbf{d} + \mathbf{G}\begin{bmatrix}\varepsilon \\ \mathbf{y}\end{bmatrix}, \varepsilon \sim \mathcal{N}(0, \mathbf{I})\right\} \text{Linear Filter}$$

$$\tilde{\mathbf{x}}^*(\tilde{\mathbf{d}}) = \arg\min_{\mathbf{x} \in \mathcal{X}} \mathcal{L}_u(\mathbf{x}, \tilde{\mathbf{d}}),$$

# Filter convergence

Theorem:

If $||\nabla L_a|| \leq \delta$ , $||\nabla L_u||_2 \leq \delta$ , $L_a$ (adversarial loss) and $L_u$ (utility loss) are convex,

As $k \to \infty$ we have

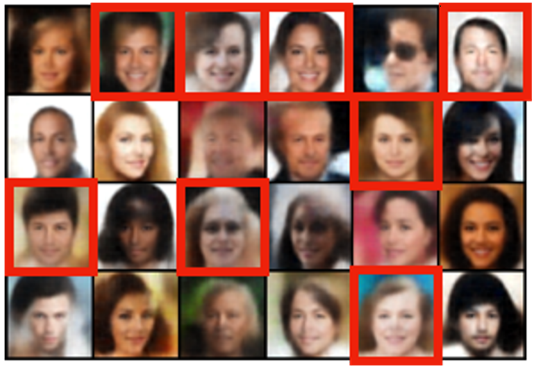
$$(L_a(G_k) - L_a^*) + (L_u(G_k) - L_u^*) \leq \frac{||G_1 - G^*||_2^2 + \sum_{k=1}^{K} \left(\eta^{(k)}\right)^2 \delta^2}{2 \sum_{k=1}^{K} \eta^{(k)}}$$

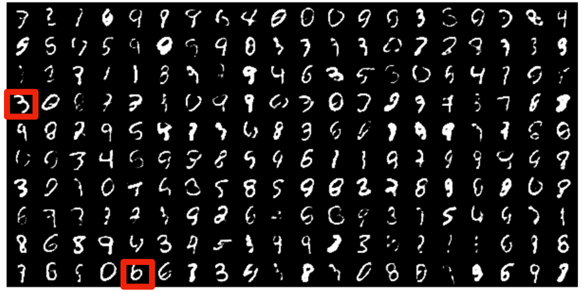
using gradient backpropagation method (e.g. SGD, Adam, etc) can converge to the optimum

# Applications of privacy-preserved DER control

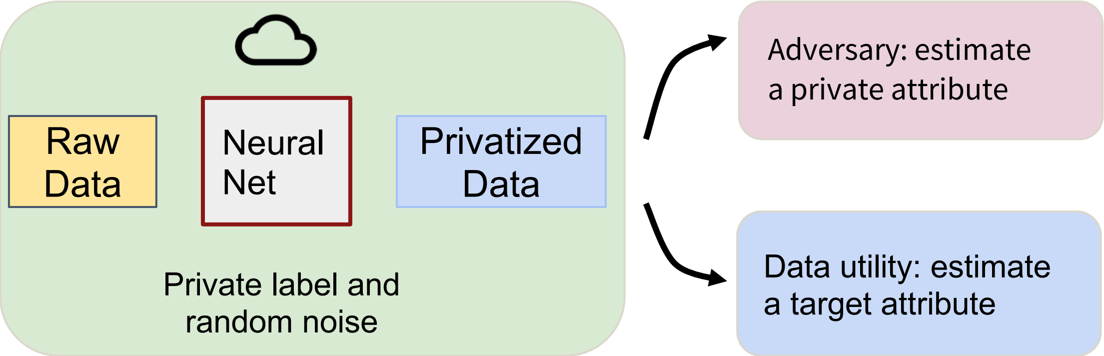
1. The linear filter is trained through the minimax optimization using an anonymous batch of private data
2. Users take the filter weights to calculate the noise injection and privatize their own data locally.
   a. Privatizing data only requires a single matrix product and can easily be run on small local hardware
3. Privatized data is sent to DER operator for use in control optimization

# Extended applications

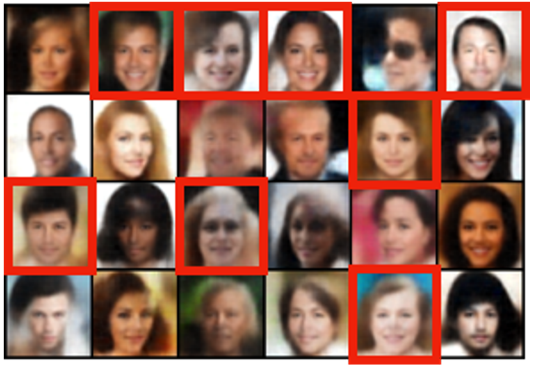




# Other applications




We use VAE + Minimax optimization



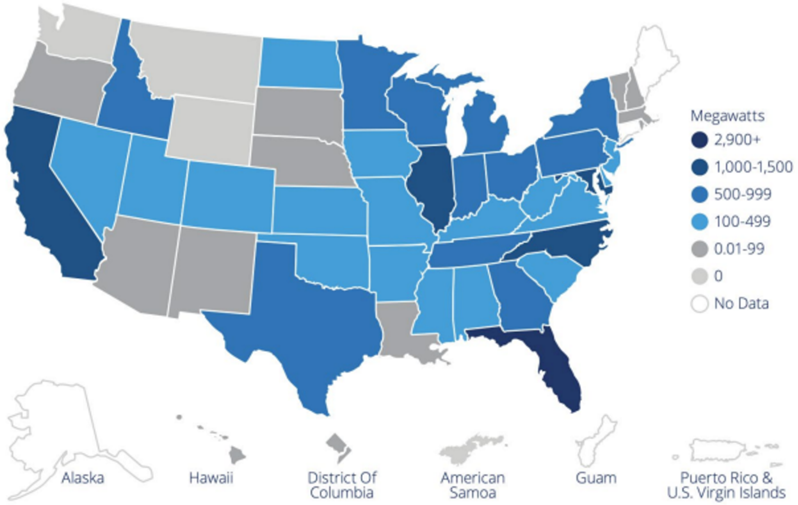
[Chen, Navidi, Ermon, and Rajagopal, ICLR 2019 SafeML]

# Break

# [Appendix] Why demand response ?

Demand response:

- large enrolled capacity
- Popular among states




Source: Smart Electric Power Alliance, 2019. N=190 Utility Survey participants.

**Top States:**

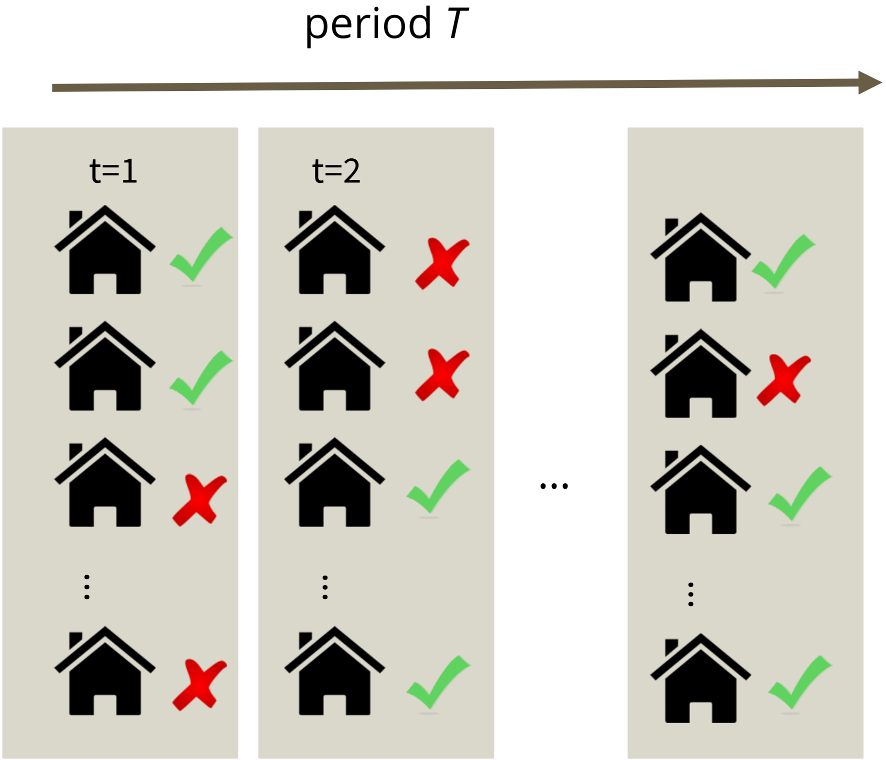
1. Florida - 2,911.4 MW
2. California - 1,335.4 MW
3. North Carolina - 1,319.8 MW
4. Maryland - 1,212.8 MW
5. Illinois - 1,146.5 MW

Certain class of demand side management

*Source: "2019 Utility Demand Response Market Snapshot ", presented by Smart Electric Power Alliance (SEPA) at 40th PLMA Conference, November 2019.*

Justification of selection

# Online optimization




- fairly select homes
- achieve the limited regret

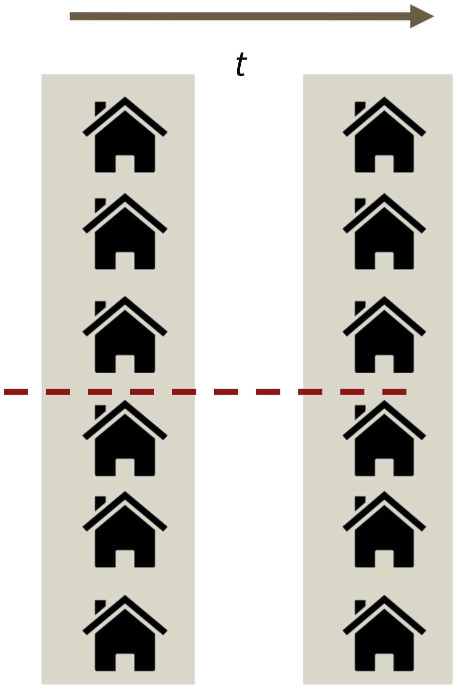
regret = $R_\pi - R_{opt}$

$$R_\pi = \mathbf{E}_{(d,c)\sim q}\left[\sum_{t=1}^{T} f_{d,t}(x_t) \,\middle|\, \{d_\tau, c_\tau, x_\tau\}_{\tau=1\ldots t}\right]$$

$$R_{opt} = \mathbf{E}_{(d,c)\sim q}\left[\min_{x\in S_x} \sum_{t=1}^{T} f_{d,t}(x_t)\right]$$

# Group Fairness quantification

equal opportunity




$J$

$$\frac{1}{TJ}\sum_{t=1}^{T}\sum_{j=1}^{J} x_{j,t}$$

$K$-$J$

$$\frac{1}{T(K-J)}\sum_{t=1}^{T}\sum_{j=J+1}^{K} x_{j,t}$$

$$\frac{4}{5} \leq \frac{P(X=1 \mid A=1)}{P(X=1 \mid A=0)} \leq 1$$

$$-\frac{1}{TJ}\sum_{t=1}^{T}\sum_{j=1}^{J} x_{j,t} + \frac{4}{5}\frac{1}{T(K-J)}\sum_{t=1}^{T}\sum_{j=J+1}^{K} x_{j,t} \leq 0$$

$$\frac{1}{TJ}\sum_{t=1}^{T}\sum_{j=1}^{J} x_{j,t} - \frac{1}{T(K-J)}\sum_{t=1}^{T}\sum_{j=J+1}^{K} x_{j,t} \leq 0$$

# Formulate a general online optimization

$$\min_{x_t} \mathbf{E}\left[\sum_{t=1}^{T} \overbrace{\left(L_t - d_t^\top x_t\right)_+}^{f_{d,t}(x_t)}\right] \longleftarrow \min_{x_t} \mathbf{E}\left[\sum_{t=1}^{T} \left(L_t - d_t^\top x_t\right)_+ + \phi\left(\frac{1}{T}\sum_{t=1}^{T} C_t x_t\right)\right]$$

$$\sum_{t=1}^{T} C_t^\top x_t \leq \mathbf{b}T$$

$$x_t \in \{0,\ 1\}^K$$

- Budget constraint
- Fairness constraints

# Formulate a general online optimization

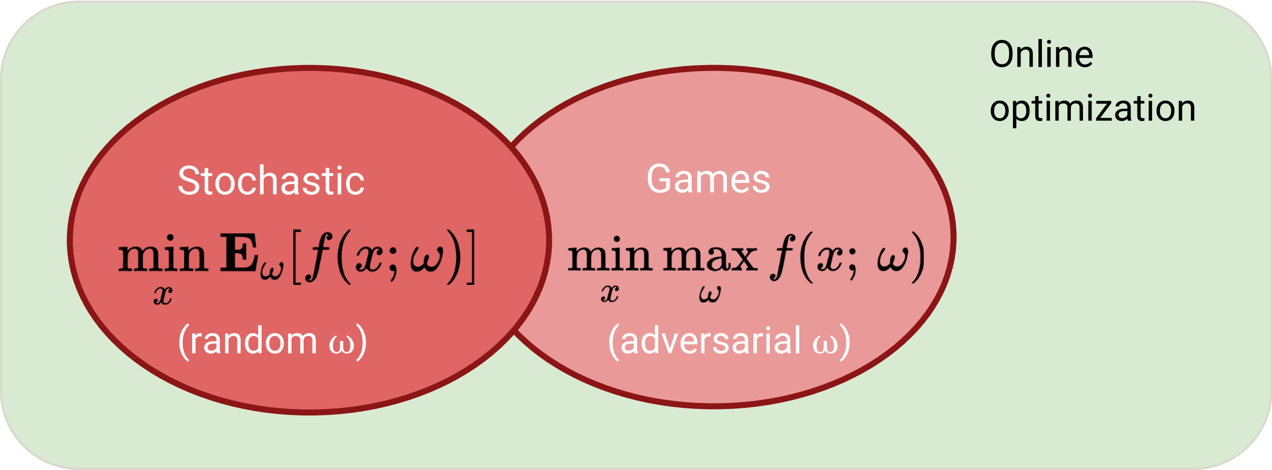
$$\min_{x_t} \mathbf{E}\left[\sum_{t=1}^{T} \overbrace{\left(L_t - d_t^\top x_t\right)_+}^{f_{d,t}(x_t)} + \phi\left(\frac{1}{T}\sum_{t=1}^{T} C_t x_t\right)\right]$$

$$\sum_{t=1}^{T} C_t^\top x_t \leq \mathbf{b}T$$

$$x_t \in \{0,\ 1\}^K$$

- Budget constraint
- Fairness constraints

# Aalternative view: Online learning/optimization



# Fairness: Related work

Machine learning:

- Demographic parity
- Equalized opportunities
- Equalized odds

Communication networks:

- Fairness in channel
- Resource allocation

# Algorithms: PDG-RF

Primal-Dual Gradient method with Regularization and Fairness:

Settings: all potential demand reduction can be observed

- introduce an auxiliary variable;
- derive the lagrangian with its multiplier;
- decouple the primal and dual;
- use the gradient updates

**Assumption 2.** *We assume :*

(i) $||\boldsymbol{y}||_2 \le ||\boldsymbol{b}||_2 \le B$ *for all* $\boldsymbol{y}$,
(ii) $||C\boldsymbol{x}||_2 \le M$ *for all* $\boldsymbol{x}$,
(iii) $||\boldsymbol{\lambda}||_2 \le \omega$ *for all* $\boldsymbol{\lambda} \in \mathcal{S}_\lambda$.

Theorem:

$$\boldsymbol{Reg}_T \le \frac{VM}{b_{min}} + \sqrt{\frac{T(B^2+M^2)}{2}(2\omega^2 + (\frac{V}{b_{min}})^2)}, + \sqrt{\frac{T(B^2+M^2)}{2}}$$

*where* $V = \phi_{max} + \omega\sqrt{B^2+M^2}$.

# PDG-RF Alg.

**Algorithm 1: Primal-Dual Gradient updates with Regularization and Fairness (PDG-RF)**

**input:** Number of homes $K$, initial guess of dual $\lambda_1$, horizon $T$, per-round budget $\boldsymbol{b}$, some demand and cost distribution $q$, initial total budget of resources $\widetilde{B}_1 = T\boldsymbol{b}$, learning rate $\eta$.

**for** $1 \leq t \leq T$ **do**

sample $(\boldsymbol{d}_t, \boldsymbol{C}_t) \sim q$

primal updates:

$$\hat{\boldsymbol{x}}_t = \arg\min_{\boldsymbol{x} \in \mathcal{S}_x} \left( f_d(\boldsymbol{x}) - \boldsymbol{\lambda}_t^\top \boldsymbol{C}_t \boldsymbol{x} \right) \tag{23}$$

$$\hat{\boldsymbol{x}}_t = \varphi_r(\hat{\boldsymbol{x}}_t) \quad \text{[randomized rounding]} \tag{24}$$

$$\boldsymbol{x}_t = \begin{cases} \hat{\boldsymbol{x}}_t & \text{if } \widetilde{B}_t - \boldsymbol{C}_t \hat{\boldsymbol{x}}_t \geq 0 \\ 0 & \text{otherwise picking violated index} \end{cases} \tag{25}$$

$$\boldsymbol{y}_t = \arg\min_{\boldsymbol{y} \leq \boldsymbol{b}} \left( \phi(\boldsymbol{y}) + \boldsymbol{\lambda}_t^\top \boldsymbol{y} \right) \tag{26}$$

$$\widetilde{B}_{t+1} = \widetilde{B}_t - \boldsymbol{C}_t \boldsymbol{x}_t \tag{27}$$

dual updates:

$$g_t = \nabla_\lambda D_q(\lambda) = -\boldsymbol{C}_t \hat{\boldsymbol{x}}_t + \boldsymbol{y} \quad \text{[gradient of dual]} \tag{28}$$

$$\lambda_{t+1} = \arg\min_{\lambda \in \mathcal{S}_\lambda} g_t^\top \lambda + \frac{1}{\eta} D_\psi(\lambda, \lambda_t) \tag{29}$$

**end**

where

$$\varphi_r(x_{j,t}) = \begin{cases} 1 & \text{If } \texttt{rand}(0{,}1) \geq x_{j,t} \\ 0 & \text{Otherwise,} \end{cases} \tag{30}$$

and $D_\psi$ is the Bregman divergence:

$$D_\psi(x, y) = \psi(x) - [\psi(y) + \nabla\psi(y)^\top (x - y)] \tag{31}$$

for some convex function $\psi$.

# Algorithms: UCB-RF

We consider the multi-armed bandit setting and use ucb-type approach.

Settings: demand reduction are observed when calling specific homes by system operator.

- Construct a UCB of demand reduction
- Construct a LCB of payment cost
- Solve optimization problem with UCB(d) and LCB(c)

$$\hat{d}_{j,t} = \left[ \overline{d_{j,t}} + 2\left( \frac{\alpha}{\sum_{\tau=1}^{t-1} x_{j,\tau}} + \sqrt{\frac{\alpha \overline{d_{j,t}}}{\sum_{\tau=1}^{t-1} x_{j,\tau}}} \right) \right] \wedge 1$$

$\wedge$ is min{} operator

# UCB-RF Regret & Alg

Thm. With a prob. (1 - $\delta$), the regret is bounded as

$$\boldsymbol{Reg}_T \leq O\Big(\sqrt{\log(\frac{KT}{\delta})KR_{opt}} + \log(\frac{KT}{\delta})K\Big),$$

where K is the number of customers.

**input:** Number of homes $K$, horizon $T$, per-round budget $\boldsymbol{b}$.

When $t = 0$, we initialize the $c_{j,0}$ and call all homes once to have $d_{j,0}, \forall j = 1, \ldots, K$.

**for** $1 \leq t \leq T$ **do**

estimate UCB of $\hat{\boldsymbol{d}}$:

**for** $1 \leq j \leq K$ **do**

$\hat{d}_{j,t} = \Big[\bar{d}_{j,t} + 2r\Big(\bar{d}_j, \sum_{\tau=1}^{t-1} x_{j,\tau} + 1\Big)\Big] \wedge 1$ [UCB of $d$]

$\hat{c}_{j,t} = \Big[\bar{c}_{j,t} - 2r\Big(\bar{c}_j, \sum_{\tau=1}^{t-1} x_{j,\tau} + 1\Big)\Big] \vee 0$ [LCB of $c$]

**end**

solve the following optimization and obtain the solution $\hat{\boldsymbol{x}}_t$:

$$\min_{\boldsymbol{x}_t} \Big(L_t - \hat{\boldsymbol{d}}_t^\top \boldsymbol{x}_t\Big)_+ + \phi(\frac{1}{t}\sum_{\tau=1}^{t} \boldsymbol{x}_\tau) \tag{36}$$

$$s.t. \sum_{\tau=1}^{t} \hat{\boldsymbol{C}}_\tau \boldsymbol{x}_\tau \leq \boldsymbol{b}t \tag{37}$$

pulling homes with probability $\boldsymbol{x}_t$ (with rescaling).

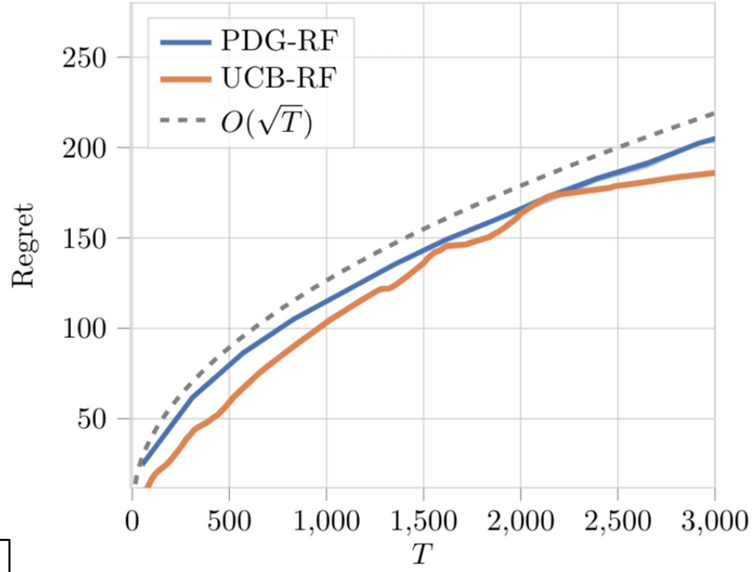
update empirical mean $\bar{d}_j$ and $\bar{c}_j$ accordingly

**end**

# Case2: simulating DR in a community

Simulation setup:

1. Demand data:
   a. 200 homes from NREL (10% reduction of peak power)
2. Curtailment target:
   a. CAISO 2021 Summer load (rescale)
3. Unit cost and budget is 20
4. Vulnerable group:
   a. 50% users



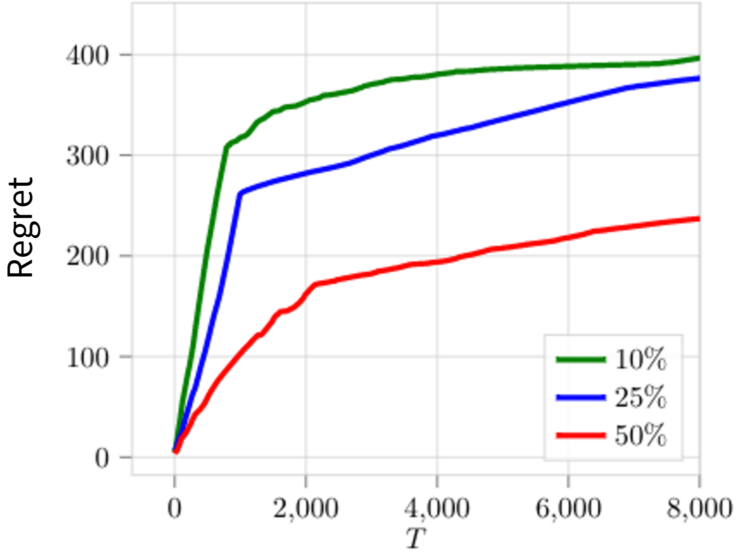

**Findings:**

- Both methods achieve sublinear regret
- UCB-RF initially behave like sqrt(T)

# Convergence with different group splits

Simulation setup:

- UCB-RF method
- The vulnerable customers account for 10/25/50% of the population




**Findings:**

- less proportion of the vulnerable group has larger regret at beginning but converge quickly after certain rounds