Robust Data-driven Pricing Scheme Design

The Problem: It is commonly believed that customized pricing schemes can best exploit the flexibility in the demand side. However, due to the large volume of customers in the electricity sector, such task is simply too overwhelming for the system operator.

Conventional Wisdom: Under the assumption that most users are homogenous, the system operator often penalizes the large consumers for high peak in the aggregate system load. This is the conventional way of "customized pricing". However, simple k-means clustering suggests this homogenous user assumption is invalid (as shown in Fig. 1).

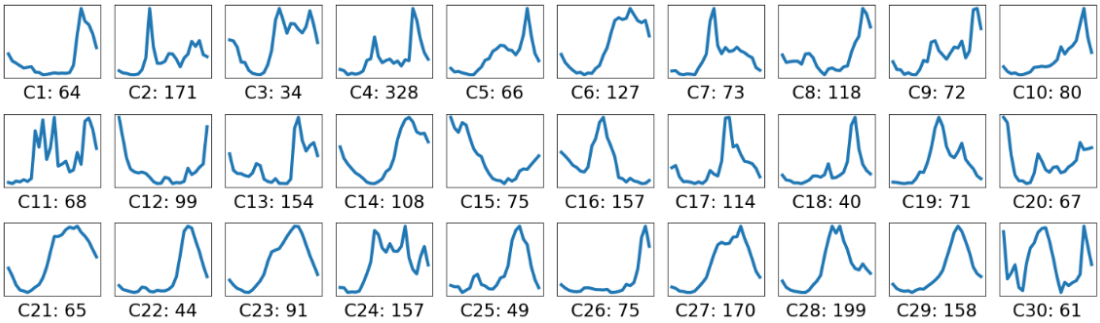

Fig. 1 Demand profile of each cluster center, the caption $(C_n{:}N)$ of each sub figure implies that it is the central profile of cluster n, with N users in the cluster (x-axis: time; y-axis: normalized load profile).

Identifying the heterogeneity in the user load profiles, Yu *et al.* proposed an index, MCI, to measure each user's system impact using l_1-norm [1]. MCI implies that the customized pricing should be uniquely determined by the shape of each user's load profile, instead of the total energy consumption. This serves as the basis that the customized pricing scheme can be implemented by k-means clustering based on users' load profiles.

Our Idea: The clustering results based on load profiles may lead to loopholes in the market, since the users in the same cluster do not exactly share the same load profile but are charged at the same rate. Such loopholes allow strategic users to bypass to other clusters for a lower rate with only minimal modification of their profiles. We conduct the vulnerability analysis for this pricing scheme and introduce a parametric definition to characterize the strategic behaviors [2]. Using real data, we highlight that under this pricing scheme, strategic users do exist. Figure 2 shows by only strategically adjusting 1% of demand, how the users may disguise and change their clusters. The arrows represent these changes in the clusters. The darker the cluster, the more stable the cluster is, which implies it is more difficult for the users in the cluster to disguise.

We then examine why the aforementioned pricing scheme is vulnerable. The first-cut answer is that the pricing scheme is not robust. A deeper reason is due to the selection of an indirect clustering criteria. Using the concept of end-to-end machine learning, we submit that the optimal k-means clustering should be directly based on MCI, instead of load profiles. Figure 3 shows while the clustering based on MCI only produces 24 clusters, clustering based on load profiles needs to

produce at least the same amount of clusters as the number of red colored rectangles.

We further prove that for MCI based *k*-means clustering algorithm, as long as maintaining the local property for each cluster, guarantees the global robustness [3]. Moreover, since this *k*-means clustering is based on a single quantity (MCI), an efficient greedy algorithm is enough to offer the optimal results!

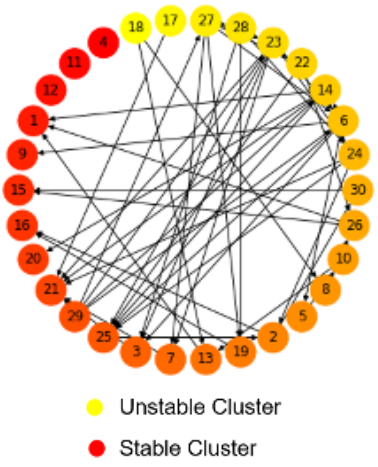

Fig. 2 Disguise trajectories

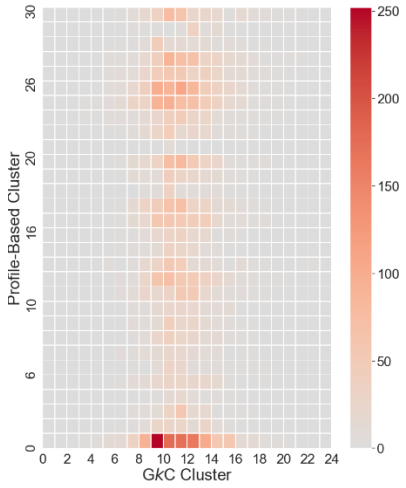

Fig. 3 Relationship Between two clustering methods.

Future Work: We intend to use the Stackelberg game formulation to better understand the dynamics between system operator and the users. It is also interesting to examine the temporal characteristics of user behaviors over a longer time span.

References

[1] Yang Yu, Guangyi Liu, Wendong Zhu, Fei Wang, et al. Good consumer or bad consumer: Economic information revealed from demand profiles, IEEE Transactions on Smart Grid, vol. 9, no. 3, pp. 2347-2358, May 2018

[2] Jingshi Cui, Haoxiang Wang, **Chenye Wu**, Yang Yu*, "Vulnerable Analysis for Data Driven Pricing Schemes", in submission to IEEE PES General Meeting 2020, https://arxiv.org/abs/1911.07453, 2019

[3] Jingshi Cui, Haoxiang Wang, **Chenye Wu***, Yang Yu,"Robust Data-driven Profile-based Pricing Schemes", in submission to IEEE Transactions on Smart Grid, Initial Submission: Nov. 2019.