Design, Benchmarking and Graphical Lasso based Explainability Analysis of an Energy Game-Theoretic Framework

Hari Prasanna Das [1] **Ioannis C. Konstantakopoulos** [1] **Aummul Baneen Manasawala** [2] **Tanya Veeravalli** [1] **Huihan Liu** [1] **Costas J. Spanos** [1]

Abstract

Energy use in buildings account for approximately half of global electricity consumption and a significant amount of CO_2 emissions. The occupants of a building typically lack the independent motivation necessary to optimize their energy usage. In this paper, we propose a novel energy game-theoretic framework for smart building which incorporates human-in-the-loop modeling by creating an interface to allow interaction with occupants and potentially incentivize energy efficient behavior. We present open-sourced dataset and benchmarked results for forecasting of energy resource usage patterns by leveraging classical machine learning and deep learning methods including deep bi-directional recurrent neural networks. Finally, we use graphical lasso to demonstrate the explainable nature on human decision making towards energy usage inherent in the dataset.

1. Introduction and Related Work

Buildings, both residential and commercial, account for more than 50% of global electricity consumption and are also responsible for 40% of worldwide CO_2 emissions (Allouhi et al., 2015). In efforts to improve energy efficiency in buildings, researchers and industry leaders have attempted to implement control and automation approaches alongside techniques like incentive design and price adjustment to more effectively regulate the energy usage (Aswani & Tomlin, 2012; Ratliff et al., 2014; Liu et al., 2019; Zou et al., 2019b). But, the occupants of a building typically lack the independent motivation necessary to optimize their energy usage and play a key role in the control of smart building infrastructure (Konstantakopoulos, 2018). So, there is a need for scalable and robust frameworks that can efficiently coordinate and control building energy resource usage in the presence of confounding dynamics such as human behavior. We present an energy game-theoretic framework aimed at incentivizing occupants to modify their behavior in a competitive game setting so that the overall energy consumption in the building is reduced. Such frameworks have been successful in many different areas such as transportation (Qin et al., 2017), medical industry (Bestick et al., 2013) etc. Our framework can also be integrated with the electricity grid (Figure 1) to facilitate the adoption of more dynamic protocols for demand response (Shariatzadeh et al., 2015). We also benchmark the results for forecasting of energy resource usage patterns by leveraging classical machine learning and deep learning methods. To make sure the data captures explainable human decision making behavior for energy usage, we perform feature correlation study using graphical lasso.

[1]Department of Electrical Engineering and Computer Sciences, UC Berkeley [2]Department of Industrial Engineering and Operations Research, UC Berkeley. Correspondence to: Hari Prasanna Das <hpdas@berkeley.edu>.

2. Design of Energy Social Game

2.1. Energy Social Game Experiment

In this section, we introduce the design and implementation of a large-scale networked energy game-theoretic framework through the utilization of cutting-edge Internet of Things (IoT) sensors, implemented with participation of dorm room occupants at an university residential housing.

The back-end of our game-theoretic framework included an

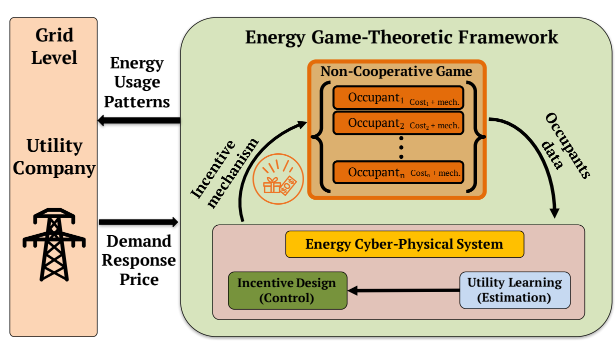

Figure 1. Interplay between electric grid and proposed framework

array of IoT sensors and a structure to assign points to the players based on their performance in the game. In each dorm room, we installed sensors which leveraged several indoor metrics like indoor illuminance, humidity, temperature, and vibrations to capture the room's energy resource (ceiling light, desk light and ceiling fan) usage, with a sampling interval of up to one minute. The players were rewarded with points based on how energy efficient their daily usage is in comparison to their peers and their usage before the social game was deployed. The baseline past usage data was gathered by monitoring occupant energy usage for approximately one month before the introduction of the game. We employed a lottery mechanism consisting of gift cards to incentivize occupants, where the probability of winning was proportional to the players points in the game, given by:

$$\hat{p}_i^d(b_i, u_i) = s_i \frac{b_i - u_i^d}{b_i} \tag{1}$$

where $\hat{p}_i^d$ is the points earned and u_i^d is the usage on day d for resource i. b_i is the resource's baseline and s_i is a points booster for inflating the points as a process of framing (Tversky & Kahneman, 1981). This process of framing is routinely used in rewards programs for credit cards among many other point-based programs. We use a discrete choice model as a core abstraction for describing occupant actions related to their dorm room resources (Konstantakopoulos et al., 2019; 2018; 2016; 2017).

The front-end of our framework included a web portal (Figure 2) as the graphical user interface to report the occupants about real-time status of the devices, their accumulated daily usage and the % of allowed baseline being used, by hovering above the utilization bars. In order to boost participation, we introduced a randomly appearing coin with the purpose of incentivizing occupants and reminding them to view their usage and optimize it. In order to have impact by visualizations, each users background in the web portal changes based on their energy efficiency, with pictures of rain forest for high and desert scenes for low energy efficient user. Detailed experiment design has been included in Konstantakopoulos et al. (2019).

2.2. Dataset Description and Open-Sourcing

The energy social game dataset so obtained consisted of per-minute time-stamped reading of each resource (desk light, ceiling light and fan) status, accumulated usage (in min/day), resource baseline, points (both from game and surveys), rank, number of visits to the portal and external weather metrics like humidity, temperature and solar radiation. Following this, we propose a pooling & picking scheme to enlarge the feature space by applying a Minimum Redundancy and Maximum Relevance (mRMR) (Peng et al., 2005) feature selection procedure to identify useful features for our predictive algorithms, such as dummy features (using one-hot encoding) which includes weekends, breaks, holidays, midterm and final exam periods, and resource features which includes daily % of resource usage. The dataset has been open-sourced[1] after proper benchmarking.

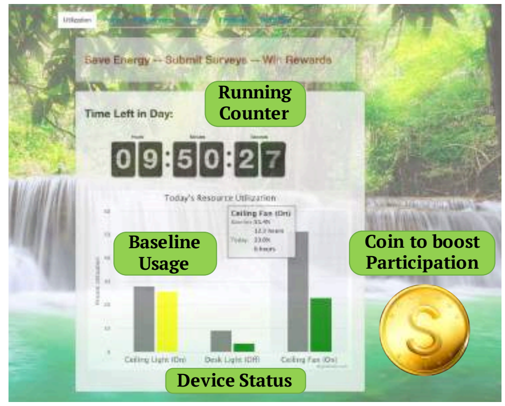

Figure 2. Illustration of web-portal displaying real-time energy resource usage by the players. The background is a picture of rain forest corresponding to a more energy efficient player.

3. Benchmarking of Social Game Dataset

In this section, we will explore the benchmarking of the social game dataset using classical and deep learning methods for accurate energy resource usage forecasts (utility estimation). Since human interaction data in general is imbalanced, we use the Synthetic Minority Over-Sampling (SMOTE) (Chawla et al., 2002) technique for providing balanced data sets for each energy resources.

3.1. Machine Learning framework for Modelling

3.1.1. Classical machine learning models

We train several classifiers as a part of the utility estimation pipeline. We propose models of logistic regression, logistic regression with l_1 penalization (Lasso), linear discriminant analysis (LDA), support vector machine and random forest classifiers. We use the Area Under Curve (AUC) (Majnik & Bosnic, 2013) as our performance metric and perform 5-fold cross validation combined with the AUC.

3.1.2. Deep Neural Networks

We also utilize the potential of deep neural networks (DNN) for utility estimation that allows us to significantly improve the accuracy. In a non-cooperative energy game setting, DNNs work as powerful models that can generalize the core model by increasing capacity for predicting agent behavior. Our proposed DNN model includes exponential linear units

[1]For open sourced social game dataset and demonstrations, please visit `https://smartntu.eecs.berkeley.edu`

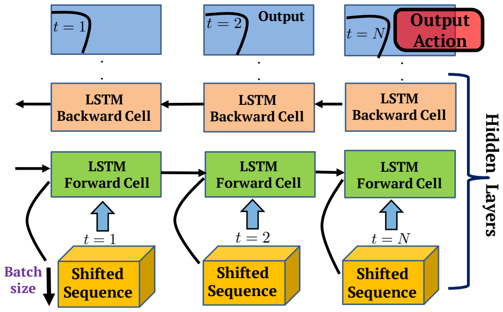

Figure 3. Architecture of Deep Bi-directional Neural Network

"Step-ahead" / "Sensor-free"	Ceiling Fan	Ceiling Light	Desk Light
Logistic regression	0.83 / 0.65	0.78 / 0.61	0.78 / 0.68
Penalized l_1 Logistic regression	0.80 / 0.65	0.77 / 0.56	0.78 / 0.64
Bagged Logistic regression	0.84 / 0.66	0.80 / 0.59	0.79 / 0.68
LDA	0.81 / 0.65	0.78 / 0.58	0.74 / 0.68
K-NN	0.76 / 0.53	0.77 / 0.56	0.74 / 0.55
Support Vector Machine	0.82 / 0.65	0.78 / 0.60	0.76 / 0.68
Random Forest	0.91 / 0.60	0.78 / 0.59	0.98 / 0.63
Deep Neural Network	0.80 / 0.55	0.76 / 0.60	0.78 / 0.64
Deep Bi-directional RNN	**0.97 / 0.71**	**0.85 / 0.66**	**0.99 / 0.76**

(a) AUC Scores for various machine learning models

	Weekday				Weekend			
Device	Before	After	p-value	Δ %	Before	After	p-value	Δ %
Ceiling Light	417.5	393.9	0.02	5.6	412.3	257.5	0	37.6
Desk Light	402.2	157.5	0	60.8	517.6	123.3	0	76.2
Ceiling Fan	663.5	537.6	0	19.0	847.1	407.0	0	51.9

(b) Energy Savings achieved in the social game

Table 1. AUC Score and Energy Savings in the Social Game

(ELUs) (Clevert et al., 2015) at each hidden layer. The usage of ELU normally adds additional hyper-parameters as a trade-off for increase in fitting accuracy. The output layer is modeled using sigmoid units. We use cross-entropy loss function and perform training using stochastic gradient descent combined with nesterov optimization. We utilize the method in (He et al., 2015) for initialization, and use batch normalization (Ioffe & Szegedy, 2015) and dropout (Srivastava et al., 2014) for efficient training.

An important challenge for sequential decision-making is the modeling of the dependence of future actions of an agent with the present and previous actions. In particular, an agent naturally tries to co-optimize around a set of discrete choices and gains the higher utility. Therefore, we leverage the time-series DNN models including recurrent neural networks (RNN) and long short term memory cells (LSTM) (Goodfellow et al., 2016) to address the issue of above time dependence. The architecture of our deep bi-directional RNN is illustrated in Figure 3. We use a sliding window of 2 hours with 0.6 dropout rate. Training was done with an exponentially decaying learning rate over 35 epochs.

3.1.3. Graphical Lasso for Explainability

To ensure the collected data incorporates explainable information about human decision making towards energy use in competitive environments, we divide the players into 3 categories based on their rank, as low, medium and high energy efficient (abbreviated as LEE,MEE and HEE) and utilize graphical lasso (GLASSO) (Hastie et al., 2015) to learn feature correlations in each category.

3.2. Experimental Results

We evaluate the performance of utility estimation under two scenarios. The first scenario involves having full information from the installed IoT sensors, called "step-ahead"prediction and second, referred to as "sensor-free", involves use of sensor-free features such as external conditions, frequency of visit to web portal and seasonal dummy variables. The AUC scores using various models are given in Table 1a. From the results, it is clear that deep RNN performs the best in terms of accuracy. In the "sensor-free"results, we have considerable accuracy even with the IoT feed decoupled. In table 1b, we present the energy savings achieved and results from hypothesis testing using energy usage before and after the game. In all of the devices, we have a significant drop in usage between the two periods. The results of feature correlation for HEE and LEE players is given in Figure 4. We observe HEE players showcase predictable behaviors of energy usage with correlation between energy resources. LEE players exhibit heedless behavior towards energy usage with use of desk light in morning. Their usage is affected by external conditions, unlike HEE players. For detailed results, please refer Das et al. (2019).

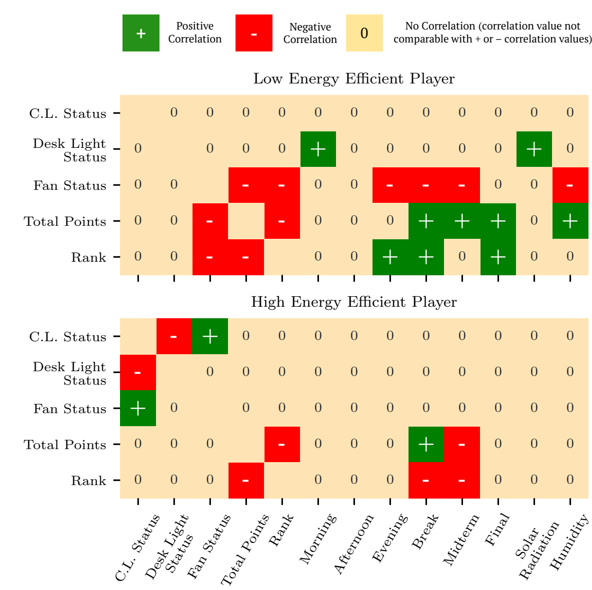

Figure 4. Feature Correlations using Graphical Lasso

4. Conclusion

In this work, we designed an energy game-theoretic framework under a non-cooperative game setting at an university housing. We used classical machine learning models and deep neural networks to benchmark the accuracy of utility estimation. Using graphical lasso, we presented the explainable information inherent in the dataset. Along with state-of-the-art smart building components (e.g. multimodal sensing (Zou et al., 2019a), thermal comfort models (Liu et al., 2018), privacy requirements (Jia et al., 2018)), frameworks such as ours can be utilized to incorporate energy efficient behavior among building occupants in large scale.

References

Allouhi, A., El Fouih, Y., Kousksou, T., Jamil, A., Zeraouli, Y., and Mourad, Y. Energy consumption and efficiency in buildings: current status and future trends. *Journal of Cleaner production*, 109:118–130, 2015.

Aswani, A. and Tomlin, C. Incentive design for efficient building quality of service. In *Proc. 50th Annu. Allerton Conf. Communication, Control, and Computing*, pp. 90–97, 2012. doi: 10.1109/Allerton.2012.6483204.

Bestick, A., Ratliff, L. J., Yan, P., Bajcsy, R., and Sastry, S. S. An inverse correlated equilibrium framework for utility learning in multiplayer, noncooperative settings. In *Proceedings of the 2nd ACM international conference on High confidence networked systems*, pp. 9–16. ACM, 2013.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.

Clevert, D.-A., Unterthiner, T., and Hochreiter, S. Fast and accurate deep network learning by exponential linear units (elus). *arXiv preprint arXiv:1511.07289*, 2015.

Das, H. P., Konstantakopoulos, I. C., Manasawala, A. B., Veeravalli, T., Liu, H., and Spanos, C. J. A novel graphical lasso based approach towards segmentation analysis in energy game-theoretic frameworks. *arXiv preprint arXiv:1910.02217*, 2019.

Goodfellow, I., Bengio, Y., Courville, A., and Bengio, Y. *Deep learning*, volume 1. MIT press Cambridge, 2016.

Hastie, T., Tibshirani, R., and Wainwright, M. *Statistical Learning with Sparsity: The Lasso and Generalizations*. Chapman & Hall/CRC, 2015. ISBN 1498712169, 9781498712163.

He, K., Zhang, X., Ren, S., and Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034, 2015.

Ioffe, S. and Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.

Jia, R., Konstantakopoulos, I. C., Li, B., and Spanos, C. Poisoning attacks on data-driven utility learning in games. In *2018 Annual American Control Conference (ACC)*, pp. 5774–5780. IEEE, 2018.

Konstantakopoulos, I. *Statistical Learning Towards Gamification in Human-Centric Cyber-Physical Systems*. PhD thesis, EECS Department, University of California, Berkeley, Nov 2018. URL http://www2.eecs.berkeley.edu/Pubs/TechRpts/2018/EECS-2018-139.html.

Konstantakopoulos, I. C., Ratliff, L. J., Jin, M., Spanos, C. J., and Sastry, S. S. Inverse modeling of non-cooperative agents via mixture of utilities. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pp. 6327–6334. IEEE, 2016.

Konstantakopoulos, I. C., Ratliff, L. J., Jin, M., and Spanos, C. J. Leveraging correlations in utility learning. In *American Control Conference (ACC), 2017*, pp. 5249–5256. IEEE, 2017.

Konstantakopoulos, I. C., Ratliff, L. J., Jin, M., Sastry, S. S., and Spanos, C. J. A robust utility learning framework via inverse optimization. *IEEE Transactions on Control Systems Technology*, 26(3):954–970, 2018.

Konstantakopoulos, I. C., Das, H. P., Barkan, A. R., He, S., Veeravalli, T., Liu, H., Manasawala, A. B., Lin, Y.-W., and Spanos, C. J. Design, benchmarking and explainability analysis of a game-theoretic framework towards energy efficiency in smart infrastructure. *arXiv preprint arXiv:1910.07899*, 2019.

Liu, S., Jin, M., Das, H., Spanos, C., and Schiavon, S. Personal thermal comfort models based on physiological parameters measured by wearable sensors. *Proceedings of the Windsor Conference*, pp. 431–441, 2018.

Liu, S., Schiavon, S., Das, H. P., Jin, M., and Spanos, C. J. Personal thermal comfort models with wearable sensors. *Building and Environment*, 162:106281, 2019.

Majnik, M. and Bosnic, Z. Roc analysis of classifiers in machine learning: A survey. *Intelligent Data Analysis*, 17:531–558, 2013.

Peng, H., Long, F., and Ding, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 27(8):1226–1238, 2005.

Qin, H., Gao, J., Guan, H., and Chi, H. Estimating heterogeneity of car travelers on mode shifting behavior based on discrete choice models. *Transportation Planning and Technology*, 40(8):914–927, 2017.

Ratliff, L. J., Jin, M., Konstantakopoulos, I. C., Spanos, C., and Sastry, S. S. Social game for building energy efficiency: Incentive design. In *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1011–1018. IEEE, 2014.

Shariatzadeh, F., Mandal, P., and Srivastava, A. K. Demand response for sustainable energy systems: A review, application and implementation strategy. *Renewable and Sustainable Energy Systems*, 45(6):343–350, 2015.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958, 2014.

Tversky, A. and Kahneman, D. The framing of decisions and the psychology of choice. *Science*, 211(4481):453–458, 1981.

Zou, H., Yang, J., Prasanna Das, H., Liu, H., Zhou, Y., and Spanos, C. J. Wifi and vision multimodal learning for accurate and robust device-free human activity recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 0–0, 2019a.

Zou, H., Zhou, Y., Yang, J., Liu, H., Das, H., and Spanos, C. Consensus adversarial domain adaptation. In *AAAI Conference on Artificial Intelligence*, 2019b.