Do Occupants in a Building exhibit patterns in Energy Consumption? Analyzing Clusters in Energy Social Games

Hari Prasanna Das [1] **Ioannis 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. To encourage energy efficient behavior among occupants in a building, energy social games have emerged to be a successful strategy leveraging human-in-the-loop strategy and engaging users in a competitive game with incentives for energy efficient behavior. Prior works involve an incentive design mechanism which is dependent on knowledge of utility functions (energy use behavior) for the users, which is hard to compute when the number of users is high, common in buildings. We propose that the utilities can be grouped to a relatively small number of clusters, which can then be targeted with tailored incentives. Proposed work performs the above segmentation by learning the features leading to human decision making towards energy usage in competitive environment. We propose a graphical lasso based approach with explainable nature for such segmentation, by studying the feature correlations in a real-world energy social game dataset.

1. Introduction and Related Work

Energy consumption of buildings, both residential and commercial, account for approximately 40% of all energy usage in the U.S. (McQuade, 2009). To achieve energy efficiency in buildings, control and automation approaches alongside techniques like incentive design and price adjustment have been implemented (Ratliff, 2015). A building manager, acting as the connection between energy utilities and the end users, can encourage participation and energy-efficient behavior among occupants in many ways. One of the successful methods proposed is a game-theoretic framework (Konstantakopoulos et al., 2019), which creates a friendly competition between occupants, motivating them to individually consider their own energy usage and hopefully, seek to improve it. Although all such frameworks aim to achieve a long term or permanent improvement in the energy usage behaviors among the users, the aim is seldom achieved after the completion of energy game, mostly attributed to the lack of an intelligent and adaptive incentive design process (Ratliff et al., 2014). The incentive design process in prior works is dependant on utility functions of every player in the game, which is hard to compute as buildings involve participation of a large number of energy users, so is often approximated using several estimation techniques (Ratliff, 2015). We propose that the utility/energy usage behavior of the players can be segmented into a relatively small number of clusters, and incentives can be designed to tailor each cluster. We utilize the potential of graphical lasso algorithm (Hastie et al., 2015) to perform such segmentation analysis in energy social games.

[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. Methods

2.1. Energy Social Game Dataset

The dataset used for our work is from a energy social game experiment in a smart residential housing, as introduced in (Konstantakopoulos et al., 2019). The dataset consists of per-minute time-stamped reading of desk light (D.L.), ceiling light (C.L.) and ceiling fan status (on/off), usage (in minutes) per day, points, rank, frequency of visits to the web portal, time of day, weekday/weekend, external features and indicators for breaks, midterms and finals for each player.

2.2. Proposed Segmentation Method

For segmentation analysis, both supervised and unsupervised segmentation methods can be implemented on the social game dataset. For supervised classification, we use rank (indicative of energy efficiency characteristics) as the label. But, such a classification groups different players together as per their overall rank, not taking the distribution of their energy efficiency across different scenarios such as

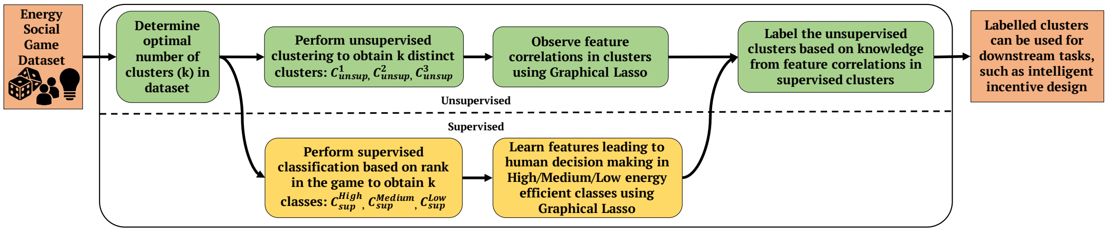


Figure 1. Overview of the proposed segmentation method

time into account. For every player, the data samples corresponding to low energy efficient behavior should be clustered separately than high energy efficient behaviors so as to have an accurate understanding of the interplay of features governing human decisions for energy usage. In this case, unsupervised clustering proves helpful, but provides unlabelled clusters. This poses a trade-off between supervised classification and unsupervised clustering methods. The trade-off signals to use the novelty of both unsupervised and supervised segmentation together to build an optimal model. Knitting together via a powerful tool, the graphical lasso algorithm(Hastie et al., 2015), we present a novel methodology to perform segmentation in energy social games. Using silhouette score, we first find the optimal number of clusters in the data. The silhouette score $\in [-1, 1]$, is a measure of how similar an object is to its own cluster compared to other clusters. The silhouette scores for number of clusters is given in Table 1, and is highest for optimal number of clusters as 3. We use K-means algorithm with k= 3 to obtain the unsupervised clusters as C^1_{unsup}, C^2_{unsup} and C^3_{unsup}. We also divide the players into three classes in a supervised way based on rank as C^{High}_{sup}, C^{Medium}_{sup} and C^{Low}_{sup}, where the superscripts signify the energy efficiency behavior of each class. We then use knowledge of feature correlations in supervised classes using graphical lasso (GLASSO) to label the unsupervised clusters as high/medium/low energy efficient. Finally, the labelled unsupervised clusters can be further explored for downstream tasks, such as intelligent incentive design and demand response. The whole process is illustrated in Figure 1.

No. of Clusters	2	3	4	5
Silhouette Scores	0.684	0.749	0.611	0.540

Table 1. Silhouette Scores for different number of clusters

3. Formulation of feature correlation learning

Let the features representing the social game data be denoted by $Y = (Y_1, Y_2, \cdots, Y_S)$. From a graphical perspective, Y can be associated with the vertex set $V = \{1, 2, \cdots, S\}$ of some underlying graph. The structure of the graph is utilized to derive inferences about the relationship between the features. We use the GLASSO algorithm (Hastie et al., 2015) to realize the underlying graph structure.

Consider the random variable Y_s at $s \in V$. We use the neighbourhood-based likelihood for graphical representation of multivariate gaussian random variables. Let the edge set of the graph be given by $E \subset V \times V$. The neighbourhood set of Y_s and collection of other random variables is:

$$\mathcal{N}(s) = \{k \in V | (k, s) \in E\} \quad (1)$$
$$Y_{V\setminus\{s\}} = \{Y_k, k \in (V - \{s\})\} \quad (2)$$

For undirected graphical models, Y_s is conditionally independent of nodes not directly connected to it given $Y_{\mathcal{N}(s)}$ by the conditional independence property. So, the problem of finding the edge set is formulated as predicting the value of Y_s given $Y_{\mathcal{N}(s)}$ (eventually given $Y_{\setminus\{s\}}$). The conditional distribution of Y_s given $Y_{\setminus\{s\}}$ is also Gaussian, so corresponding optimization problem for vertex s is formulated:

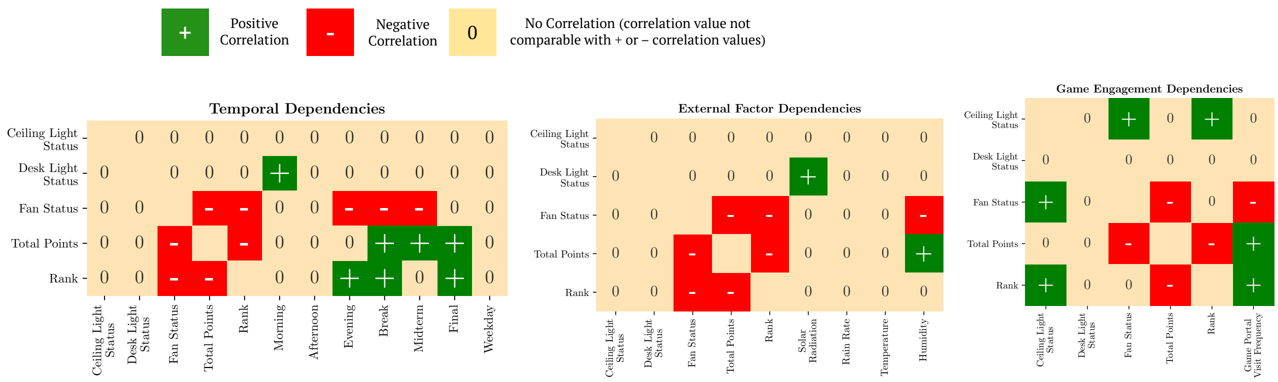
$$\hat{\beta^s} \in \underset{\beta^s \in \mathbb{R}^{S-1}}{\operatorname{argmin}} \left\{ \frac{1}{2N} \sum_{j=1}^{N} (y_{js} - y^T_{j,V\setminus s}\beta^s)^2 + \lambda \|\beta^s\|_1 \right\} \quad (3)$$

The β^s terms dictate the edge set for node s in the graph.

4. Results

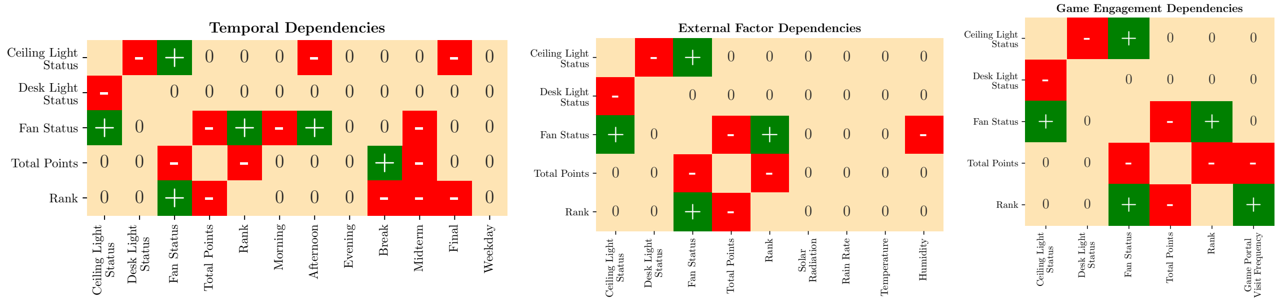
4.1. Supervised Feature Correlation Learning

The feature correlations obtained using GLASSO are presented in Figure 2,3,4. The feature correlations reveal that a low energy efficient player tends to use each resource independently as observed with no correlation between the resource usage identifiers. There is a positive correlation between morning and desk light usage indicating heedless behavior towards energy savings. External parameters play a significant role in energy usage behavior of this class. A medium energy efficient player showcases predictable behaviors and co-optimizes the usage by alternating the use of ceiling and desk light. A high energy efficient player opportunistically saves energy during breaks and midterms. The decrease in absolute amount of points does not reflect in increase of ranks, completely opposite to that of low energy efficient class. The player is neither affected by the time of the day, nor by the external factors showing a dedicated effort to save energy. To enhance the explainable nature of



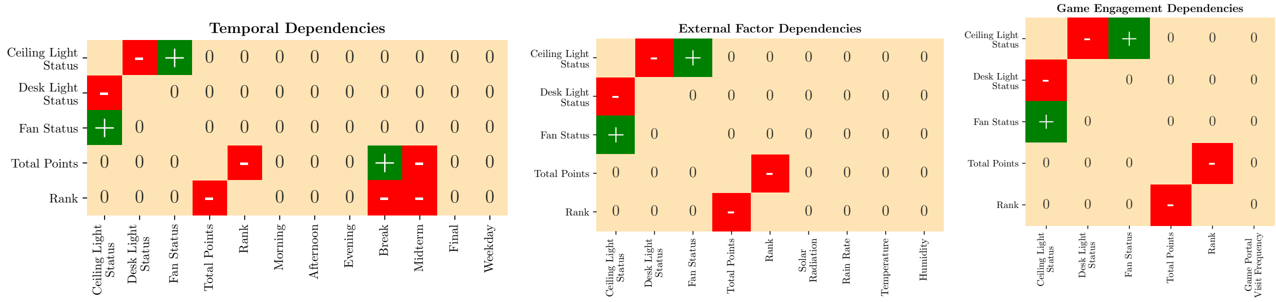
(c) (d) (e)

Figure 2. Feature correlations for a Low Energy Efficient Player ($\in C_{sup}^{Low}$)



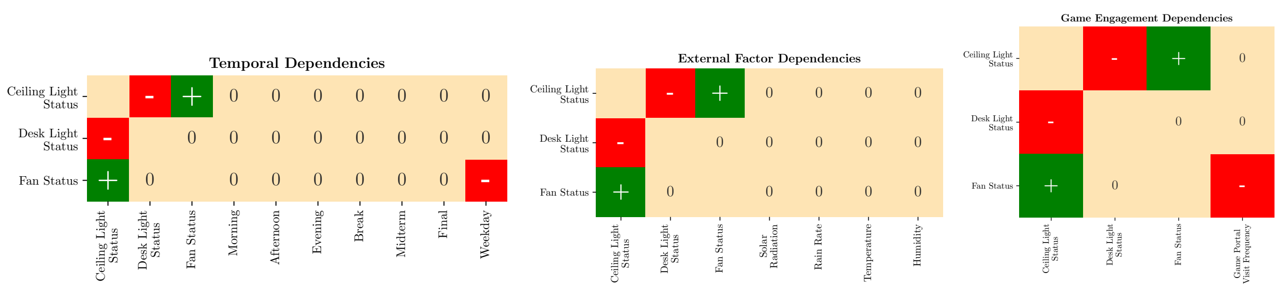
(a) (b) (c)

Figure 3. Feature correlations for a Medium Energy Efficient Player ($\in C_{sup}^{Medium}$)



(a) (b) (c)

Figure 4. Feature correlations for a High Energy Efficient Player ($\in C_{sup}^{High}$)



(a) (b) (c)

Figure 5. Feature correlations for energy usage behaviors in C_{unsup}^{3}. The labels “Total Points”and “Rank”are removed for unsupervised clustering.

Test whether X causes Y	Fan ⇒ Ceiling Light		Humidity ⇒ Fan		Desk Light ⇒ Fan		Ceiling Light ⇒ Desk Light		Morning ⇒ Desk Light		Afternoon ⇒ Fan		Evening ⇒ Ceiling Light	
Player type	p-value	F-statistic	p-value	F-statistic	p-value	F-statistic	p-value	F-statistic	p-value	F-statistic	p-value	F-statistic	p-value	F-statistic
Low Energy Efficient	0.54	0.37	**0.004**	8.12	0.06	3.55	0.81	0.06	0.4	0.71	**0.01**	6.1	**0**	25.3
Medium Energy Efficient	**0**	21.2	**0.008**	7.06	**0**	113.6	**0**	25.8	0.23	1.41	0.46	0.55	**0.0007**	11.5
High Energy Efficient	**0**	21.9	0.12	2.36	0.99	0.003	0.93	0.007	0.63	0.22	**0.04**	4.2	0.52	0.41

Table 2. Causality test results among various potential causal relationships using grangers causality method

our model, we studied the causal relationship between features using granger causality test (Table 2). The p-values for which granger causality is established are highlighted in the table. For medium and high energy efficient building occupants, ceiling fan usage causes ceiling light usage indicating predictive behavior. In both low and medium energy efficient building occupants, external humidity causes ceiling fan usage unlike a high energy efficient player. The above results confirm the explainability of the proposed model.

4.2. Labelling unsupervised clusters

We also learn the feature correlations in clusters obtained from unsupervised clustering of data in Section 2.2. Based on the feature correlation knowledge gained from different supervised classes, we label the clusters as having low, medium or high energy efficient data. As an illustration, the feature correlations for C^3_{unsup} in Fig 5, it is evident that it exhibits predictability in behavior and energy savings. This is indicative of the similarity between the energy efficiency characteristics manifested by C^3_{unsup} and C^{High}_{sup}. So, C^3_{unsup} is labelled as the high energy efficient cluster. Following the same comparison, the labelling is done as, $\{C^1_{unsup} \sim$ Medium Energy Efficient$\}$,$\{C^2_{unsup} \sim$ Low Energy Efficient$\}$ and $\{C^3_{unsup} \sim$ High Energy Efficient$\}$. We also compute the similarity using Pearson Correlation and RV coefficient (Robert & Escoufier, 1976) between the feature correlation matrices (Figure 6) which indicate the same assignment. The labelled unsupervised clusters can be used for downstream tasks as discussed in Section 5.

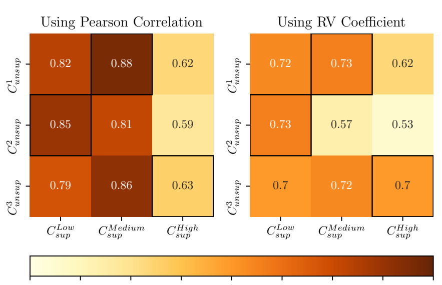


Figure 6. Similarity between feature correlation matrices. The highest value in each column is highlighted and corresponds to the matching of supervised classes to the unsupervised clusters

5. Conclusion and Future Work

A novel GLASSO based approach for segmentation analysis in energy social games was presented in this work. The analysis included clustering of agent behaviors and an explainable statistical model towards human decision-making for energy usage in competitive environments. The proposed method can provide characteristic clusters demonstrating different energy usage behaviors. More details on this research work can be found at (Das et al., 2019).

There are several directions for future research. An improved version of energy social game, similar in structure to that of (Konstantakopoulos et al., 2019) but with intelligent incentive design and privacy preserving techniques can be implemented, with building occupants and managers interaction modeled as a reverse stackelberg game (leader-follower) in which there are multiple followers that play in a non-cooperative game setting (Ratliff et al., 2014). By leveraging proposed segmentation analysis, an adaptive model can be formulated that learns how user preferences change over time, and thus generate the appropriate incentives. Furthermore, the learned preferences can be adjusted through incentive mechanisms (Ratliff & Fiez, 2018) and a tailored mean-field game approach (Gomes & Saúde, 2018) to enact improved energy efficiency. Above two operations can be carried out in a tree structure, with segmentation carried out in regular intervals in each of the tree branches, as depicted in Figure 7. This can be coherently designed with other smart building systems (Zou et al., 2019a;b;c; Liu et al., 2019; 2018b; Jin et al., 2018). Summing up, this would result in a novel mechanism design, effectively enabling variation in occupant's behaviors, in order to meet, for instance, the requirements of a demand response program. Another line of future work can be to study the delayed impacts of energy social game and design it accordingly to achieve long term energy efficiency, like a research in same line (Liu et al., 2018a).

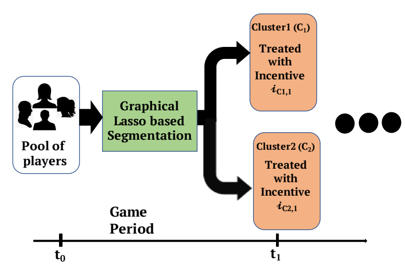


Figure 7. Tree based incentive design with tailored incentives for clusters employing proposed GLASSO based segmentation.

6. Acknowledgments

This research is funded by the Republic of Singapore's National Research Foundation through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore-Berkeley Building Efficiency and Sustainability in the Tropics (SinBerBEST) Program. BEARS has been established by the University of California, Berkeley as a center for intellectual excellence in research and education in Singapore. The work of I. C. Konstantakopoulos was supported by a scholarship of the Alexander S. Onassis Public Benefit Foundation.

References

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. In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, pp. 1702–1709. IEEE, 2019.

Gomes, D. and Saúde, J. A mean-field game approach to price formation in electricity markets. *arXiv preprint arXiv:1807.07088*, 2018.

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

Jin, M., Jia, R., Das, H., Feng, W., and Spanos, C. Biscuit: Building intelligent system customer investment tools. In *Proc. 10th International Conference on Applied Energy (ICAE)*, 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, L. T., Dean, S., Rolf, E., Simchowitz, M., and Hardt, M. Delayed impact of fair machine learning. In *Proceedings of the 35th International Conference on Machine Learning*, pp. 3150–3158. PMLR, 2018a. URL http://proceedings.mlr.press/v80/liu18c.html.

Liu, S., Jin, M., Das, H. P., Spanos, C., and Schiavon, S. Personal thermal comfort models based on physiological parameters measured by wearable sensors. 2018b.

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.

McQuade, J. A system approach to high performance buildings. *United Technologies Corporation, Tech. Rep*, 2009.

Ratliff, L. J. *Incentivizing efficiency in societal-scale cyber-physical systems*. PhD thesis, UC Berkeley, 2015.

Ratliff, L. J. and Fiez, T. Adaptive incentive design. *arXiv preprint arXiv:1806.05749*, 2018.

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.

Robert, P. and Escoufier, Y. A unifying tool for linear multivariate statistical methods: The rv- coefficient. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 25(3):257–265, 1976. ISSN 00359254, 14679876.

Zou, H., Das, H. P., Yang, J., Zhou, Y., and Spanos, C. Machine learning empowered occupancy sensing for smart buildings. 2019a.

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, 2019b.

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