Real-time Contextual Anomaly Detection in Buildings Energy Consumption

Haroon Rashid, Pushpendra Singh
Indraprastha Institute of Information Technology Delhi

INTRODUCTION

Buildings consume 50% of electric energy worldwide. A study by Roth et al. [1] shows that 20% of this energy gets waste due to faults. These faults result in abnormal usage patterns called as anomalies. In this work, we propose a real-time anomaly detection method.

PREVIOUS WORK

Work by [2] initially predicts energy usage and finally if the actual usage is greater than 2 standard deviations, then usage is considered as anomalous

- Work does not consider most recent days for prediction
- Does not remove anomalous readings in prediction step

CHALLENGES

- Accuracy in energy prediction
- Avoiding anomalous historical readings during prediction

DATASET

- Open Source dataset, Dataport
  - Energy usage of 4 homes in Austin, USA for a duration of 90 days (1st June - 31st Aug. 2014)
  - Data at both aggregate and device level
  - Data sampling rate used: 10 minutely average

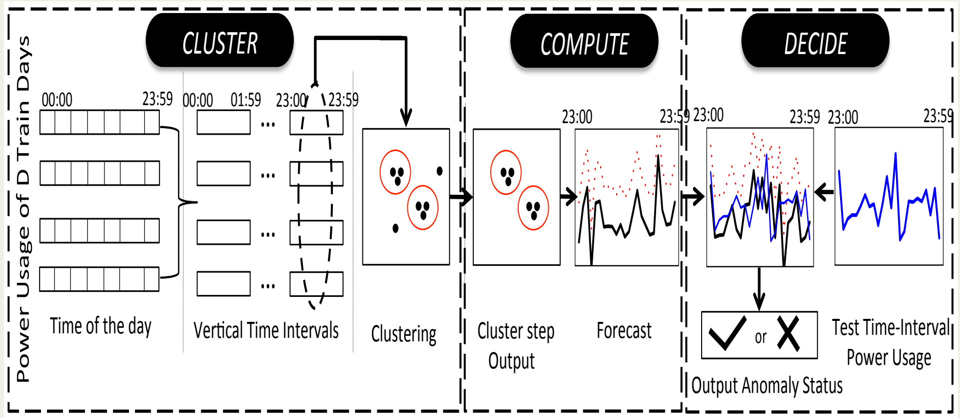
BASELINE APPROACH

- Anomalous - Seasonal Auto-regressive Integrated Moving Average (*a*-SARIMA)
- *a*-SARIMA works in two steps:
  - Prediction using SARIMA
  - Anomaly detection using 2 standard deviations

1. Roth et al. Energy impact of commercial building controls and performance diagnostics, US dept. of Energy, 2005
2. J.-S. et al., Real-time detection of anomalous power consumption, Ren. And Sus. Energy Reviews 2014

OUR APPROACH

- Our approach, temporally contextual based approach is called as Cluster, Compute and Decide (CCD).
- Uses historical train data contextually to predict the usage. Usage found greater than the prediction interval for continuous defined duration is considered as anomalous.



EXPERIMENTAL SETUP

- *a*-SARIMA
  - Autoregressive term (p) = 4, Moving Average (q) = 2, Seasonal p = 2, Seasonal q = 0, Difference (d) = 1, periodicity = 1 day
- Common parameter values to *a*-SARIMA and CCD
  - Number of training days = 20
  - Anomaly status checking Interval = 60 minutes
  - Considered Continuous Anomalous Usage = 40 minutes

Method	Home 1	Home 2	Home 3	Home 4
CCD	0.83	0.95	0.56	0.92
a-SARIMA	0.73	0.76	0.86	0.80

Table 1: F1 score for CCD and *a*-SARIMA on four homes

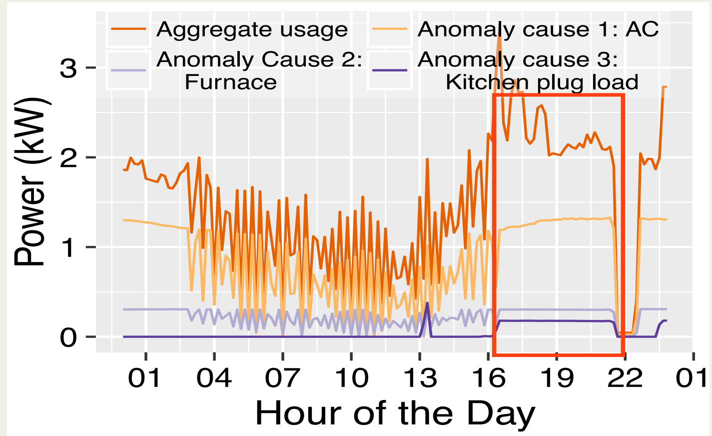
ACKNOWLEDGEMENT

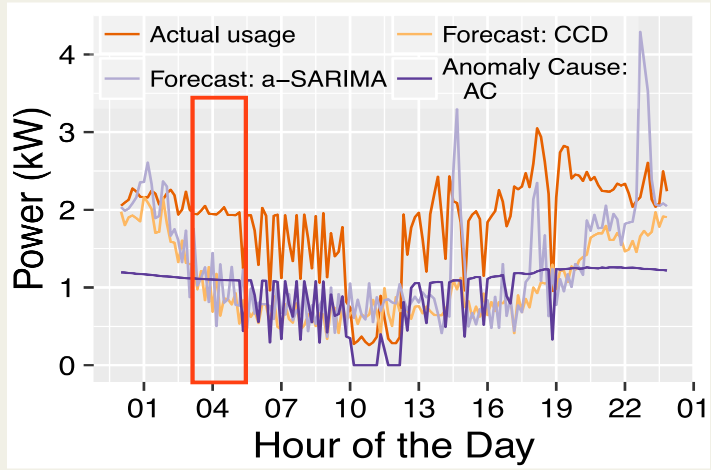
Authors will like to acknowledge the support provided by ITRA project, funded by DEITy, Government of India, under grant with Ref. No. ITRA/15(57)/Mobile/HumanSense/01

EXPERIMENTAL RESULTS

Method	Metric	Home 1	Home 2	Home 3	Home 4
CCD	SMAPE	0.275	0.190	0.339	0.275
	MASE	1.534	0.952	1.369	1.535
a-SARIMA	SMAPE	0.337	0.228	0.442	0.406
	MASE	1.983	1.180	1.565	2.452

Table 2: Prediction Accuracy of CCD and a-SARIMA using SMAPE and MASE





Figures showing anomalous energy usage with bounded rectangles

CONCLUSION

CCD performs better in both energy prediction and the anomaly detection in real-time as compared to the baseline *a*-SARIMA.