List of Figures

1.1 The sankey diagram showing the energy flow from sources of production to consumptions across different sectors in the USA for the year 2015 [13]. It is estimated that 35% of energy is wasted in both residential and commercial buildings. 2

1.2 Electricity consumption of residential [39] and commercial [7] buildings by end-use in the USA for the year 2010. 3

2.1 SensorAct system architecture showing various layers and system components. 13

2.2 Components of the VPDS and Broker in SensorAct. 15

2.3 Sensor data representation using WaveSeg format. 18

2.4 An example of a guard rule and a macro for selective sharing of sensor data. 18

2.5 Scripting framework workflow with an example Tasklet. 21

2.6 Sample browser and mobile application user interface. 25

2.7 Deployment scenario of SensorAct architecture showing VPDSes deployed across four different locations and a common broker hosted at IIIT-Delhi. 27

2.8 Daily electricity usage pattern of street lights (6:30pm to 6:00am every day) for 12 days. Some abnormal energy usage events are marked in red. 29

2.9 Electricity usage patterns of sports area lights for 12 days. Few street lights around the sports area consuming about 0.9 kilo-watts of power are used from 6pm to 6am every day. Each spike in this plot corresponds to the usage of flood lights in the sports area. 29
2.10 Campus-wide total commercial electricity usage (from two transformers) for 1 week. The spikes in transformer 1 corresponding to the electricity consumption of HVAC systems, that consumes over 100 kilo-watts.

3.1 Evolution of building middleware systems based on support for processing sensor data. (a) primitive middleware provides no support for sensor data analytics, (b) rule-based middleware provides trigger-actions based on thresholds, and (c) context-based middleware provides sophisticated analytics for inferring context from sensory data.

3.2 OpenBAN system architecture showing various system components.

3.3 The workflow of the Context Inference Engine for training and execution mode.

3.4 OpenBAN user interface showing the “Aggregate-Analyze-Act” pipeline for a sensor data analytical application.

3.5 Integration of SensorAct and OpenBAN systems for energy disaggregation application.

3.6 Water usage pattern for six sprinkler stations.

3.7 Indirect occupancy count prediction from network activity using SVM classifier.

3.8 Comparison of execution time between local and cloud hosted analytics engine for the energy forecasting contextlet.

4.1 Hourly power usage (normalized) of different buildings with in a large commercial building complex (neighborhood) in Sweden for a year from 1st Feb 2013 to 31st Jan 2014. It shows, (a) daily and weekly power usage cycles with seasonal variations during summer, winter and holidays, (b) examples of single point anomaly (marked in black), and (c) examples of sequence anomaly (marked in red). X-axis denotes the day of the year while Y-axis is hour of the day.
4.2 Logical flowchart of the proposed anomaly detection algorithm. The function $f$ computes self anomaly score for each meter and for each temporal context set separately followed by function $g$ concatenates them. Function $p$ computes the adjusted anomaly score for each meter data based on the available neighborhood information.

4.3 The baseline correlation between 10 buildings for a year in the Sweden commercial building data set. Meters are arranged using hierarchical clustering algorithm.

4.4 The baseline correlation between 18 apartments for a year in the Indian residential buildings dataset. Meters are arranged using hierarchical clustering algorithm.

4.5 Hourly meter readings of a Sweden commercial building with computed anomaly score by different baseline and proposed anomaly detection methods. It shows several instances of point and sequence anomalies and how the computed anomaly score differs using the temporal and neighborhood information.

4.6 Anomaly score comparison of (a) self anomaly score without using any context verses self anomaly score using only the temporal context, (b) self-anomaly score only using the temporal context verses using available neighborhood information, and (c) a violin plot (a combination of box and density plot) shows the differences between anomaly scores (self minus adjusted), by using different adjustment weights for the Sweden commercial building dataset.

4.7 Adjusted anomaly score difference for different weights over time for the Swedish commercial building data set. The curve with the smallest magnitude corresponds to a weight of 10% and the one with the highest magnitude corresponds to a weight of 100%. Positive values indicate a reduction in the anomaly score after neighborhood comparison and vice versa.
4.8 Hourly meter readings of an Indian residential building with computed anomaly score by different baseline and proposed anomaly detection methods. It shows several instances of point and sequence anomalies and how the computed anomaly score differs using the temporal and neighborhood information. 86

4.9 Anomaly score comparison of (a) self anomaly score without using any context verses self anomaly score using only the temporal context, (b) self-anomaly score only using the temporal context and using available neighborhood information, and (c) violin plot (a combination of box and density plot) shows the differences between anomaly scores (self minus adjusted), by using different adjustment weights for the Indian residential building dataset. 87

4.10 Adjusted anomaly score difference for different percentage of weights over time for the Indian residential building data set. The curve with the smallest magnitude corresponds to a weight of 10% and the one with the highest magnitude corresponds to a weight of 100%. Positive values indicate a reduction in the anomaly score after neighborhood comparison. 88
List of Tables

2.1 Primitive API functions in Scripting framework available to use in Tasklet Scripts. 22
2.2 A list of APIs supported by different components of the SensorAct architecture. 24
2.3 Different deployment details of SensorAct system. Ambient sensors include Temperature, light intensity, motion, and door contact status 28
3.1 Motivating energy management applications which require complex features and analytics on top of the collected sensor data. 42
3.2 List of system requirements for designing an analytics middleware and the corresponding system components. 45
3.3 List of different categories of features that can be identified from various sensor data streams to infer a wide range of context information of a building. These features are computed for each time window spanning a N-seconds interval. 48
4.1 Details about the injected abnormal energy usage events into the residential building dataset. 80