Energy Wastage Detection in Smart Buildings

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF REQUIREMENT FOR THE DEGREE OF

Bachelor of Science in Computer Science and Engineering

Submitted by

SAMIUN-RAJI SIFAT (134403)

NABIL SHAD (134405)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Islamic University of Technology (IUT)

GAZIPUR, BANGLADESH

NOVEMBER 2017

Declaration of Authorship

We, Samiun-Raji Sifat (134403) and Nabil Shad (134405), declare that this thesis titled, 'Energy Wastage Detection in Smart Buildings' and the works presented in it are our own. We confirm that:

- This work has been done for the partial fulfillment of the Bachelor of Science in Computer Science and Engineering degree at this university.
- Any part of this thesis has not been submitted anywhere else for obtaining any degree.
- Where we have consulted the published work of others, we have always clearly attributed the sources.

Submitted by:

Samiun-Raji Sifat (134403)

Nabil Shad (134405)

Energy Wastage Detection in Smart Buildings

Approved By:

Ashraful Alam Khan
Thesis Supervisor,
Lecturer
Department of Computer Science and Engineering
Islamic University of Technology

Acknowledgements

First and foremost, we offer gratitude to the almighty Allah (SWT) for giving us the capabilities to do this work with good health.

We are grateful to our research supervisor, Ashraful Alam Khan, for the support and guidance throughout our research at Islamic University of Technology (IUT). He created a nice research environment for which we were able to explore many ideas without constraints. We have gained a wealth of knowledge and experience in science and engineering though his direction that is beyond value to our future endeavor. For all his effort as our true mentor, we express our heartfelt gratitude to him.

We would like to thank all the faculty members of the department of CSE, IUT for their inspiration and help.

And last but not the least we are thankful to our family, friends and well-wishers for their support and inspiration. Without them it would never have been possible for us to make it this far.

Table of Contents

List of Figures

No table of figures entries found.

Abstract

Energy is wasted due to unconsciousness and negligence of the users. Buildings are the major source of energy consumption. So, most of the energy wastage detection techniques are designed for buildings. In our proposed method we use the occupancy based sensors for the detection of the unnecessary consumption. Along with current data, the stored data is also used for the decision making. To attain accuracy and efficiency, the reinforcement learning algorithm is used in data processing. A user interface is there to take the user feedback. This user feedback is used for the learning growth of reinforcement learning algorithm.

1. Introduction

At present electricity crisis is one of the most severe problems of our country, because the energy infrastructure of Bangladesh is quite small, insufficient and poorly managed. Only 62% of our total population has access to electricity. Yet it is not possible to provide electricity to all of them simultaneously, because the demand is much greater than the supply. That's why Bangladesh is still reeling under 600 - 1200 MW of 'load-shedding'. A situation which deteriorates during irrigation seasons, when the demand-supply gap reaches up to 1500 MW. Every year Bangladesh lose almost $1 billion and 0.5% of our national GDP due to load-shedding.

Now, if we think what the causes are of this electricity crisis. We will find that there are three reasons behind it.

- Poor and backdated electricity distribution system.
- Illegal connections.
- Wastage of electricity.

The first two problems can only be solved by proper initiatives from the government. So we decided to focus only on the third problem.

Residential and industrial sectors consume about 43% and 44% electrical energy respectively, i.e. a total of about 87% of power consumption occurs in these two sectors.[1] All power sector experts acknowledge that a large part of electrical energy is consumed for lighting. Lighting and other electronic device management system in commercial buildings is very inefficient. As a result electronic devices are often turned on even though it is not necessary. This causes wastage of large amount of electricity. Besides in residential buildings people sometimes waste electricity due to lack of awareness.

These are the main reasons why a significant amount of electricity is being wasted. A way to manage the load is the introduction of *Smart and Intelligent* energy efficient electricity management system.

From various sources we gathered enough data to be ensure that if an efficient energy usage system can be ensured then it can have huge impact on the electricity crisis that the world is facing right now. Some statistics below will make the impact more clear.

- 1-4% more investment in energy efficiency can save $60 billion.[2]

- Intelligent efficiency measures applied to just 35% of eligible commercial floor area in buildings can save 50 TWh by 2030.[1]
- Over 3 billion units of electricity, or a day's national consumption, were wasted in 2014-15 in India.[3]

As established before, a huge amount of electricity is being wasted due to lack of awareness and inefficient management of electronic devices. These problems occur because the electronic devices are manually controlled. If the electronic devices were controlled automatically by a central intelligent management software, then the usage of electricity would be more efficient and as a result the wastage of electricity would decrease. That is why our target is to minimize the loss of electricity in our country due to poor management and lack of awareness by introducing an intelligent and efficient distributed occupancy based sensor network.

We are proposing a distributed sensor based network system where we will collect the wastage information from this network system and rectify through actions with the help of actuator network system. Our system will be able to analyze the usage after a period of time as well.

2. Background Knowledge

Determining the abnormal energy consumption is an active area of research. A lot of research works are done on this subject and many end products are there in the market. We will have a detailed discussion in the following sections about these approaches.

2.1. Occupancy based line of sight model

This model has an occupancy sensor connected to each smart device in the building. These devices have a one to one relationship with the sensors designated to it. The used sensor sends a signal based on the occupancy of the room. Based on the signals from the sensor, the system detects anomaly and start or stop a device.

One of the very general area of using the occupancy based sensor is the lights and fans in the room. The smart lights and fans automatically turns on or off with the occupancy in the room. The system in the line of sight basis. When a sensor gets some reading in a room only then I will send signal to its related device. And a device is acting based on the input from its sensor. The different sensors in the house has no connection among themselves.

The model has its own advantages and disadvantages. The main advantage is the invention of an automated system for controlling different electric devices around the house. This creates a lot of convenience for the users. Another important advantage of this system is that it is very accurate. As it work on the signal from a dedicated sensor, the probability of error is very less if the equipment are working properly. The model is very simple and easy to implement.

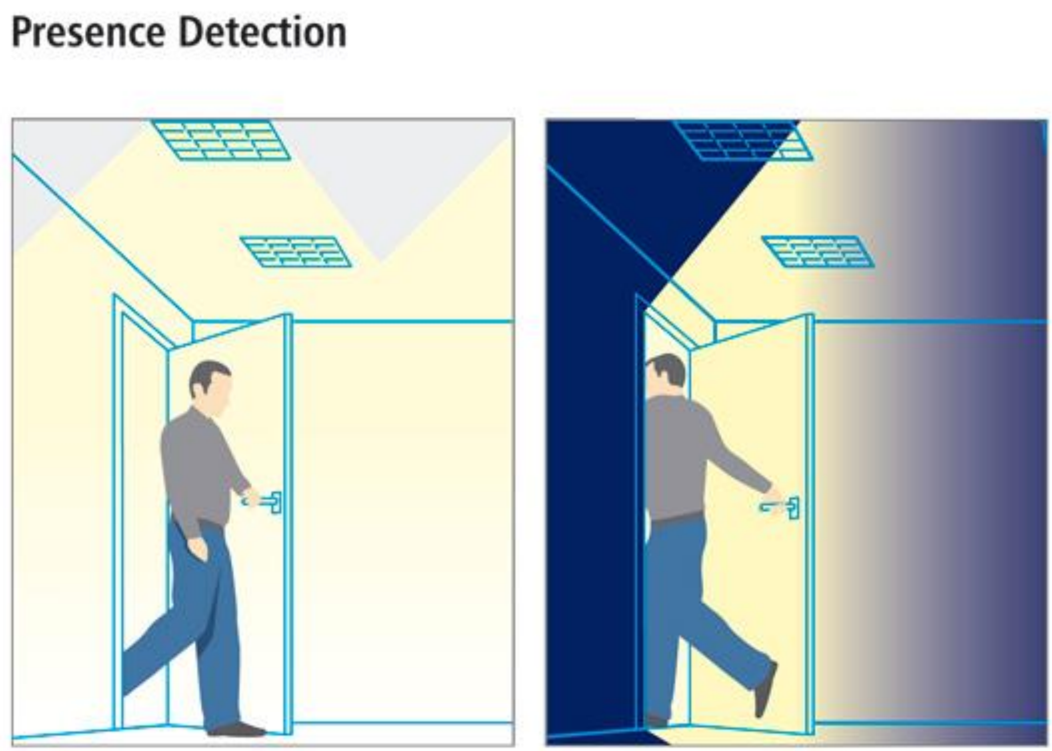

Fig. 1: Occupancy based sensing model.

The model has some disadvantages also. The system is not efficient. Let us think of a scenario of a flat with two rooms. A child is there in the building. The flat has occupancy based line of sight model implemented. So, when someone enters a room, the light lit up. Let, the child is running between the rooms again and again. As a result, the lights in the rooms are going on and off again and again. This situation is very inefficient. Again, this model does not any control to the users about the usage. So, user may not feel easy with the system in some cases.

2.2. Occupancy based distributed model

Distributed model of occupancy based sensing comes with solution of being inefficient in the line of sight model. Here different sensors placed across the building are connected to a single network. The network controller is a sink node who gives different decision about abnormal uses. We can solve the problem scenario given in the previous section. When a child is running between two rooms, the lights will not be on and off frequently. This model can state that when someone is within the certain distance of a room, the lights will not go off. And, the location of the child will be given by the occupancy sensors of the other room. The sensor data will be shared by the sink node.

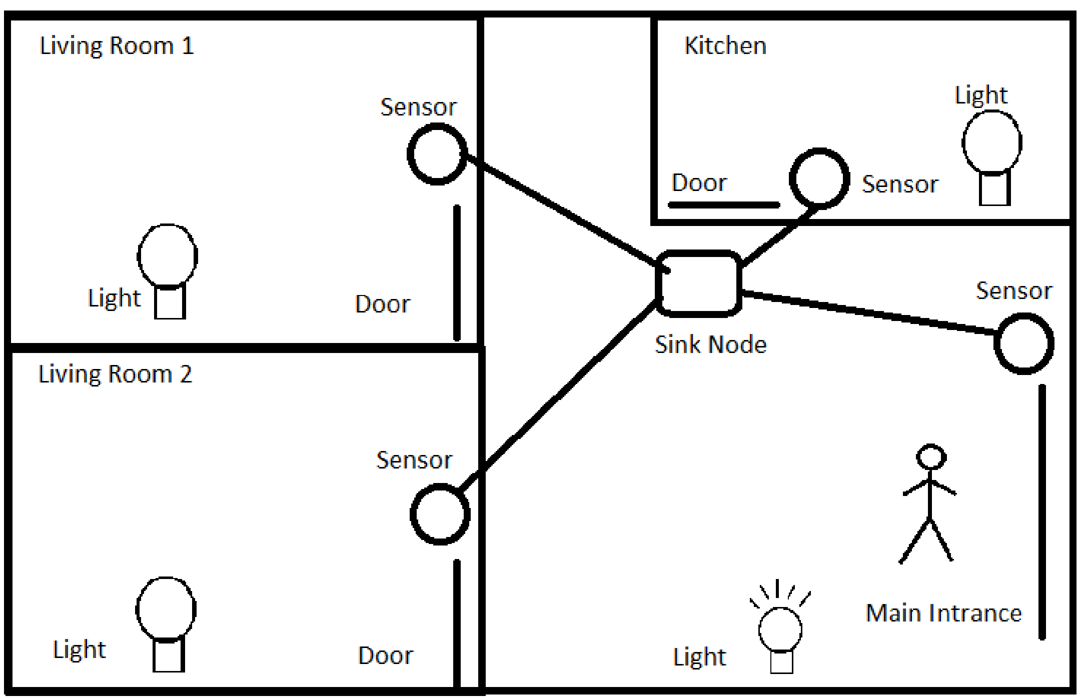

Fig 2: Distributed occupancy based model.

The problem with this model is it works with the current data generated from sensors. The model does not count the historical usage patterns and other contextual features. So, there will be some problems with the accuracy of the model.

The general shortcomings of all the occupancy based model is that they perform well with the rarely occupied rooms. However, when working in a frequently occupied room, the occupancy based models have a relatively poor performance.

3. Related Works

The solution of the occupancy based models comes with the models with historical data and usage pattern. These models integrates the occupancy sensor data with historical data and deferent contexts to detect anomaly. Some of these models are discussed subsections.

3.1. Strip, Bind and Search[5]

A typical large building contains thousands of sensors, monitoring the HVAC system, lighting, and other operational sub-systems. With the increased push for operational efficiency, operators are relying more on historical data processing to uncover opportunities for energy-savings. However, they are overwhelmed with the deluge of data and seek more efficient ways to identify potential problems. This paper, we present a new approach called the Strip, Bind and Search (SBS); a method for uncovering abnormal equipment behavior and in-concert usage patterns. SBS uncovers relationships between devices and constructs a model for their usage pattern relative to other devices.

This method consists of two key components. Strip and Bind, the first part of the proposed method mines the raw sensor data, identifying inter-device usage patterns. We first strip the underlying traces of occupancy-induced trends. Then we bind devices whose underlying behavior is highly correlated. Search, the second part of the method monitors devices relationships over time and reports deviations from the norm. It learns the normal inter-device usage using a robust, longitudinal analysis of the building data and detect anomalous usages.

Discovering devices that are used in concert is non-trivial. SBS decomposes each signal into an additive set of components, called Intrinsic Mode Functions (IMF) and reveals the signal patterns at different frequency bands. IMFs are obtained using Empirical Mode Decomposition.

Empirical Mode Decomposition (EMD) is a technique that decomposes a signal and reveals intrinsic patterns, trends, and noise. This technique has been widely applied to a variety of datasets, including climate variables, medical data, speech signals, and image processing. EMD's effectiveness relies on its empirical, adaptive and intuitive approach. In fact, this technique is designed to efficiently decompose both non-stationary and non-linear signals without requiring any a priori basis functions or tuning. EMD decomposes a signal into a set of oscillatory components called intrinsic mode functions (IMFs). An IMF satisfies two conditions:

(1) It contains the same number of extrema and zero crossings (or differ at most by one).

(2) The two IMF envelopes defend by its local maxima and local minima are symmetric with respect to zero. Consequently, IMFs are functions that directly convey the amplitude and frequency modulations.

EMD is an iterative algorithm that extracts IMFs step by step by using the so-called sifting process; each step seeks for the IMF with the highest frequency by sifting, then the computed IMF is removed from the data and the residual data are used as input for the next step. The process stops when the residual data becomes a monotonic function from which no more IMF can be extracted.

By applying EMD to energy consumption signals we obtain a set of IMFs that precisely describe the devices consumption patterns at different frequency bands. Therefore, we can focus our analysis on the smaller time scales, ignoring the dominant patterns that prevent us from effectively analyzing raw signals.

There are numerous techniques to retrieve IMF frequencies [15]. In this work we take advantage of the Generalized Zero Crossing (GZC) [13] because it is a simple and robust estimator of the instantaneous IMF frequency [15]. GZC is a direct estimation of IMF instantaneous frequency using critical points defined as the zero crossings and local extrema.

Hereafter, we refer to the time scale of an IMF as the average of the instantaneous periods along the whole IMF. Because the time scale of each IMF depends on the original signal, we propose the following to efficiently compare IMFs from different signals. We cluster IMFs with respect to their time scales and partially reconstruct each signal by aggregating its IMFs from the same cluster. Then, we directly compare the partial signals of different devices.

All the devices are compared pairwise at the four different time scale ranges. Consequently, we obtain four correlation matrices that convey device similarities at different time scales. Each line of these matrices (or column, since the matrices are symmetric) reveals the behavior of a device - its relationships with the other devices at a particular time scale. The matrices form the basis for tracking the behavior of devices and to search for misbehavior.

Search aims at identifying misbehaving devices in an unsupervised manner. Device behavior is monitored via the correlation matrices presented in the previous section. Using numerous observations SBS computes a specific reference that exhibits the normal inter-device usage pattern. Then, SBS compares the computed reference with the current data and reports devices that deviate from their usual behavior.

3.2. Collective Contextual Anomaly Detection[6]

Buildings have a great potential for helping to meet energy efficiency targets. Hence, energy saving goals that target buildings can have a significant contribution in reducing environmental impact. Today’s smart buildings achieve energy efficiency by monitoring energy usage with the aim of detecting and diagnosing abnormal energy consumption behaviour. This research proposes a generic collective contextual anomaly detection (CCAD) framework that uses sliding window approach and integrates historic sensor data along with generated and contextual features to train an autoencoder to recognize normal consumption patterns. Subsequently, by determining a threshold that optimizes sensitivity and specificity, the framework identifies abnormal consumption behaviour.

One approach to building energy efficiency is to monitor energy usage with the aim of detecting and diagnosing abnormal consumption behaviour. In recent decades, modern buildings have been equipped with an increasing number of sensors and smart meters. By analyzing data from these devices, normal consumption profiles can be identified. Subsequently, when patterns that do not conform to the normal profiles are detected, the building manager is notified, and appropriate energy-saving measures are taken. More importantly, for safety-critical building services such as gas consumption, early detection and notification of anomalous behaviour (gas leakage) can help prevent potentially life-threatening disasters.

Most modern buildings are equipped with a built-in control system referred to as a building automation system (BAS). A BAS is a part of what is referred to as an intelligent or smart building, and it enables building managers to automate and oversee the energy efficiency aspect of a building. By providing early detection and diagnosis of abnormal building behavior, contextual collective anomaly detection helps not only to reduce financial cost, but also, on a larger scale, to reduce the environmental impact of electric power generation. This research proposes a framework to identify collective contextual anomalies. The Collective Contextual Anomaly Detection (CCAD) framework uses a sliding window approach and integrates historic sensor data along with generated and contextual features to identify contextually abnormal patterns in sensor data. The framework is flexible and can adapt to requirements of the anomaly detection domain. This provides an anomaly detection platform that can be tuned to stringent or lenient requirements with regards to sensitivity, specificity or an optimal overall value of these two metrics. Moreover, two anomaly detection models are compared, one trained using historic sensor data, generated features and contextual features, and the other model trained with all these except the generated features.

The following sections describe the components of this framework.

Data Pre-processing

The term "sensor data" in this case represents a time stamped record of consumption data recorded at regular intervals. These sensor data usually suffer from noise and incompleteness caused by faulty devices or communication errors; hence, the data requires cleaning.

1) Data Cleaning: To mitigate the negative impact of noisy and incomplete data on CCAD framework performance, these data must be removed from the dataset. Depending on the problem domain, noisy data are indicated by values outside the valid range. For instance, in the building energy domain, negative electric consumption values are considered noisy. Incomplete data in this context refers to the existence of missing data within an hourly sliding window data. Moreover, the proposed CCAD framework uses a sliding window to identify anomalies; a sliding window in this case is a specific window size that includes a set of consecutive values. Hence, if the data in a sliding window are incomplete, then that specific input set is removed from the dataset.

2) Feature Preparation: Given a clean dataset, the feature preparation component focuses on the arrangement and generation of features and involves two sub-steps: data reorganization and feature generation.
Data Reorganization: This sub-step involves reorganizing the dataset so that an input data instance is represented not by a single consumption value, but instead by a set of consecutive consumption values.

Feature Generation: This component introduces contextual or behavioural features into the CCAD framework. In the building energy consumption domain, the context can be spatial, temporal or weather-related. Moreover, by deriving additional sensor data features such as the mean and median, more insights can be obtained from the sliding window sensor data. The generated features are described in Table II. The temporal contextual features day of year, season, month, day of week, and hour of day are selected because energy consumption exhibits temporal seasonality. The generated features (_x, s, Sn-S1, (_xi-_x(i+1)), (_x(i+1)-_xi), Q1, Q2, Q3, and IQR) are selected to explore whether or not these features affect the performance of the CCAD framework.

3) Normalization: To avoid suppressing the information within the smaller-valued features, the features must be normalized. The data were normalized by rescaling the features of range in [0 1].

Model Training and Testing

The CCAD framework uses unlabelled sensor data; hence, it relies on an unsupervised learning algorithm to identify collective contextual anomalies. The basic assumption is that historic sensor data are predominantly normal. Because of this assumption, the historical dataset ("real dataset") can be split into training and testing data with the objective of using the testing data to evaluate the capacity of the CCAD framework to correctly identify normal behaviour.

1*) Real Training Data:* This part of the real data is used to train the Anomaly Detection Engine to recognize normal input data patterns.

2) Real Testing Data: Once the Anomaly Detection Engine is trained, the real testing data are used to test the specificity or true negative rate (TNR) of the model. True negative (TN) is the number of normal consumption patterns that are correctly identified; TNR is the ratio of TN and total negatives in the dataset.

3) Artificial Data: This research generates artificial anomalous data to test the sensitivity or true positive rate (TPR) of the model. True positive (TP) is the number of anomalous consumption patterns that are correctly identified; TPR is the ratio of TP and total positives in the dataset.

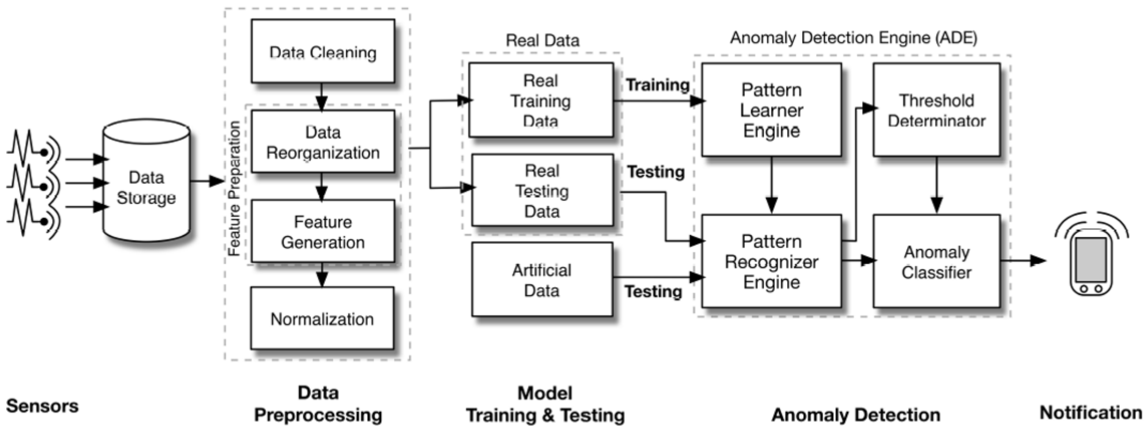

Fig 3: Collective Contextual Anomaly Detection (CCAD) framework

Anomaly Detection Engine (ADE)

Anomaly detection using dimensionality reduction relies on the assumption that data contain variables that are correlated with each other and hence can be reduced to a lower dimensional subspace where normal and abnormal data appear substantially different.

1) Pattern Learner Engine (PLE): The pattern learner engine (PLE) uses an autoencoder to train a model to reconstruct input data patterns. During feature generation, to provide more information to the CCAD framework, attempts were made to introduce more contextual and other generated features. However, as the dimension increases, the engine falls into the trap of the curse of dimensionality where scalability and over-fitting issues become apparent. High-dimensional data impose performance strains on machine learning algorithms. By using an autoencoder, which performs non-linear dimensionality reduction, the CCAD framework gains computational efficiency and better classification accuracy compared to other dimensionality reduction techniques such as PCA or Kernel PCA. The PLE is generic and can be replaced by other dimensionality reduction techniques.

2) Pattern Recognizer Engine (PRE): Once the PLE has trained a model using the normal consumption patterns, the pattern recognizer engine (PRE) tests the model using the real testing dataset as well as the artificially generated anomalous dataset. The output of the PRE engine is a reconstruction error, which is a measure of how close the input data pattern is to the normal data pattern on which the model was initially trained. The PRE engine serves two purposes: the first is to help the threshold determinator (TD) find a suitable threshold value, and the second is to test whether new sensor data patterns conform to the normal consumption pattern.

3) Threshold Determinator (TD): This component uses the outputs of the PRE component to determine a threshold value that optimizes the sensitivity and specificity of the CCAD framework. The TP and TN values obtained from the PRE engine are used to determine the true positive rate (TPR) and true negative rate (TNR).

Lower threshold values yield higher true positive rate, but increase false positives (FP), which refer to the number of normal consumption patterns that are incorrectly identified. To find a threshold value that optimizes the trade-off between high true positive rate and low false positive rate (the proportion of normal consumption patterns incorrectly identified), the receiver operating characteristics (ROC) curve was explored. ROC is a plot in a unit square of the true positive rate (TPR) versus false positive rate (FPR) across varying threshold values. The ROC curve was chosen to analyze the performance of the anomaly detection model because, it considers all possible threshold values to evaluate both TPR and FPR of the CCAD framework. Several threshold determination techniques were explored.

4) Anomaly Classifier and Notifier: When new instances of data patterns are entered into the CCAD framework, their reconstruction error values are determined using the trained model in PLE. These values are then compared with the threshold, and patterns with a reconstruction error value greater than are classified as anomalous. Anomalous values trigger the notifier component to raise an alarm that notifies the building manager, who then performs appropriate procedures.

3.3. DayFilter[7]

The amount of sensor data generated by modern building systems is growing rapidly. Automatically discovering the structure of diurnal patterns in this data supports implementation of building commissioning, fault detection and retrofit analysis techniques. Additionally, these data are crucial to informing design professionals about the efficacy of their assumptions and strategies used in performance prediction simulation models. In this paper, we introduce DayFilter, a day-typing process that uses symbolic aggregate approximation (SAX), motif and discord extraction, and clustering to detect the underlying structure of building performance data.

Performance and energy data generation in the built environment is rapidly growing. Modern building controls and management systems are improving in their ability to acquire and store measured data as the technology improves. This phenomenon results in vast portfolios of collected data from heterogeneous buildings. Figure 1 illustrates a general example of various types of measured data from a conventional commercial building. Whole building performance is inuenced by layers of complex measurement systems. Aggregated performance metrics or sensors are often measured or calculated at each level of this hierarchy in order to condense the exponential detailed sensor data downstream.

DayFilter is an application of temporal data mining to building performance data. It includes five steps designed to incrementally filter structure from daily raw measured performance data.

1. Data pre-processing:

As in any data mining approach, data pre-processing is an important step to clean and standardize the data. In our approach, we first remove extreme point measurements that fall outside of three standard deviations, 3_, of the mean of the selected univariate data stream x(t). The data are then normalized in order to create a dataset, Z(t) with an approximate 0 mean and a standard deviation of close to 1.

2. Symbolic Aggregate appoXimation (SAX) transformation:

In the second step, we transform Z(t) into a symbolic representation using SAX. It is one of the many means of representing time-series data to enhance the speed and usability of various analysis techniques. SAX is a type of Piecewise Aggregate Approximation (PAA) representation and it has been used extensively in numerous applications.

In brief, the SAX transformation is as follows. The normalized time-series, Z(t), is first broken down into N individual non-overlapping subsequence. This step is known as chunking, and the period length N is based on a context-logical specific period.

Each chunk is then further divided into W equal sized segments. The mean of the data across each of these segments is calculated and an alphabetic character is assigned according where the mean lies within a set of vertical breakpoints. These breakpoints are calculated according to a chosen alphabet size.

3. Daily profile tagging and filtering

Once the SAX words are created, we are interested in visualizing each pattern and tagging each type as either a motif or discord. The results of applying the SAX process to a two-week sample power dataset. The diagram shows how each daily chunk of high frequency data is transformed into a set of SAX characters.

The more frequent patterns are categorized as motifs, or patterns which best describe the average behavior of the system. One can see the patterns with the lower frequencies and their indication as discords, or subsequence that are least common in the stream.

4. Clustering

After dividing the profiles into motif and discord candidates, we go on to cluster the motif candidates to create general daily performance phenotypes of the targeted data stream. This step is supplementary if the SAX transformation process produces too many motif candidates based on the input parameter settings. Clustering would be useful, for example, if 15 motif candidates are created and the user wants to further aggregate those candidates into 4 or 5 more typical profiles for simulation calibration purposes. This feature gives the user additional control to further aggregate the performance characterization, which can be useful when choosing large values of A or W in the SAX process. It should be noted that in some simplified cases or small datasets, this step may be redundant with SAX aggregation.

We use the k-means algorithm to cluster the daily profiles after removing the discord candidate day-types. This ensures load profile patterns that are not inuenced by the less frequent discords. Time series clustering can be approached as a raw-data-based, feature based, or model-based solution.

5. Expressive visualization for interpretation

As the final step, interpretation and visualization are important for DayFilter in order for a human analyst to visually extract knowledge from the results, and to make decisions regarding further analysis. We utilize insight from the Overview, zoom and filter, details-on-demand approach and the previously mentioned VizTree tool. The hidden structures of building performance data is revealed through the SAX process and we use visualization to communicate this structure to an analyst. The process uses a modified sankey diagram to visualize the augmented suffix tree in a way which the count frequency of each SAX word can be distinguished.

3.4. Shortcomings

One of the major shortcomings of the methods discussed above is the systems are not implemented in any real-life network. So, they are not tested products. And, their performance on the real network is not beyond question.

Another shortcoming is the discussed methods uses unsupervised learning technique. So, they have no learning growth. There is no way to increase the performance measure of the systems.

4. Our Approach

We are proposing an efficient model for energy wastage detection. In our proposed approach, we tried to overcome the shortcomings of the previous methods and tried to introduce new features.

The input of our proposed method comes from both smart meters and different sensors. For the analyzing, both historical sensor data and current occupancy based data is used. This will increase the reliability and efficiency of our system.

An important addition to our proposal is the addition of learning growth. The proposed model uses the reinforcement learning algorithm to improve performance. The feedback for reinforcement learning will come from the user. A user feedback tool has been introduced for this purpose.

The proposed approach will also have a network of actuators which will be used to prevent the anomaly. This is an interesting addition as most of the other methods only concentrated on the detection of anomaly.

5. Objectives

- Building a distributed wireless sensor network for data collection.
- Developing a system that can take intelligent decision analyzing the current and historical data
- Taking action based on the decision made by analyzing data
- Creating an interface in order to give the users the control to process of decision making.

6. Necessary Technologies

6.1. TellosB

TelosB Mote is developed by Memsic technology. It is an open-source platform designed to enable cutting-edge experimentation for the research community. It offers many features such as:

- IEEE 802.15.4 compliant RF transceiver
- 2.4 to 2.4835 GHz, a globally compatible ISM band
- 250 kbps data rate
- Integrated onboard antenna
- 8 MHz TI MSP430 microcontroller with 10kB RAM
- Low current consumption
- 1MB external flash for data logging

- Programming and data collection via USB Sensor suite including integrated light, temperature and humidity sensors

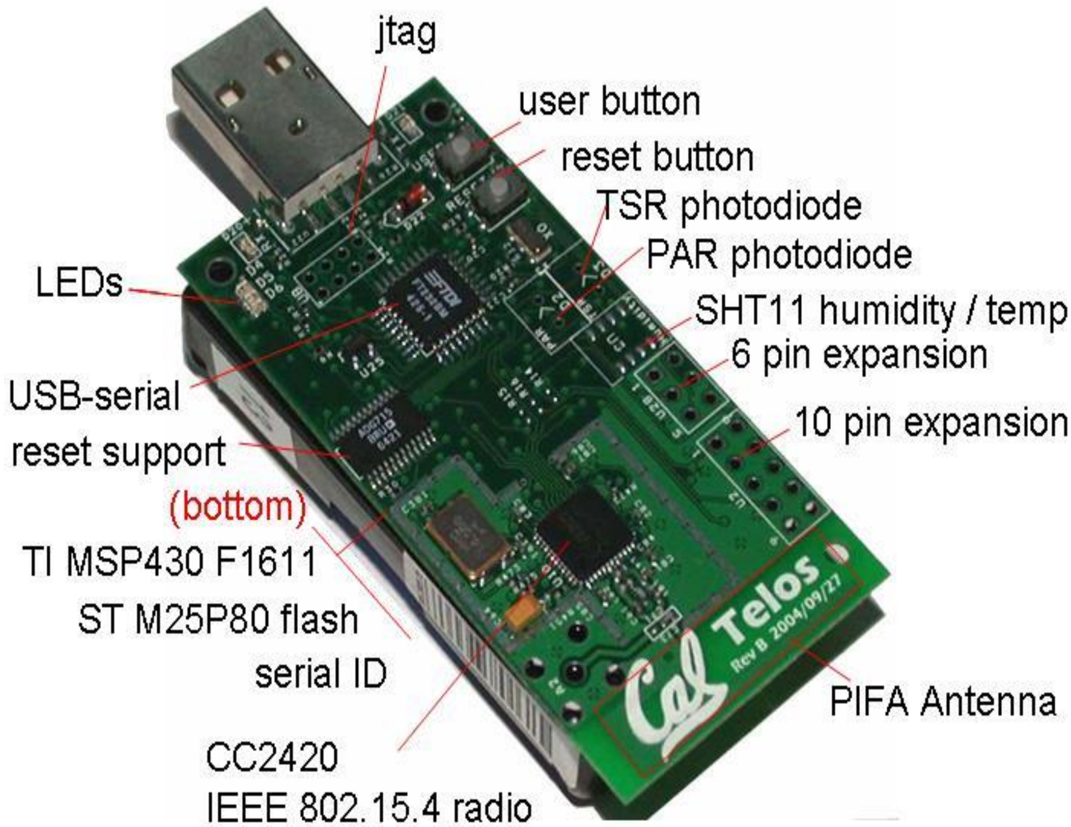

Fig 4: A TelosB mote

The TelosB platform was developed and published to the research community by UC Berkeley. This platform delivers low power consumption allowing for long battery life as well as fast wakeup from sleep state. It is powered by two AA batteries. If the mote is plugged into the USB port for programming or communication, power is provided from the host computer. It provides users with the capability to interface with additional devices. The two expansion connectors and on board jumpers may be configured to control analog sensors, digital peripherals and LCD displays.

6.2. Relay Switch

A relay is an electrically operated switch. Many relays use an electromagnet to mechanically operate a switch, but other operating principles are also used such as solid-state relays. Relays are used where it is necessary to control a circuit by a separate low-power signal, or where several circuits must be controlled by one signal. The first relays were used in long distance telegraph circuits as amplifiers. They repeated the signal coming in from one circuit and re-transmitted it on another circuit. Relays were used extensively in telephone exchanges and early computers to perform logical operations.

We used Relay Switches as digital switches to turn electronic devices on/off. The basic mechanism of relay is very simple. It has 3 ports:

- Normally Connected (NC)
- Common (Comm)
- Normally Open (NO)

The Comm port is normally connected with the NC port but when voltage is applied the Comm port gets attracted by the NO port and as a result the Comm port no longer keeps connected with the NC port rather it remains connected with the NO port. So if the circuit is built in such a way that one end of the circuit is connected with the Comm port another en is connected with the NO port then normally the circuit will be open and the device will be turned off but when power applied the circuit will be complete and the device will be turned on.

6.3. TinyOS

TinyOS is an embedded, component-based operating system and platform for low-power wireless devices, such as those used in wireless sensor networks (WSNs), smartdust, ubiquitous computing, personal area networks, building automation, and smart meters. Fundamentally, it is a work scheduler and a collection of drivers for microcontrollers and other ICs commonly used in wireless embedded platforms. It began as a collaboration between the University of California, Berkeley, Intel Research, and Crossbow Technology, was released as free and open-source software under a BSD license, and has since grown into an international consortium, the TinyOS Alliance.

TinyOS programs are built of software components, some of which present hardware abstractions. Components are connected to each other using interfaces. TinyOS provides interfaces and components for common abstractions such as packet communication, routing, sensing, actuation and storage.

TinyOS components offer three types of elements: Commands, Events and Tasks. These are basically normal C functions but they differ significantly in terms of who calls them and when they get called.

6.4. nesC

nesC (pronounced "NES-see") is a component-based, event-driven programming language used to build applications for the TinyOS platform. It is a dialect of the C language optimized for the memory limits of sensor networks. Its supplementary tools are mainly in the form of Java and shell script front-ends. Associated libraries and tools are mostly written in C.

nesC programs are built out of components, which are assembled to form whole programs. Components have internal concurrency in the form of tasks. Threads of control may pass into a component through its interfaces. These threads are rooted either in a task or a hardware interrupt.

In nesC, interfaces are bidirectional. They specify a set of functions to be implemented by the interface's provider (commands) and a set to be implemented by the interface's user (events). This allows a single interface to represent a complex interaction between components. This is critical because all lengthy commands in TinyOS are non-blocking; their completion is signaled through an event (send done). By specifying interfaces, a component cannot call the send command unless it provides an implementation of the sendDone event. Typically commands call downwards, i.e., from application components to those closer to the hardware, while events call upwards. Certain primitive events are bound to hardware interrupts.

7. Implementation

7.1. Data collection

In this phase sensor data such as the temperature or the ambient light intensity is collected from the sensors. This layer consists of TelosB motes. Necessary number of motes were scattered throughout the building. Each mote is powered by 2 batteries (each battery with 1.5v power). We used light and temperature sensor of TelosB for sensing light and temperature and we added external sensors to the expansion ports of TelosB motes for sensing other things such as human presence or movement. Collected data was directly sent to the processing layer through the internal radio of the TelosB motes without any kind of processing. This helps to save the power of the motes.

7.2. Networking

The sensed data is sent by the sensor node to a central sink node. All the sensors in a system sends the data to the same sink. Nodes send data to the receiver without any processing.

The distance between sender and receiver may be too large. As sensors are low powered devices, their signal strength is not too high. So, relays can be used at the suitable positions to get the correct signal.

7.2.1. Sensor Medium Access Control (S-MAC) protocol

This protocol tries to reduce energy consumption due to overhearing, idle listening and collision. In this protocol also every node has two states, sleep state and active state. A node can receive and transmit data during its listen period.

The main concept in SMAC is that, all the neighbouring nodes form virtual clusters and synchronize their sleep and listen periods. They communicate during their listen periods and sleep rest of the time. The immediate neighbours of nodes, which are transmitting and receiving, sleep until the communication is completed. A long message is divided into many fragments and all the fragments are sent as burst. S-MAC contributes in these ways; reduction of idle listening, collision and overhearing avoidance by using RTS and CTS and saving energy and time by sending a series of fragments of a long message together, rather than going for contention after sending every fragment.

7.3. Data processing

At this phase newly collected data is processed at the sink node. This sink mote is connected to a processing device (A desktop computer in this case) through a USB port. The sink mote gets power from the computer through the USB port and also sends all the data to the computer via Serial Communication. The computer then processes all the data according to some pre-set parameters and threshold values and decides which electronic device should be turned on and which should be turned off. After taking decisions the computer tells the sink mote to send on signal to the TelosB motes.

7.3.1. Reinforcement Learning Algorithm

Reinforcement learning is a type of dynamic machine learning algorithm that uses the system of reward and punishment.

A reinforcement learning algorithm, or agent, learns by interacting with its environment. The agent receives rewards by performing correctly and penalties for performing incorrectly. The agent learns without intervention from a human by maximizing its reward and minimizing its penalty.

In our system, we used brute force approach for reinforcement learning. The brute force approach entails two steps:

- For each possible policy, sample returns while following it
- Choose the policy with the largest expected return

One problem with this is that the number of policies can be large, or even infinite. Another is that variance of the returns may be large, which requires many samples to accurately estimate the return of each policy.

These problems can be ameliorated if we assume some structure and allow samples generated from one policy to influence the estimates made for others. The two main approaches for achieving this are value function estimation and direct policy search.

7.4. Actuators

This phase works with the actuators. Different relays are used to act as the actuators. This actuators act based on the signals sent from the processor. Actuators are attached with different electronic devices. The relays can turn these devices on or off.

If any electronic device needs to be turned on or off is decided by the processor phase. The sink mote sends 1 as the on signal and 0 as the off signal along with the coded name of the respective electronic devices to the TelosB motes of the electronics layers based on the decision taken by the computer.

7.5. User Interface

This is the final layer of the system architecture. The main purpose of this layer is to providing the users an overview and the status of the whole system at any given time. This layer also gives the users some control to set some parameters and threshold based on which the Processing layer will take decision.

The interface is made in Java programming language. The interface contains some labels that shows the values of the reading of all the sensor. There are also some text box where the users can input a particular threshold value.

The User Interface layer communicates with the Processing layer through a process which is called Serial Forwarding. By Serial Forwarding a TelosB mote can set and get data of the java program.

8. Future Improvements

- Real-life implementation
- Improving learning algorithm
- User interface improvement
- Improving efficiency

9. Conclusion

We want to reduce the wastage of energy consumption through our system in a smart way. We will take our decision based on collective data. Among the related approaches our proposed approach brings the real life implementation through actuator network system. Statistical analysis will bring the output of our system to the users which will lead a better connection establishment with the users. As there would be user feedback tool so there are more scopes for the users to contribute to the system and control it as well. We hope a better improvement of our system in near future.

References

1. http://www.buildings.com/article-details/articleid/19537/title/how-smart-buildings-save-energy.aspx
2. American Council for an Energy Efficient Economy.
3. http://timesofindia.indiatimes.com/india/Three-billion-units-of-power-wasted-in-one-year/articleshow/47942237.cms
4. https://en.m.wikipedia.org/wiki/Electricity_sector_in_Bangladesh
5. R. Fontugne, J. Ortiz, N. Tremblay, P. Borgnat, P. Flandrin, K. Fukuda, D. Culler, and H. Esaki, “Strip, bind, and search : a Method for identifying abnormal energy consumption in buildings,” 12th International Conference on Information Processing in Sensor Networks, pp. 129–140, 2013.
6. Araya, Daniel B., et al. "Collective contextual anomaly detection framework for smart buildings." *Neural Networks (IJCNN), 2016 International Joint Conference on*. IEEE, 2016.
7. Miller, Clayton, Zoltán Nagy, and Arno Schlueter. "Automated daily pattern filtering of measured building performance data." *Automation in Construction* 49 (2015): 1-17.