Strip, Bind, and Search: A Method for Identifying Abnormal Energy Consumption in Buildings

Romain Fontugne[1,5], Jorge Ortiz[2], Nicolas Tremblay[3], Pierre Borgnat[3]
Patrick Flandrin[3], Kensuke Fukuda[4], David Culler[2] and Hiroshi Esaki[1]

[1]The University of Tokyo [2]University of California, Berkeley
[3]CNRS, Ecole Normale Supérieure de Lyon [4]National Institute of Informatics
[5]Japanese-French Laboratory for Informatics

ABSTRACT

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. In 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. It then flags deviations from the model. We run SBS on a set of building sensor traces; each containing hundred sensors reporting data flows over 18 weeks from two separate buildings with fundamentally different infrastructures. We demonstrate that, in many cases, SBS uncovers misbehavior corresponding to inefficient device usage that leads to energy waste. The average waste uncovered is as high as 2500 kWh per device.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

Keywords

Building; Energy Consumption; Anomaly Detection

1. INTRODUCTION

Buildings are one of the prime targets to reduce energy consumption around the world. In the United States, the second largest energy consumer in the world, buildings account for 41% of the country's total energy consumption [26]. The first measure towards reducing the building's energy consumption is to prevent electricity waste due to the improper use of the buildings equipment.

IPSN'13, April 8–11, 2013, Philadelphia, Pennsylvania, USA.

Large building infrastructure is usually monitored by numerous sensors. Some of these sensors enable building administrators to view device power-draw in real time. This allows administrators to determine proper device behavior and system-wide inefficiencies. Detecting misbehaving devices is crucial, as many are sources of energy waste. However, identifying these saving opportunities is impractical for administrators because large buildings usually contain hundreds of monitored devices producing thousands of streams and it requires continuous monitoring. As such, the goal of this work is to establish a method that automatically reports abnormal device-usage patterns to the administrator by closely examining all of the continuous power streams.

The intuition behind the proposed approach is that each service provided by the building requires a minimum subset of devices. The devices within a subset are used at the same time when the corresponding service is needed and a savings opportunity is characterized by the partial activation of the devices. For example, office comfort is attained through sufficient lighting, ventilation, and air conditioning. These are controlled by the lighting and HVAC (Heating, Ventilation, and Air Conditioning) system. Thus, when the room is occupied both the air conditioner (heater on a cold day) and lights are used together and should be turned off when the room is empty. In principal, if a person leaves the room and turns off *only* the lights then the air conditioner (or heater) is a source of electricity waste.

Following this basic idea we propose *Strip, Bind and Search* (SBS), an unsupervised methodology that systematically detects electricity waste. Our proposal 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. This allows us to differentiate between devices that are used together (high correlation), used independently (no correlation), and used mutually exclusively (negative correlation).

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. Such abnormalities usually present an opportunity to reduce electricity waste or events that deserve careful attention (e.g. faulty device).

SBS overcomes several challenges. First, noisy sensor traces that all share a similar trend, making direct correlation analysis non-trivial. Device energy consumption is mainly driven by occupancy and weather, all the devices display a similar daily pattern, in roughly overlapping time intervals and phases. Therefore, one of the main contributions of this work is uncovering the intrinsic device relationships by filtering out the dominant trend. For this task we use Empirical Mode Decomposition [14], a known method for de-trending time-varying signals.

Another key contribution of this work is in using SBS to practically monitor building energy consumption. Moreover, the proposed method is easy to use and functions in any building, as it does not require prior knowledge of the building nor extra sensors. It is also tuned through a single intuitive parameter.

We validate the effectiveness of our approach using 10 weeks of data from a modern Japanese building containing 135 sensors and 8 weeks of data from an older American building containing 70 sensors. These experiments highlight the effectiveness of SBS to uncover device relationships in a large deployment of 135 sensors. Furthermore, we inspect the SBS results and show that the reported alarms correspond to significant opportunities to save energy. The major anomaly reported in the American building lasts 18 days and accounts for a waste of 2500 kWh. SBS also reports numerous small anomalies, hidden deep within the building's overall consumption data. Such errors are very difficult to find without SBS.

In the rest of this paper, we detail the mechanisms of SBS (Section 3) before evaluating it with real data (Section 5) then we discuss different outcomes of the proposed methodology (Section 7) and conclude.

2. PROBLEM DESCRIPTION

The primary objective of SBS is to determine *how* device usage patterns are correlated across all pairs of sensors and discover when these relationships change. The naive approach is to run correlation analysis on pairs of sensor traces, recording their correlation coefficients over time and examining when there is a statistically-significant deviation from the norm. However, this approach does not yield any useful information when applied to *raw data traces*. For example, the two raw signals shown in Figure 3 are from two independent HVAC systems, serving different rooms on different floors. Since each space is independently controlled, we expect their power-draw signals to be uncorrelated (or at least distinguishable from other signal pairs). However, their correlation coefficient (0.57), is not particularly informative – it is statistically similar to the correlation between itself and other signals in the trace.

Using a larger set of devices, Figure 1 shows a correlation matrix with 135 distinct lighting and HVAC systems serving numerous rooms in a building (described later on in Section 4.1). The indices are selected such that their index-difference is indicative of their relative spatial proximity. For example, a device in location 1 is closer in the building to a device in location 2 than it is to a device in location 135. The color of the cell is the average pairwise correlation coefficient for devices in the row-column index. The higher the value,

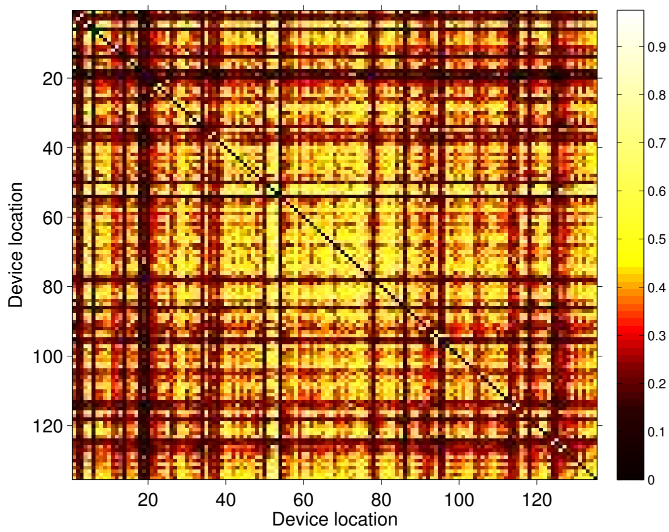


Figure 1: Correlation coefficients of the raw traces from the Building 1 dataset (Section 4.1). The matrix is ordered such as the devices serving same/adjacent rooms are nearby in the matrix.

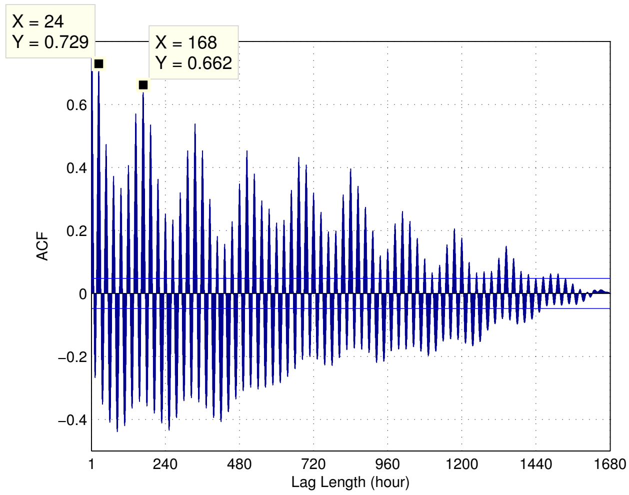


Figure 2: Auto-correlation of a usual signal from the Building 1 dataset. The signal features daily and weekly patterns (resp. $x = 24$ and $x = 168$).

the lighter the color. Devices serving the same room are along the diagonal. Because these devices are used simultaneously, we expect high average correlation scores, lighter shades, along the diagonal figure. However, we observe no such pattern. Most of the signals are correlated with all the others and we see no discernible structure.

An explanation for this is that the daily occupant usage patterns drive these results. Figure 3 demonstrates this more clearly. It shows two 1-week raw signals traces which feature the same diurnal pattern. This trend is present in almost every sensor trace, and, it hides the smaller fluctuations providing more specific patterns driven by local occupant activity. Upon deeper inspection, we uncovered several dominant patterns, common among energy-consuming devices in buildings [27]. Figure 2 depicts the auto-correlation of a usual electric power signal for a device. The two highest values in the figure correspond to a lag of 24 hours and 168 hours (one week). Therefore, the signal has some periodicity and similar (though not equal) values are seen at daily and weekly time scales. The daily pattern is due to daily office hours and the weekly pattern corresponds to weekdays and weekends. Correlation analysis on *raw* signals cannot be used to determine meaningful inter-device relationships because periodic components act as non-stationary trends for high-frequency phenomenon, making the correlation function irrelevant. Such trends must be removed in order to make meaningful progress towards our aforementioned goals.

In the next section we describe SBS. We discuss *strip and bind* in section 3.1, which addresses de-trending and relationship-discovery. Then, we describe how we *search* for changes in usage patterns, in section 3.2, to identify potential savings opportunities.

3. METHODOLOGY

3.1 Strip and Bind

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), that reveals the signal patterns at different frequency bands. IMFs are obtained using Empirical Mode Decomposition (see Figure 3 and Section 3.1.1). We only consider IMFs with time scales shorter than a day, since we are interested in capturing short-scale usage patterns. Consequently, SBS aggregates the IMFs that fall into this specific time scale (see *IMF agg.* in Figure 3). The resulting partial signals of different device power traces are compared, pairwise, to identify the devices that show un/correlated usage patterns (see *Corr. Coeff.* in Figure 3).

3.1.1 Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) [14] 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 [20], medical data [4], speech signals [12, 11], and image processing [21]. 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 defined 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 [14]; 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.

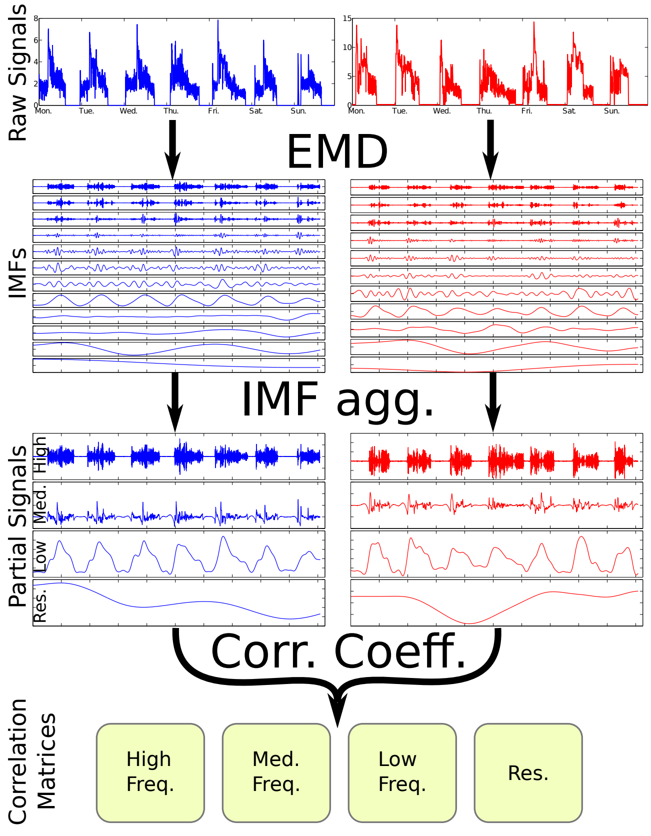


Figure 3: *Strip and Bind* using two raw signals standing for one week of data from two different HVACs. (1) Decomposition of the signals in IMFs using EMD (top to bottom: c_1 to c_n); (2) aggregation of the IMFs based on their time scale; (3) comparison of the partial signals (aggregated IMFs) using correlation coefficient.

We formally describe the EMD algorithm as follows:

1. Sifting process: For a current signal $h_0 = X$, let m_0 be the mean of its upper and lower envelopes as determined from a cubic-spline interpolation of local maxima and minima.

2. The estimated local mean m_0 is removed from the signal, giving a first component: $h_1 = h_0 - m_0$

3. The sifting process is iterated, h_1 taking the place of h_0. Using its upper and lower envelopes, a new local mean m_1 is computed and $h_2 = h_1 - m_1$.

4. The procedure is repeated k times until $h_k = h_{k-1} - m_{k-1}$ is an IMF according to the two conditions above.

5. This first IMF is designated as $c_1 = h_k$, and contains the component with shortest periods. We extract it from the signal to produce a residual: $r_1 = X - c_1$. Steps 1 to 4 are repeated on the residual signal r_1, providing IMFs c_j and residuals $r_j = r_{j-1} - c_j$, for j from 1 to n.

6. The process stops when residual r_n contains no more than 3 extrema.

The result of EMD is a set of IMFs c_i and the final residue r_n, such as:

$$X = \sum_{i=1}^{n} c_i + r_n$$

where the size of the resulting set of IMFs (n) depends on the original signal X and r_n represents the trend of the data (see *IMFs* in Figure 3).

For this work we implemented a variant of EMD called Complete Ensemble EMD [25]. This algorithm computes EMD several times with additional noise, it allows us to efficiently analyze signals that have flat sections (i.e. consuming no electricity in our case).

3.1.2 IMF aggregation

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.

However, comparing the IMFs obtained from different signals is also not trivial, because EMD is empirically uncovering IMFs from the data there is no guarantee that the two IMFs c_i^1 and c_i^2 obtained from two distinct signals S^1 and S^2 represent data at the same frequency domain. Directly comparing c_i^1 and c_i^2 is meaningless unless we confirm that they belong to the same frequency domain.

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 (round dots in Figure 4). Formally, given a data point p, GZC measures the quarter (T_4), the two halves (T_2^x), and the four full periods (T_1^y), p belong to (see Figure 4) and the instantaneous period is computed as:

$$T = \frac{1}{7}\{4T_4 + (2T_2^1 + 2T_2^2) + (T_1^1 + T_1^2 + T_1^3 + T_1^4)\}$$

Since all points p between two critical points have the same instantaneous period GZC is local down to a quarter period. 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.

The IMFs are clustered using four time scale ranges:

- The *high frequencies* are all the IMFs with a time scale lower than 20 minutes. These IMFs capture the noise.

- The *medium frequencies* are all the IMFs with a time scale between 20 minutes and 6 hours. These IMFs convey the detailed devices usage.

- The *low frequencies* are all the IMFs with a time scale between 6 hours and 6 days. These IMFs represent daily device patterns.

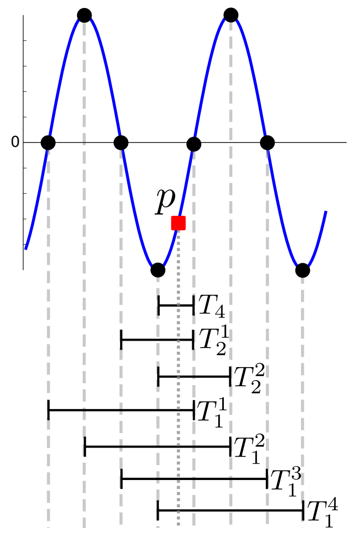


Figure 4: Generalized Zero Crossing: the local mean period at the point p is computed from one quarter period T_4, two half periods T_2^x and four full periods T_1^y (where $x = 1, 2$, and, $y = 1, 2, 3, 4$).

- The *residual data* is all data with a time scale higher than 6 days. This is mainly residual data obtained after applying EMD. Also, it highlights the main device trend.

These time scale ranges are chosen based on our experiments and goal. The 20-minute boundary relies on the sampling period of our dataset (5 minutes) and permits us to capture IMFs with really short periods. The 6-hour boundary allows us to analyze all patterns that have a period shorter than the usual office hours. The 6-day boundary allows us to capture daily patterns and weekday patterns.

Aggregating IMFs, within each time scale range, results in 4 partial signals representing different characteristics of the device's energy consumption (see *Partial Signals* in Figure 3). We do a pairwise device trace comparison, calculating the correlation coefficient of their partial signals. In the example shown in Figure 3, the correlation coefficient of the raw signals suggests that they are highly correlated (0.57). However, the comparison of the corresponding *partial signals* provides new insights; the two devices are poorly correlated at high and medium frequencies (respectively −0.01 and −0.04) but highly correlated at low frequencies (0.79) meaning that these devices are not "intrinsically" correlated. They only share a similar daily pattern.

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.

3.2 Search

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.

3.2.1 Reference

We define four reference matrices, which capture normal device behavior at the four time scale ranges defined in Section 3.1.2. The references are computed as follows: (1) we retrieve the correlation matrices for n consecutive time bins. (2) For each pair of devices we compute the median correlation over the n time bins and obtain a matrix of the median device correlations.

Formally, for each time scale range the computed reference matrix for d devices and n time bins is:

$$R_{i,j} = \text{median}(C^1_{i,j}, ..., C^n_{i,j})$$

where i and j ranges in $[1, d]$.

Because anomalies are rare by definition, we assume the data used to construct the reference matrix is an accurate sample of the population; it is unbiased and accurately captures the range of normal behavior. Abnormal correlation values, that could appear during model construction, are ignored by the median operator thanks to its robustness to outlier (50% breakdown point). However, if that assumption does not hold (more than 50% of the data is anomalous), our model will flag the opposite – labeling abnormal as normal and vice-versa. From close inspection of our data, we believe our primary assumption is sound.

3.2.2 Behavior change

We compare each device behavior, for all time bins, to the one provided by the reference matrix. Consider the correlation matrix C^t obtained from the data for time bin t ($1 \leq t \leq n$). Vector $C^t_{i,*}$ is the behavior of the i^{th} device for this time bin. Its normal behavior is given by the corresponding vector in the reference matrix $R_{i,*}$. We measure the device behavior change at the time bin t with the following Minkowski weighted distance:

$$l^t_i = \left(\sum_{j=1}^{d} w_{ij} \left(C^t_{i,j} - R_{i,j} \right)^p \right)^{1/p}$$

where d is the number of devices and w_{ij} is:

$$w_{ij} = \frac{R_{i,j}}{\sum_{k=1}^{d} R_{i,k}}.$$

The weight w enables us to highlight the relationship changes between the device i and those highly correlated to it in the reference matrix. In other words, our definition of behavior change is mainly driven by the relationship among devices that are usually used in concert. We also set $p = 4$ in order to inhibit small differences between $C^t_{i,j}$ and $R_{i,j}$ but emphasize the important ones.

By monitoring this quantity over several time bins the abnormal device behaviors are easily identified as the outlier values. In order to identify these outlier values we implement a robust detector based on median absolute deviation (MAD), a dispersion measure commonly used in anomaly detection [16, 7]. It is a measure that robustly estimates the variability of the data by computing the median of the absolute deviations from the median of the data. Let $l_i = [l^1_i, ..., l^n_i]$ be a vector representing the behavior changes of device i over n time bins, then its MAD value is defined as:

$$\text{MAD}_i = b\,\text{median}(|l_i - \text{median}(l_i)|)$$

where the constant b is usually set to 1.4826 for consistency with the usual parameter σ for Gaussian distributions. Consequently, we define anomalous behavior, for device i at time t, such that the following equation is satisfied:

$$l^t_i > \text{median}(l_i) + \tau\,\text{MAD}_i$$

Note, τ is a parameter that permits to make SBS more or less sensitive.

The final output of SBS is a list of alarms in the form (t, i) meaning that the device i has abnormal behavior at the time bin t. The priority of the alarms in this list is selected by the building administrator by tuning the parameter τ.

4. DATA SETS

We evaluate SBS using data collected from buildings in two different geographic locations. One is a new building on main campus of the University of Tokyo and the other is an older building at the University of California, Berkeley.

4.1 Engineering Building 2 - Todai

Engineering building 2, at the University of Tokyo (Todai), is a 12-story building completed in 2005 and is now hosting classrooms, laboratories, offices and server rooms. The electricity consumption of the lighting and HVAC systems of 231 rooms is monitored by 135 sensors. Rather than a centralized HVAC system, small, local HVAC systems are set up throughout the buidling. The HVAC systems are classified into two categories, EHP (Electrical Heat Pump) and GHP (Gas Heat Pump). The GHPs are the only devices that serve numerous rooms and multiple floors. The 5 GHPs in the dataset serve 154 rooms. The EHP and lighting systems serve only pairs of rooms and which are directly controlled by the occupants. In addition, the sensor metadata provides device-type and location information (room number), therefore, the electricity consumption of each pair of rooms is separately monitored.

The dataset contains 10 weeks of data starting from June 27, 2011 and ending on September 5, 2011. This period of time is particularly interesting for two reasons: 1) in this region, the summer is the most energy-demanding season and 2) the building manager actively works to curtail energy usage as much as possible due to the Tohoku earthquake and Fukushima nuclear accident.

Furthermore, this dataset is a valuable ground truth to evaluate the Strip and Bind portions of SBS. Since the light and HVAC of the rooms are directly controlled by the room's occupants, we expect SBS to uncover verifiable devices relationships.

4.2 Cory Hall - UC Berkeley

Cory Hall, at UC Berkeley, is a 5-story building hosting mainly classrooms, meeting rooms, laboratories and a datacenter. This building was completed in 1950, thus its infrastructure is significantly different from the Japanese one. The HVAC system in the building is centralized and serves several floors per unit. There is a separate unit for an internal fabricated laboratory, inside the building.

This dataset consists of 8 weeks of energy consumption traces measured by 70 sensors starting on April 5^{th}, 2011. In contrast to the other dataset, a variety of devices are monitored, including, electric receptacles on certain floors, most of the HVAC components, power panels and whole-building consumption.

These two building infrastructures are fundamentally different. This enables us to evaluate the practical efficacy of the proposed, unsupervised method in two very different environments.

4.3 Data pre-processing

Data pre-processing is not generally required for the proposed approach. Nevertheless, we observe in a few exceptional cases that sensors reporting excessively high values (i.e. values higher than the device actual capacity) that greatly alter the performance of SBS by inducing a large bias in the computation of the correlation coefficient. Therefore, we remove values that are higher than the maximum capacity of the devices, from the raw data.

5. EXPERIMENTAL RESULTS

In this section we evaluate SBS on our building traces. We demonstrate the benefits of striping the data by monitoring patterns captured at different time scales. Then, we thoroughly investigate the alarms reported by SBS.

5.1 Shortcomings

Because our analysis is done on historical data, some of the faults found by SBS could not be fully corroborated. In order to fully examine the effectiveness of our approach, we must run it in real time and physically check that the problem is actually occurring. When a problem is detected in the historical trace, months after it has occurred, the current state of the building may no longer reflect what is in the traces. Some of the anomalies discussed in this section uncover interpretable patterns that are difficult to find in practice. For example, simultaneous heating and cooling is a known, recurring problem in buildings, but it is very hard to identify when it is occurring. Some of the anomalies we could not interpret might be interpretable by a building manager, however, we did not consult either building manager for this study. Therefore, the results of this study do not examine the true/false positive rate exhaustively.

The true/false negative rate is impractical to assess. It may be examined through synthetic stimulation of the building via the control system. However, getting cooperation from a building manager to hand over control of the building for experimentation in non-trivial. Therefore, we forgo a full true/false negative analysis in our evaluation.

Because of these challenges, the evaluation of SBS focuses on comparing the output with known fault signatures. We examine anomalies, in either building, where the anomaly is easily interpretable but difficult to find by the building manager. We forego a comparison of SBS with competing algorithms because related algorithms require detailed knowledge of the building, *a priori*. The advantage of SBS is that it requires no such information to provide immediate value.

5.2 Device behavior at different time scales

The Strip and Bind part of SBS is evaluated using the data from Eng. Bldg 2. This dataset is appropriate to measure SBS's performance, since lighting and HVAC systems serving the same room are usually used simultaneously. Consequently, we analyze this data using SBS and verify that the higher correlations at medium frequencies correspond to devices located in the same room.

The dataset is split into 10, one-week bins and each bin is processed by SBS. Using the 10 correlation matrices at each time scale range, SBS uncovers the four reference matrices depicted in Figure 5.

High frequencies.

In this work the high frequencies correspond to the signals *noise*, therefore, we do not expect any useful information from the corresponding matrix (Figure 5a). Indeed, the corresponding reference matrix does not provide any help to determine a device's relative location. Thus, we emphasize that high frequency data should be ignored for uncovering device relationships (in contrast to [10]). Interestingly, we find that the sensors monitoring the lights generate consistent noise.

Medium frequencies.

Our main focus is on the medium frequencies as it is designed to capture the intrinsic device relationships. Figure 5b shows the correlation matrix at medium frequencies. It is significantly different from the one obtained with the raw signals (Figure 1): high correlation coefficients are concentrated along the matrix diagonal. Since devices serving the same or adjacent rooms are placed nearby in the matrix it validates our hypothesis: *high correlation scores within the medium frequency band shows strong inter-device relationships.*

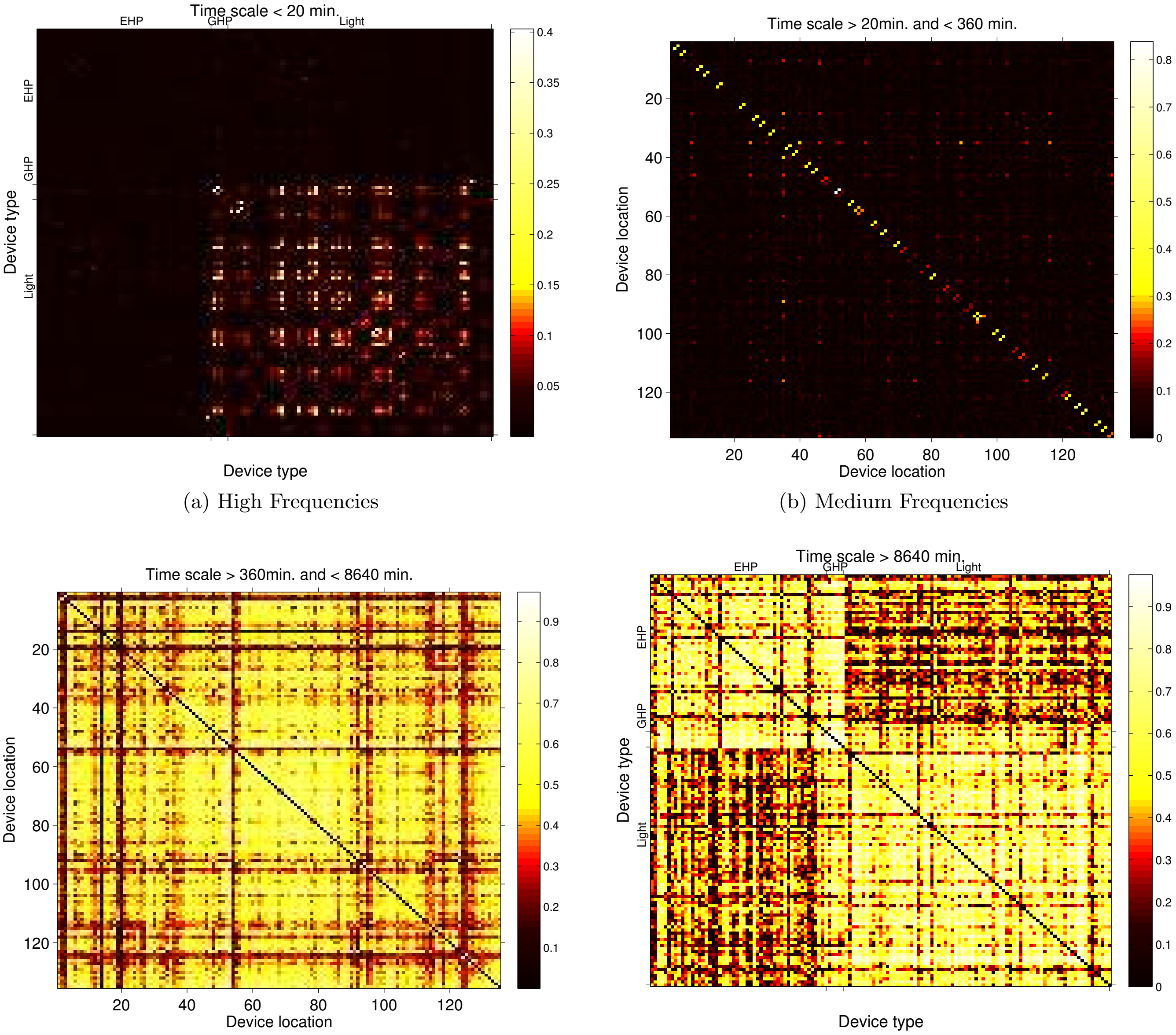
Considering this reference matrix as an adjacency matrix of a graph, in which the nodes are the devices, we identify the clusters of correlated devices using a community mining algorithm [5]. As expected we obtain mainly clusters of only two devices (light and HVAC serving the same room), but we also find clusters that are composed of more devices. For example a cluster contains 3 HVAC systems serving the three server rooms. Although these server rooms are located on different floors, SBS shows a strong correlation between these devices. Coincidentally, they are managed similarly. Interestingly, we also observe a couple of clusters that consist of independent devices serving adjacent rooms belonging to the same lab. The bigger cluster contains 33 devices that are 2 GHP devices and the corresponding lights. This correlation matrix and the corresponding clusters highlight the ability for SBS to identify such hidden inter-device usage relationships.

Low frequencies.

Low frequencies capture daily patterns, embedded in all the device traces. Figure 5c depicts the corresponding reference matrix which is similar to the one of raw signal traces (Figure 1) and it shows no particular structure.These partial signals are discarded as they do not help us in identifying inter-device usage patterns.

Residual data.

The residual data shows the weekly trend, which gives us no information about device relationships. But, surprisingly, by reordering the correlation matrix based on the type



(a) High Frequencies

(b) Medium Frequencies

(c) Low Frequencies

(d) Residual data

Figure 5: Reference matrices for the four time scale ranges (the diagonal $x = y$ is colored in black for better reading). The medium frequencies highlight devices that are located next to each other thus intrinsically related. The low frequencies contains the common daily pattern of the data. The residual data permits to visually identify devices of the similar type.

of the devices (Figure 5d) we can visually identify two major clusters. The first cluster consists of HVAC devices (see EHP and GHP in Figure 5d) and the second one contains only lights. An in-depth examination of the data reveals that long-term trends are inherent to the device types. For example, as the consumption of both the EHP and GHP devices is driven by the building occupancy and the outside temperature, these two types of devices follow the same trend. However, the use of light is independent from the outside temperature thus the lighting systems follow a common trend different from the EHP and GHP one.

We conduct the same experiments by splitting the dataset in 70 bins of 1 day long and observe analogous results at high and medium frequencies but not at lower frequencies. This is because the bins are too short to exhibit daily oscillations and the residual data captures only the daily trend.

5.3 Anomalies

We evaluate the *search* performance of SBS using the traces from the Eng. Bldg 2 and Cory Hall. Due to the lack of historical data, such as room schedule or reports of energy waste, the evaluation is non-trivial. Furthermore, getting ground truth data from a manual inspection of the hundreds traces of our data sets is impractical. The lack of ground truth data prevents us from producing a systematic

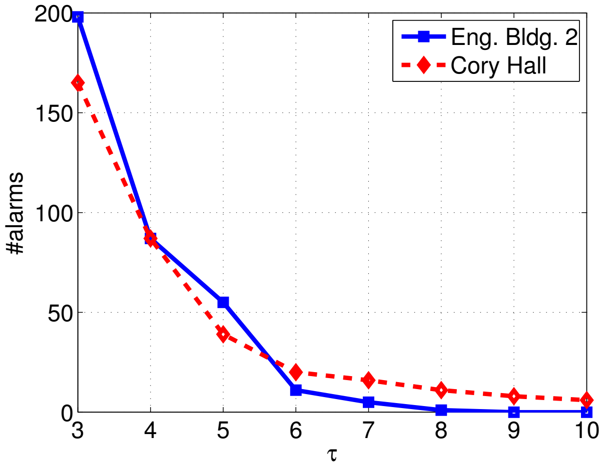


Figure 6: Number of reported alarms for various threshold value ($\tau = [3, 10]$).

	High	Low	Punc.	Missing	Other
Eng. Bldg 2	9 (5)	6 (5)	1 (1)	36 (1)	3 (3)
Cory Hall	25 (7)	7 (3)	4 (4)	0 (0)	3 (3)

Table 1: Classification of the alarms reported by SBS for both dataset (and the number of corresponding anomalies).

analysis of the anomalies missed by SBS (i.e. false negatives rate). Nevertheless, we exhibit the relevance of the anomalies uncovered by SBS (i.e. high true positive rate and low false positive rate) by manually checking the output of SBS.

Anomaly classification.

To validate SBS results we manually inspect the anomalies detected by the algorithm. For each reported alarm (t, i) we investigate the device trace i and the devices correlated to it to determine the reason for the alarm. Specifically, we retrieve the major relationship change that causes the alarm (i.e. $\max(|w_j(C^t_{i,j} - R_{i,j})|)$, see Section 3.2) and examine the metadata associated to the corresponding device. This investigation allows us to classify the alarms into five groups:

- *High power usage*: alarms corresponding to electricity waste.
- *Low power usage*: alarms representing the abnormally low electricity consumption of a device.
- *Punctual abnormal usage*: alarms standing for short term (less than 2.5 hours) raise or drop of the electricity consumption.
- *Missing data*: alarms raised due to a sensor failure.
- *Other*: alarms whose root cause is unclear.

Experimental setup.

For each experiment, the data is split in time bins of one day, starting from 09:00 a.m. – which is approximately the office's opening time. We avoid having bins start at midnight since numerous anomalies appear at night and they are better highlighted if they are not spanning two time bins. Only the data at medium frequencies are analyzed, the other frequency bands are ignored, and the reference matrix is computed from all time bins.

The threshold τ tunes the sensitivity of SBS, hence, the number of reported alarms. Furthermore, by plotting the number of alarms against the value of τ for both datasets (Figure 6) we observe an elbow in the graph around $\tau = 5$. With thresholds lower than this pivot value ($\tau < 5$), the number of alarms significantly increases, causing less important anomalies to be reported. For higher values ($\tau > 5$), the number of alarms is slowly decreasing, providing more conservative results that consist of the most important anomalies. This pivot value provides a good trade off for either data set.

Table 1 classifies the alarms reported by SBS on both datasets. Anomalies spanning several time bins (or involving several devices) may raise several alarms. We display these in Table 1 as numbers in brackets – the number of anomalies corresponding to the reported alarms.

5.3.1 *Engineering Building 2*

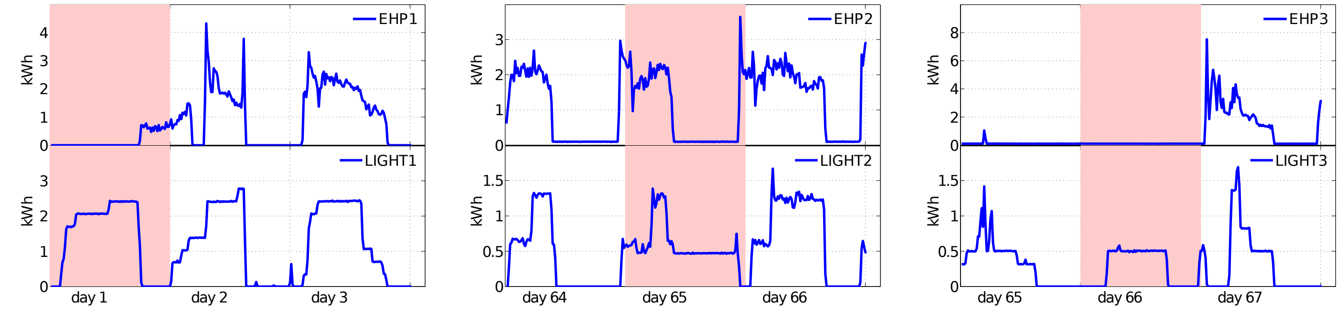
SBS reported 55 alarms over the 10 weeks of the Eng. Bldg 2 dataset. However, 36 alarms are set aside because of sensor errors; one GHP has missing data for the first 18 days. Since this device is highly correlated to the GHP in the reference matrix, their relationship is broken for the 18 first bins and for each bin one alarm per device is raised.

In spite of the post-Fukushima measures to reduce Eng. Bldg 2's energy consumption, SBS reported 9 alarms corresponding to high power usage (Table 1). Figure 7a depicts the electricity consumption of the light and EHP in the same room where two alarms are raised. Because the EHP was not used during daytime (but is turned on at night, when the light is turned off) the relationship between the two devices is "broken" and an alarm is raised for each device. Figure 7b shows another example of energy waste. The light is on at night and the EHP is off. The top-priority anomaly reported by SBS is caused by the 10 days long constant use of an EHP (Figure 7d) and this waste of electricity accounts for 165 kWh. SBS partially reports this anomaly but lower values of τ permits us to identify most of it.

We observed 6 alarms corresponding to abnormally low power use. Upon further inspection we notice that it corresponds to energy saving initiatives from the occupants – likely due to electricity concerns in Japan. This behavior is displayed in Figure 7c. The room is occupied at the usual office hours (indicated by light usage) but the EHP is not on in order to save electricity.

5.3.2 *Cory Hall*

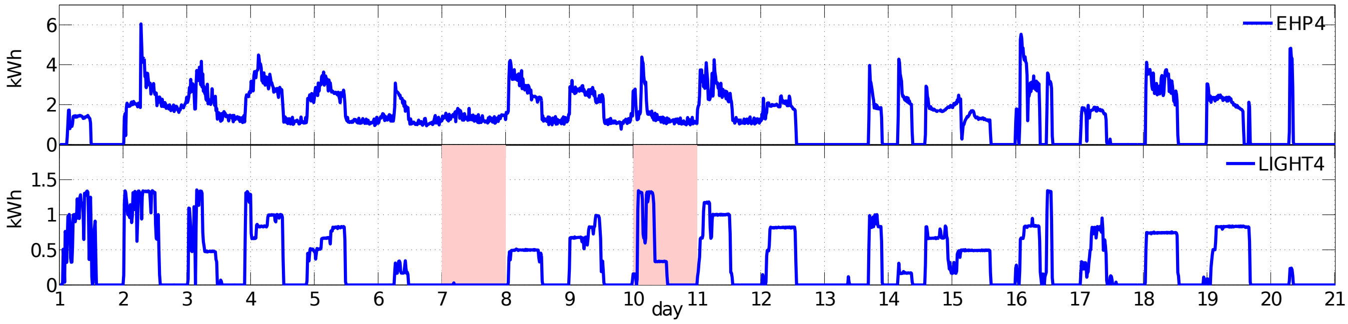
SBS reported 39 alarms for the Cory Hall dataset (Table 1). 7 are classified as low power usage, however, our inspection revealed that the root causes are different than for the Eng. Bldg 2 dataset. We observe that the low power usage usually corresponds to device failures or misconfiguration. For example, Figure 8a depicts the electricity consumption of the 2^{nd} floor chiller and a power riser that comprises the consumption of multiple systems, including the chiller. As the chiller suddenly stops working, the correlation between



(a) High power usage where the HVAC (EHP) is turned on at night

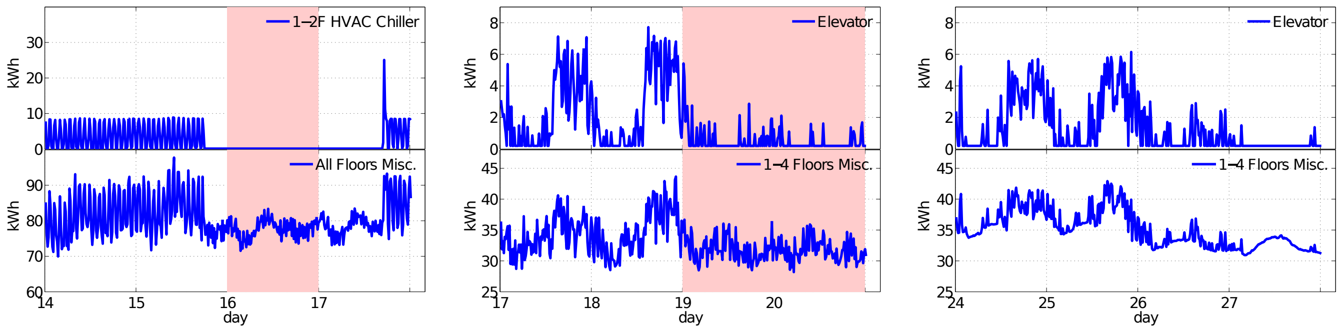
(b) High power usage where the light is left on at night

(c) Low power usage where the HVAC (EHP) is not used during office hours



(d) Long term high power usage partially detected

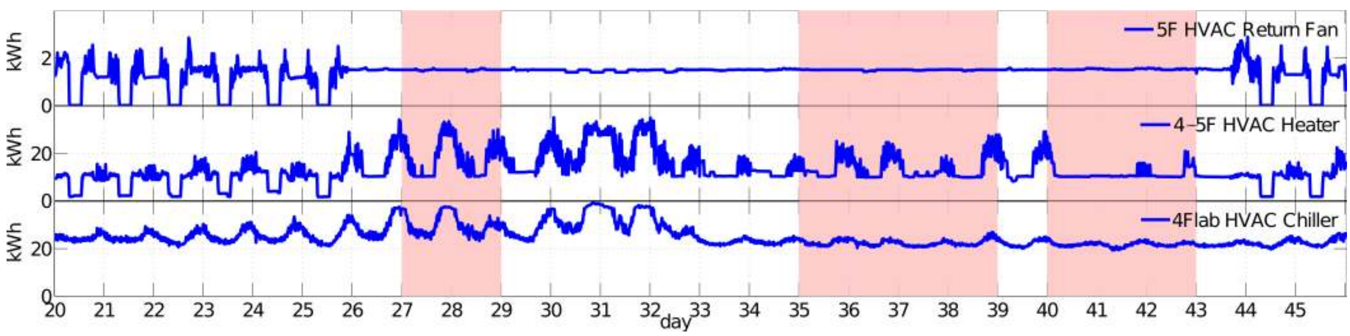
Figure 7: Example of alarms (red rectangles) reported by SBS on the Eng. Bldg 2 dataset



(a) Low power usage due to a chiller failure

(b) High power usage highlighted by the elevator usage

(c) Normal power and elevator usage



(d) Long term high power usage due to competing heating and cooling

Figure 8: Example of alarms (red rectangles) reported by SBS on the Cory Hall dataset

both measurements is significantly altered and an alarm for each device is raised.

SBS also reports 25 alarms corresponding to high power usage. One of the identified anomalies is particularly interesting. We indirectly observe abnormal usage of a device from the power consumption of the elevator and a power panel for equipment from the 1^{st} to the 4^{th} floor. Figure 8b and 8c show the electricity consumption for both devices. SBS uncovers the correlation between the these two signals, as the amount of electricity going through the panel fluctuates along with the elevator power consumption (Figure 8c). In fact, the elevator is a good indicator of the building's occupancy. Anomalous energy-consumption is identified during a weekend as the consumption measured at the panel is independently fluctuating from the elevator usage. These fluctuations are caused by a device that is not directly monitored. Therefore, we could not identify the root cause more precisely. Nevertheless, the alarm is worthwhile for building operators to start investigating.

The most important anomaly identified in Cory Hall is shown in Figure 8d. This anomaly corresponds to the malfunctioning of the HVAC heater serving the 4^{th} and 5^{th} floors. The heater is constantly working for 18 consecutive days, regardless of the underlying occupant activity. Moreover, in order to maintain appropriate temperature this also results in an increase of the 4^{th} floor HVAC chiller power consumption and several fans, such as the one depicted in Figure 8d. This situation is indicative of simultaneous heating and cooling – whereby heating and cooling systems are competing – and it is a well-know problem in building management that leads to significant energy waste. For this example, the electricity waste is estimated around 2500 kWh for the heater. Nevertheless, as the anomaly spans over 18 days, it is hidden in the building's overall consumption, thus, it is difficult to detect by building administrators without SBS.

6. RELATED WORK

The research community has addressed the detection of abnormal energy-consumption in buildings in numerous ways [17, 18].

The rule-based techniques rely on a priori knowledge, they assert the sustainability of a system by identifying a set of undesired behaviors. Using a hierarchical set of rules, Schein et al. propose a method to diagnose HVAC systems [23]. In comparison, state machine models take advantage of historical training data and domain knowledge to learn the states and transitions of a system. The transitions are based on measured stimuli identified through a domain expertise. State machines can model the operation of HVAC systems [22] and permit to predict or detect the abnormal behavior of HVAC's components [3]. However, the deployment of these methods require expert knowledge and are mostly applied to HVAC systems.

In [24], the authors propose a simple unsupervised approach to monitor the average and peak daily consumption of a building and uncover outlier, nevertheless, the misbehaving devices are left unidentified.

Using regression analysis and weather variables the devices energy-consumption is predicted and abnormal usage is highlighted. The authors of [6] use kernel regression to forecast device consumption and devices that behave differently from the predictions are reported as anomalous. Regression models are also used with performances indices to monitor the HVAC's components and identify inefficiencies [28]. The implementation of these approaches in real situations is difficult, since it requires a training dataset and non-trivial parameter tuning.

Similar to our approach, previous studies identify abnormal energy-consumption using frequency analysis and unsupervised anomaly detection methods. The device's consumption is decomposed using Fourier transform and outlier values are detected using clustering techniques [2, 27, 8]. However, these methods assume a constant periodicity in the data and this causes many false positives in alarm reporting. We do not make any assumption about the device usage schedule. We only observe and model device relationships. We take advantage of a recent frequency analysis technique that enables us uncover the inter-device relationships [10]. The identified anomalies correspond to devices that deviate from their normal relationship to other devices.

Reducing a building's energy consumption has also received a lot of attention from the research community. The most promising techniques are based on occupancy model predictions as they ensure that empty rooms are not over conditioned needlessly. Room occupancy is usually monitored through sensor networks [1, 9] or the computer network traffic [19]. These approaches are highly effective for buildings that have rarely-occupied rooms (e.g. conference room) and studies show that such approaches can achieve up to 42% annual energy saving. SBS is fundamentally different from these approaches. SBS identifies the abnormal usage of any devices rather than optimizing the normal usage of specific devices. Nevertheless, the two approaches are complementary and energy-efficient buildings should take advantage of the synergy between them.

7. DISCUSSION

SBS is a practical method for mining device traces, uncovering hidden relationships and abnormal behavior. In this paper, we validate the efficacy of SBS using the sensor metadata (i.e. device types and location), however, these tags are not needed by SBS to uncover devices relationships. Furthermore, SBS requires no prior knowledge about the building and deploying our tool to other buildings requires no human intervention – neither extra sensors nor a training dataset is needed.

SBS is a best effort approach that takes advantage of all the existing building sensors. For example, our experiments revealed that SBS indirectly uncovers building occupancy through device use (e.g. the elevator in the Building 2). The proposed method would benefit from existing sensors that monitor room occupancy as well (e.g. those deployed in [1, 9]). Savings opportunities are also observable with a minimum of 2 monitored devices and building energy consumption can be better understood after using SBS.

SBS constructs a model for normal inter-device behavior by looking at the usage patterns over time, thus, we run the risk that a device that constantly misbehaves is labeled as normal. Nevertheless, building operators are able to quickly identify such perpetual anomalies by validating the clusters of correlated devices uncovered by SBS. The inspection of these clusters is effortless compare to the investigation of the numerous raw traces. Although this kind of scenario is possible it was not observed in our experiments.

In this paper, we analyze only the data at medium frequencies, however, we observe that data at the high frequencies and residual data (Figure 5) also permits us to determine the device type. This information is valuable to automatically retrieve and validate device labels – a major challenge in building metadata management.

This paper aims to establish a methodology to identify abnormalities in device power traces and inter-device usage patterns. In addition, we are planning to apply this method to online detection using, for example, a sliding window to compute an adaptive reference matrix that evolve in time. However, designing such system raises new challenges that are left for future work.

8. CONCLUSIONS

The goal of this article is to assist building administrators in identifying misbehaving devices in large building sensor deployments. We proposed an unsupervised method to systematically detect abnormal energy consumption in buildings: the Strip, Bind, and Search (SBS) method. SBS uncovers inter-device usage patterns by striping dominant trends off the devices energy-consumption trace. Then, it monitors device usage and reports devices that deviate from the norm. Our main contribution is to develop an unsupervised technique to uncover the true inter-device relationships that are hidden by noise and dominant trends inherent to the sensor data. SBS is used on two sets of traces captured from two buildings with fundamentally different infrastructures. The abnormal consumption identified in these two buildings are mainly energy waste. The most important one is an instance of a competing heater and cooler that caused the heater to waste around 2500 kWh.

Acknowledgments

The authors thank Hideya Ochiai for providing the data from the University of Tokyo. This research was partially supported by the JSPS fellowship program and the CNRS/JSPS Joint Research Project. This work is also supported in part by the National Science Foundation under grants CPS-0932209 and CPS-0931843.

9. REFERENCES

[1] Y. Agarwal, B. Balaji, S. Dutta, R. K. Gupta, and T. Weng. Duty-cycling buildings aggressively: The next frontier in hvac control. In *IPSN'11*, pages 246–257, Chicago, IL, USA, 2011.

[2] G. Bellala, M. Marwah, M. Arlitt, G. Lyon, and C. E. Bash. Towards an understanding of campus-scale power consumption. Buildsys'11, page 6, Seattle, WA, Nov. 1, 2011.

[3] G. Bellala, M. Marwah, A. Shah, M. Arlitt, and C. Bash. A finite state machine-based characterization of building entities for monitoring and control. pages 153–160, 2012.

[4] M. Blanco-Velasco, B. Weng, and K. E. Barner. Ecg signal denoising and baseline wander correction based on the empirical mode decomposition. *Computers in biology and medicine*, 38(1):1–13, jan. 2008.

[5] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *J.STAT.MECH.*, 2008.

[6] M. Brown, C. Barrington-Leigh, and Z. Brown. Kernel regression for real-time building energy analysis. *Journal of Building Performance Simulation*, 5(4):263–276, 2012.

[7] P. Chan, M. Mahoney, and M. Arshad. Learning rules and clusters for anomaly detection in network traffic. In *Managing Cyber Threats*, volume 5 of *Massive Computing*, pages 81–99. Springer US, 2005.

[8] C. Chen and D. J. Cook. Energy outlier detection in smart environments. In *Artificial Intelligence and Smarter Living*, volume WS-11-07 of *AAAI Workshops*. AAAI, 2011.

[9] V. L. Erickson, M. Á. Carreira-Perpiñán, and A. Cerpa. Observe: Occupancy-based system for efficient reduction of hvac energy. In *IPSN'11*, pages 258–269, Chicago, IL, USA, 2011.

[10] R. Fontugne, J. Ortiz, D. Culler, and H. Esaki. Empirical mode decomposition for intrinsic-relationship extraction in large sensor deployments. In *IoT-App'12, Workshop on Internet of Things Applications*, Beijing, China, 2012.

[11] T. Hasan and M. Hasan. Suppression of residual noise from speech signals using empirical mode decomposition. *Signal Processing Letters, IEEE*, 16(1):2 –5, jan. 2009.

[12] H. Huang and J. Pan. Speech pitch determination based on hilbert-huang transform. *Signal Processing*, 86(4):792 – 803, 2006.

[13] N. E. Huang. Computing frequency by using generalized zero-crossing applied to intrinsic mode functions. *U.S. Patent 6,990,436 B1*, 2006.

[14] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A*, 454(1971):903–995, 1998.

[15] N. E. Huang, Z. Wu, S. R. Long, K. C. Arnold, X. Chen, and K. Blank. On instantaneous frequency. *Advances in Adaptive Data Analysis*, pages 177–229, 2009.

[16] P. Huber and E. Ronchetti. *Robust Statistics*. Wiley Series in Probability and Statistics. Wiley, 2009.

[17] S. Katipamula and M. Brambley. Review article: Methods for fault detection, diagnostics, and prognostics for building systemsâĂŤa review, part i. *HVAC&R Research*, 11(1):3–25, 2005.

[18] S. Katipamula and M. Brambley. Review article: Methods for fault detection, diagnostics, and prognostics for building systemsâĂŤa review, part ii. *HVAC&R Research*, 11(2):169–187, 2005.

[19] Y. Kim, R. Balani, H. Zhao, and M. B. Srivastava. Granger causality analysis on ip traffic and circuit-level energy monitoring. BuildSys'10, pages 43–48, Zurich, Switzerland, Nov. 2, 2010.

[20] T. Lee and T. B. M. J. Ouarda. Prediction of climate nonstationary oscillation processes with empirical mode decomposition. *Journal of Geophysical Research*, 116, 2011.

[21] J. C. Nunes, S. Guyot, and E. Delechelle. Texture analysis based on local analysis of the bidimensional

empirical mode decomposition. *Machine Vision and Applications*, 16:177–188, 2005.

[22] D. Patnaik, M. Marwah, R. Sharma, and N. Ramakrishnan. Temporal data mining approaches for sustainable chiller management in data centers. *ACM Transactions on Intelligent Systems and Technology*, 2(4), 2011.

[23] J. Schein and S. Bushby. A hierarchical rule-based fault detection and diagnostic method for hvac systems. *HVAC&R Research*, 12(1):111–125, 2006.

[24] J. E. Seem. Using intelligent data analysis to detect abnormal energy consumption in buildings. *Energy and Buildings*, 39(1):52 – 58, 2007.

[25] M. Torres, M. Colominas, G. Schlotthauer, and P. Flandrin. A complete ensemble empirical mode decomposition with adaptive noise. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4144 –4147, May 2011.

[26] U.S. Energy Information Administration. Annual Energy Review 2011, 2012.

[27] M. Wrinch, T. H. EL-Fouly, and S. Wong. Anomaly detection of building systems using energy demand frequency domain anlaysis. In *IEEE Power & Energy Society General Meeting*, San-Diego, CA, USA, 2012.

[28] Q. Zhou, S. Wang, and Z. Ma. A model-based fault detection and diagnosis strategy for hvac systems. *International Journal of Energy Research*, 33(10):903–918, 2009.