Smart-Energy Group Anomaly Based Behavioral Abnormality Detection

Mohammad Arif Ul Alam, Nirmalya Roy, Michelle Petruska, Andrea Zemp
Department of Information Systems
University of Maryland, Baltimore County
alam4@umbc.edu, nroy@umbc.edu, michell5@umbc.edu, zemp1@umbc.edu

Abstract— **Monitoring behavioral abnormality of individuals living independently in their own homes is a key issue for building sustainable healthcare models in smart environments. While most of the efforts have been directed towards building ambient and wearable sensors-assisted activity recognition based behavioral analysis models for remote health monitoring, energy analytics assisted behavioral abnormality prediction have rarely been investigated. In this paper, we propose a data analytic approach that helps detect energy usage anomalies corresponding to the behavioral abnormality of the residents. Our approach relies on detecting everyday appliances usage from smart meter and smart plug data traces in regular activity days and then learning the unique time segment group of each appliance's energy consumption. We focus on detecting behavioral anomalies over a set of energy source data points rather than pinpointing individual odd points. We employ hierarchical probabilistic model-based group anomaly detection [7] to interpret the anomalous behavior and therefore, detect potential tendency towards behavioral abnormality. We apply daily activity logs to evaluate our approach using two real-world energy datasets pertaining to staged functional behaviors, and show that it is possible to detect max. 97% of anomalous days with max. 87% of meaningful micro-behavioral abnormal events generating 1.1% of false alarms. However, we show that our detected abnormality can be meaningfully represented to different stakeholders such as caregivers and family members to understand the nature and severity of abnormal human behavior for sustaining better healthcare.**

I. INTRODUCTION

Human Behavior is considered abnormal when it is atypical, out of the ordinary, and causes some kind of impairment, or consists of undesirable consequences [1]. Advances in wireless sensor networks have enabled the monitoring of daily activity behaviors of individuals. The goal of these monitoring applications is to learn normal behavior in terms of daily activities and look for any deviation, i.e., anomalies, so that alerts can be sent to relatives or caregivers. Human behavior is very complex, and many existing anomaly detection systems are too simplistic which cause many false alarms, resulting in unreliable systems. Existing approaches have typically proposed remotely monitoring the behavior of an occupant in a smart home environment through the use of additional ambient sensors [6], [11], wearable sensors [3], [14] or both [4]. However, the costs and privacy concerns of such sensors have significantly limited their potential for widespread adoption.

Smart meters have grown in popularity among energy providers and home owners interested in reducing overhead and energy costs. They have become a stalwart of energy saving campaigns, but their values have not been exploited outside this domain. Smart meters provide not only the capability to analyze green homes' energy footprint, but also help different stakeholders to better understand the behaviors of inhabitants. This indicates that smart meters can be harnessed to support medical outreach programs. In this work, we propose the use of smart meters and plugs to monitor abnormal behaviors of the inhabitants in order to provide better healthcare support. Successful applications could potentially result in a reduction of wandering, insomnia, forgetfulness, suicides and self-inflicted injuries, as well as their associated costs.

Our proposed application aims to achieve several goals including abnormal behavior predictions, improved emergency response times, support of at risk populations and reduction of the financial burden on hospitals and families. Energy monitoring enables analysts to rely on real-time behavioral patterns to interpret the behavioral and mental health state of an individual at risk, rather than relying only on that individual's verbal feedback. In addition, since the application is deployed at the individual's home, feedback can be obtained constantly, rather than intermittently retrieved during office visits. Furthermore, energy monitoring applications have the potential to support at risk populations reinforcing greater independence while still allowing for a caregiver support system. Employing energy monitoring applications in this manner would be a cost effective alternative to in-hospital observation for vulnerable individuals and a long-term remedy to the national shortage of hospital beds and funds for psychiatric care. Several recent works proposed to detect anomalous activities of daily events based on energy usage [5], [9] or activity recognition [8]. However, no existing works focus on detecting the anomalous events from energy source disaggregation by sub-grouping the energy consumption of a number of appliances rather than identifying individual appliance usage level. We propose a hierarchical group anomaly detection algorithm to efficiently capture the underlying generative phenomena associated with the energy data streams and interpret the anomalous events, if any, interms of behavioral abnormality. Additionally, we propose a novel representation of energy-based usage behavior to which can be easily interpreted by caregivers and family members to understand the any anomalous lifestyle pattern.

Our research in this paper consequently tackles the following key research questions:

- What algorithmic approach can help differentiate

between the normal Activities of Daily Living (ADLs) and abnormal ADLs which are inferred from the energy disaggregation-based appliance usage model?

• What approach is suitable for detecting anomalous data points in-terms of energy consumption of appliances when either those individual appliance's energy consumption is *usual* or *unusual*?

• How can we represent such appliance-specific energy usage-based anomaly events that allows different stakeholders to interpret abnormal activities and risky behaviors of vulnerable residents? How can we apply such underpinning data analytics in solving specific behavioral health problem?

To tackle these challenges, our proposed framework extends upon Mixture of Gaussian Mixture Model (MGMM) [7] to find the most unusual time of a specific day hierarchically and helps group the unusual energy data points as a distribution to validate any abnormal ADLs. We also organize the group anomaly detection events of appliance energy usage with meaningful representation so that caregivers can interpret the possible behavioral abnormality conditions of the residents. More specifically, we make the following **key contributions:**

• We extend the MGMM algorithm in four layers of hierarchical structure for detecting group-based appliance usage anomalies and investigate the anomalous point of each appliance usage in-terms of behavioral abnormality.

• We represent the detected anomalies with human understandable format and explore the feasibility of engaging different stakeholders for assessing and validating the usability of behavioral abnormality representation of smart home residents.

II. RELATED WORKS

The primary purpose of the smart plug is to provide greater protection against loose connections, corrosion and increase energy usage awareness. To the best of our knowledge, no known research focuses directly on the application of smart plugs to perform behavioral health abnormality, although several studies have explored how they can be used in the healthcare domain in future [5], [8], [9]. Studies relevant to the the area of smart environment based behavioral health monitoring include the application of smart applications in elder care [9] or assisted living environments focusing on activities of daily living [8], [13]. 'Holmes' [13] explored DBSCAN clustering based temporal ADLs rules learning in predicting hierarchical anomalous events using ambient motion sensor assisted smart home. The motivation behind many of these studies is to find a way to harness smart applications and technologies in order to provide more independence to the aging population [5]. Previous studies showed promise in tracking patient well being and ADLs in smart home environments based on ambient wireless sensors networks in the presence of single [13] or multiple inhabitants [15]. The application of smart meters and behavioral models have been explored in order to understand activity of elderly individuals in single apartments [10]. [2] focused on optimizing the media experiences within a multi-person household across

resident categories by using an ANT-based service selection framework. A neural network and clustering based technique to analyze residential smart meter data in order to improve customer experience in load forecasting has been proposed in [12]. While our approach is closest to [5], the fundamental difference is that prior study considers an energy disaggregation based appliance usage detection model and application of a periodic log Gaussian Cox process to build an appliance usage schedule, while we propose a hierarchical mixture model capable of handling the usage of a group of appliances to characterize any unusual longitudinal changes over a set of energy data points. While the prior study fails to find the exact anomaly *start-end* points, our framework helps find the anomalous events, even if the individual appliance usage points are normal but their distribution as a group over a set of appliances is unusual or vice versa. Another salient feature of our framework is to represent energy-meter based abnormality representation thus crowds/caregivers can understand the abnormality nature easily. Our framework considers both a group of appliances' cumulative energy consumption pattern and an individual appliance's energy consumption to correlate the anomalies hierarchically so that it can provide the maximum anomaly detection accuracy with reduced false alarm.

III. OVERVIEW

Our framework assumes a smart home that includes *(a)* Enmetric smart meter with multiple power ports, and *(b)* Single-inhabitant environment. As a specific exemplar, this paper considers a smart home environment with *(i)* three Enmetric smart energy meters with 8 power ports where each port is being used by specific appliance (24 appliances), and *(ii)* each appliance has specific properties in terms of behavioral health conditions. For example, a bedroom light which consumes more power at midnight signifies that one inhabitant of the smart home may be suffering from a specific behavioral abnormality called insomnia.

A. Architectural Overview

Fig. 1 explains the basic steps of our overall anomaly detection pipeline.

• **Data Collection and Segmentation:** Raw data are collected using the Powerhouse Dynamic SiteSage and Enmetric energy monitoring systems in two homes. Different power ports provide a multi-modal power reading streams where we apply sliding window mechanism for power spectrum segmentation. We segment the energy data streams into two segments: (i) Day of week, where each of the appliance's aggregate energy consumption has been segmented day-wise; and (ii) Hours of day where each day of week aggregate energy consumption as has been enumerated for each appliance is segmented hourly. We segment the energy usage data streams for both of the behavioral cases, normal and abnormal days.

• **Learning parameters:** Normal days and abnormal days energy data are used for learning our hierarchical MGMM algorithm parameters. We learn two types of parameters: (i)

Day of week where each appliance's energy consumption baseline parameters have been learned considering the day of the week; and (ii) Hours of day where each appliance's energy consumption baseline parameters have been learned considering the hours of the day of the week.

• **Anomaly detection:** We propose a four layer hierarchical group anomaly detection algorithm extending upon the existing MGMM algorithm. Our proposed algorithm first helps find the probabilistic anomaly scores for the day of week that deviate maximum from the day of week baseline. Then it finds the anomalous days of the week scores that go beyond a specific threshold value. Finally, it applies the same anomaly detection method on the detected anomalous days in terms of hour of day baseline. Sometimes the algorithm terminates at the second layer, if it fails to find any anomalous days of week which helps reduce anomaly detection time, providing more fine-grained information about the unusual data points.

• **Behavioral Abnormality Representation for Caregivers:** We define each appliance usage in terms of behavioral abnormality. Our modified hierarchical MGMM framework helps provide anomalous hours of the day. We sort the appliances in terms of anomaly scores and define the anomalies in terms of the energy usage model (excessive usage or reduced usage of a particular appliance). If this defined energy usage representation matches with any of our pre-defined abnormal behaviors, we predict that as a similar type. We validate our predicted behavioral abnormality condition type from the user logs of abnormal days.
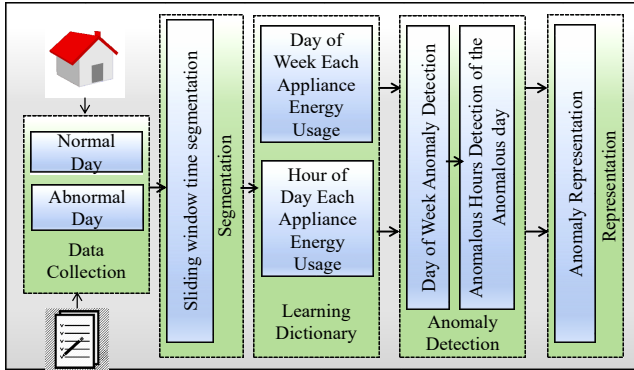

Fig. 1. An overview of our hierarchical MGMM group anomaly detection model

IV. Behavioral Abnormality Group and Appliance Usage

Abnormal behavior is defined as deviating from what is considered normal or most prevalent in a sociocultural context. There are five main criteria of abnormality: statistical criterion, social criterion, personal discomfort, maladaptive behavior and deviation from ideal [1]. We re-define the above criterion in five *Behavioral Abnormality Group (BAG)* categories and their possible indicators in terms of appliance usage as follows:

(1) *Loss of interest in daily activities (LA):* If we are able to build a baseline of daily normal behaviors of appliance's

Group (Segment)	Definition
LA (Day)	Dishwasher-, Disposal-, MBR-, Dryer-, GFI-
IG (Night)	Refrigerator+, Kitchen Outlet+, Fan+, Bed Light+
SI (Day)	All appliances through entire day
CH (Day&Night)	All appliances (+/-) at regular sleep/eat time
EU (Day)	Dryer+, Heater+, Water Heater+, AC+, Fan+

usage of special interest (TV, Cassette Player, treadmill etc.), we can also detect the deviations from the norm that may carry the indicators of loss of interest in daily activities. For example, an individual typically does the following activities in 6.00 AM to 8.00 AM in every working day (Mon–Fri): waking up (bedroom light usage), using bathroom (bathroom light usage), running (treadmill usage), making breakfast (microwave oven usage), coffee (coffee maker usage) etc. Significant decrease/increase of appliance usage (say no coffee, no breakfast and no running) creates more abnormality scores in terms of LA during that time of day.

(2) *Increased eating or insomnia (IG):* Excess usage of the refrigerator, or keeping the lights of any room on for an unusually long time, etc. For example, excessive refrigerator use could be correlated with overeating which sometimes accompanies depression (if occurred in day time) or insomnia (if occurred in nighttime) which increases the IG scores of that time of day.

(3) *Withdrawing from friends and family members i.e., increasing social isolation (SI):* Social isolation i.e., staying inside for entire days can be detected using the energy usage monitoring of everyday appliances which help detect if something is unusual. For example, person stays entire day at home and uses a group of appliances (say TV, refrigerator, computer etc.) which significantly deviate from the normal behavioral of the occupant.

(4) *Changes in sleeping or eating habits (CH):* The lights of the apartment, eating related appliance usage etc., can be postulated as indicators of sleeping and eating habits. Deviation from the normal habits may indicate any abnormal mental conditions.

(5) *Excessive usage of some appliances than normal (EU):* A significant increase in usage of TV, computer, air condition etc., over a period of time compared to their regular usage may indicate potential mental health problems. To build the above profile for behavioral abnormality in terms of energy usage, we categorize the set of appliances in our smart home testbed setup in different groups and enumerate the energy usage with possible behavioral abnormality mapping. A sample behavioral abnormality profile is shown in Table I according to our SiteSage datasets (explained later) where '+' indicates the excessive usage and '-' indicates reduced usage of appliances compared to normal usage levels.

V. Smart Home Energy Meter Setup

In this section, we discuss the process of setting up real energy metering facilities in two real homes.

A. Enmetric Systems Infrastructure

The Enmetric energy monitoring system (as shown in Fig 2(a)) is installed in a single-member household. Like the SiteSage system, the Enmetric system has three components: the first captures the energy data, the second is a wireless device that transfers the data from the sensor to the third component that houses the data on a cloud server. Enmetric's three components are: (i) Enmetric PowerPort, (ii) Enmetric Wireless Bridge, (iii) Energy Management Software. Unlike SiteSage's data collecting sensors that attach directly to a home's electric panel, Enmetric's data collecting component is an intermediary device between an appliance and an electrical socket. Each PowerPort has four channels to draw energy usage data directly from household devices. Data is measured at a frequency of once per plug/sec with an input voltage of 120V 60 Hz. The data captured by the PowerPorts is wirelessly transferred to the bridge which transfers the data on to Enmetric servers in the cloud.

B. Powerhouse Dynamic SiteSage Infrastructure

The Powerhouse Dynamic SiteSage energy monitoring system (as shown in Fig 2(b)) is installed in a multi-inhabitant townhome. The SiteSage system has three components: (i) Energy Monitor Controller and Sensors, (ii) Central Gateway, and (iii) Web-based Software. The Energy Monitor Controller and Sensors are installed on the participant's home electric panel. We attached 24 sensors to a single circuit breaker (electrical panel) in the townhome. The controller and sensors help detect electric usage data from every load. The Central Gateway transfers data captured by the Energy Monitor Sensors to the Powerhouse Dynamic Server which we accessed through a web-based portal to enumerate the appliance-specific energy usage.

Fig. 2.   (a) Enmetric Power Strip system and (b) Powerhouse Dynamic SiteSage System

VI. FORMALIZING ANOMALY DETECTION

We consider a day-wise dictionary learning for each appliance's energy usage. We segment the dictionary into several groups based on the time of the day, day of the week and holiday events. We label each time segment as group and the property of the day as group type. Each group consists of several appliances' energy consumption readings and we refer each appliance as an object. For example, Saturday (group type weekend) 8.00 AM to 10.00 AM is a group that consists of 4 objects: bed room light, mobile charger, bathroom light and air conditioner. Assume that we have M groups denoted by $G_1, \ldots, G_M$ and each group G_m consists of N_m objects, represented by $X_{m,n}, \mathcal{R}^f, n = 1; \ldots; N_m$. These are our observations, e.g., $X_{m,n}$ is the $f = 50$ dimensional energy usage spectrum of the nth appliance in the mth segment group (say Friday 10.00 PM to 12.00 PM),

where these time segment groups are created based on the group type. Assume further that these $X_{m,n}$ feature vectors are generated by a mixture of K Gaussian distributions, that is, each object (appliance) energy usage $X_{m,n}$ belongs to one of these K types distribution, and if we know its type Z_m, $n \in \{1, \ldots, K\}$, then $X_{m,n} \sim \mathcal{N}(\beta^\mu_{Z_{m,n}}, \beta^\sigma_{Z_{m,n}})$. $\beta = \{\beta^\mu_k, \beta^\sigma_k\}^K_{k=1}$ is a dictionary of the possible mean values and covariance matrices for the above mentioned Gaussian mixture of energy usage, where $\beta^\mu_k \in \mathcal{R}^f$ and $\beta^\sigma_k \in \mathcal{R}^{f \times f}$ are positive semi definite matrix. For example, when K = 3, then we can think of these objects (appliances) as 'bedroom light', 'livingroom light', 'kitchen light' devices, and each group G_m is a set of N_m objects where each object can be one of the K different group types. We introduce $S^K = \{s \in \mathcal{R}^K | s_k \geq 0, \sum^K_{k=1} s_k = 1\}$ notation for the K-dimensional probability simplex, and let $\chi_t \in \mathcal{S}^K$ for all $t = 1, \ldots, T$ and $\chi = \{\chi_1, \ldots, \chi_T\}$ denote the set of T possible non-anomalous distributions (proportions) of the K different objects ('bedroom light', 'livingroom light', 'kitchen light') in the M groups. We pose the question whether in group G_m, the distribution of these 'bedroom light', 'livingroom light', 'kitchen light' appliances look normal, that is, if they look similar to a distribution in $\chi = \{ \chi_1, \ldots, \chi_T \}$, or in another words, did we find a group where this distribution seems far from the other emerging distributions that we can observe in the other potential groups.

VII. HIERARCHICAL GROUP ANOMALY DETECTION FRAMEWORK

We employ the Mixture of Gaussian Mixture Model (MGMM) model [7] for group anomaly detection that extends Gaussian Latent Dirichlet Allocation (GLDA) algorithm [17]. We extends upon the MGMM to a four layered Hierarchical MGMM to tackle our energy usage-based group of appliance anomaly detection problem. In this section, we first briefly describe GLDA, MGMM, learning parameters and scoring functions. Finally, we describe our proposed hierarchical approach towards detecting group of appliance usage anomalies.

A. Gaussian Latent Dirichlet Allocation (GLDA)

In the original LDA model the data set is a text corpus, that is a collection of M documents. Each document G_m is a set of N_m words, and each document is represented by a random mixture over latent topics, which is characterized by a distribution (K) over words. Formally, let $Dir(\pi)$ denote the Dirichlet distribution with parameter π, and let $M(\theta)$ be the multinomial distribution with parameters $\theta \in S^K$. In the LDA model given some non-negative hyperparameters $\pi \in \mathcal{R}^K_+$, we generate first some $\theta_m \in \mathcal{R}^K$ from the $Dir(\pi)$ distribution. Having these K dimensional θ_m vectors we generate $Z_{m,n} \sim M(\theta_m)$ variables $(n = 1, \ldots, N_m)$ indicating which topic is active out of K when we generate the word $X_{m,n} \sim P(|Z_{m,n}, \beta)$. Here β is a dictionary of K f-dimensional probability vectors $(\beta_k \in S_f$). We assume $\beta_i = \beta^\mu_i, \beta^\sigma_i$ to be a mean value and a covariance matrix

respectively. Then, the GLDA model observations are given by:

$$X_{m,n} \sim P(.|Z_{m,n},\beta) = \mathcal{N}(\beta^{\mu}_{Z m,n}, \beta^{\sigma}_{Z m,n}) \tag{1}$$

The above GLDA learns that each group is a certain mixture of K Gaussian components, but it also assumes that there is only one "best" mixture (topic distribution) for all groups. In case of appliances, if we consider the topic distribution as day of week, then using GLDA, we are only able to identify one single anomalous day of the weekly dictionary. To solve this problem of multi-modal group anomaly, we further modify this model to MGMM.

B. Mixture of Gaussian Mixture Model (MGMM)

Let again $\chi_t \in S^K, \forall t = 1, \ldots, T$ and $\chi = \chi_1, \ldots, \chi_T$ denote the set of possible non-anomalous probability distributions of the K different topics (*normal, CH, SI, IG* etc. segments) in the M groups. Let $\pi \in S^T$ denote a distribution vector on the set χ, and let $\beta = \{\beta^{\mu}_k, \beta^{\sigma}_k\}^K_{k=1}$ be a dictionary of the possible mean values and covariance matrices. Using the training set $X_{m,n}$, we first learn the hyperparameters π, χ, β of the model. If a group G is not compatible with our model, then it will lead to a small likelihood $P(G|\pi,\chi,\beta)$ compared to that of the other groups, and we can detect it as an anomalous group.

C. Learning Parameters and Scoring Function

We can represent the marginal likelihood of $G = G_m$ [7] as follows:

$$P(G_m|\pi,\chi,\beta) = \sum_{t=1}^{T} \pi_t \prod_{n=1}^{N_m} \sum_{k=1}^{K} \chi_{tk} P(x_{m,n}|z_{m,n},\beta) \tag{2}$$

where the hyperparameters $\Theta = \{\pi,\chi,\beta\}$ can be learned using maximum likelihood estimation with the following optimization:

$$\arg\max_{\Theta} \prod_{m=1}^{M} P(G_m|\pi,\chi,\beta) \tag{3}$$

We define scoring function as $F = -lnP(G|\Theta)$ which can detect group likelihood of the test case towards detecting anomalous group. We can define this scoring function as follows:

$$F = -\int_{\theta_m} P(\theta_m|\Theta,G_m)lnP(\theta_m|\Theta)d\theta \tag{4}$$

We use the group likelihood score to find aggregations of group anomalous points.

D. Modified Hierarchical MGMM

We extend Hierarchical MGMM by decomposing it to four layers: (1) automatic model selection; (2) day of week group anomaly detection; (3) hour of day of week group anomaly detection; and (4) behavioral abnormality probabilities.

(1) Automatic model selection: To automatically determine the values of T and K, we use the modified Bayesian Information Criterion method [19]. The definition of BIC score is given by $BIC(X,\Theta) = lnL(X,\Theta)\frac{1}{2}ln(|X|)|\Theta|$, where $||$ denotes the number of free parameters. We represent above unidimensional BIC scores as a multiple regression problem to extend it to multi-dimensional. We investigate the behavior

of the Schwarz BIC [18] by explaining the phenomenon of the overestimation and modification of BIC that allows the detection of main effects and pairwise interactions in a backcross population [19].

(2) Day of week group anomaly detection: We first consider each day of week as a group type (topics in MGMM algorithm i.e., $M = 7$) that consists with several appliances (objects in MGMM algorithm i.e., if number of appliances is 24 then, $N_m = 24$) and their aggregate energy usage on each day as features ($X_{m,n}$ denotes the mth day and n^{th} number of object energy usage). Then we apply the hierarchical MGMM algorithm to find the most anomalous days of the week. We use automatic model selection for the above MGMM algorithm on the training dataset.

(3) Hour of day of week group anomaly detection: If we find any anomalous days in the test dataset considering day of week group anomaly detection for MGMM, then we port those anomalous days for further analysis. In this step, we consider each hour of the day of the week as a group type (topics in MGMM algorithm i.e., $M = 24$) that consists with several appliances (objects in MGMM algorithm i.e., if number of appliances is 24 then, $N_m = 24$) and their aggregate energy usage on each hour as features ($X_{m,n}$ denotes the mth hour and n^{th} number of object energy usage of each day of the week). We use automatic model selection for the above MGMM algorithm on the training dataset.

(4) Behavioral Abnormality Group (BAG) scores: We define the BAG scores in Table I with two time segments: day: from morning 6.00 AM to 10.00 PM and night: from 10.00 PM to 6.00 AM. To calculate the BAG scores for each anomalous hour, we first calculate individual appliance anomaly scores (A_m where $m = 1, \ldots 24$) using GLDA and calculate BAG probability given A_m, $P(B_p|A_m)$ as follows.

$$P(B_p|A) = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{m=1}^{M} F_{p,q} \times A_m \tag{5}$$

where $P = \{LA, IG, SI, CH, EU\}$ and Q_p is all appliances of pth BAG and $F_{p,q}$ denotes the anomaly score of pth BAG and q^{th} denotes appliances in terms of energy usage. We then take the average of the concerned time segments of all BAG scores i.e., final IG score will be average of each hour IG scores from 10.00 PM to 6.00 AM (Night). We consider a BAG as significant if the score goes above the threshold τ. The hierarchical MGMM learning algorithm for group anomaly detection is illustrated in Algorithm 1.

VIII. Data Collection

For data collection, we set up two single inhabitant houses equipped with two different types of energy metering systems: (1) Enmetric and (2) SiteSage. For abnormal days, participants are provided instructions on how to emulate BAG scores. These instructions include recommendations to simulate insomnia (waking up late, staying up late at night), poor hygiene, a loss of interest in normal activities, excess usage of appliances due to guest visits or special event etc. Participants are asked to keep an activity log consists of

Algorithm 1 Hierarchical MGMM Learning

```
 1: procedure MGMM LEARNING(X, Y, M)              ▷ M = Group
 2:                           ▷ X = Energy Consumption, Y = Group Type
 3:     T, K ← msBIC(X, M)                         ▷ Modified Schwarz
 4:                     ▷ Bayesian Information Criterion (msBIC) algorithm
 5:     for m = 1 to M do
 6:         Y_m ∼ M(π)                  ▷ Choose group type Y_m where
 7:                                              ▷ Y_m ∈ 1, . . . , T
 8:         θ ← χY_m                    ▷ Choose topic distribution
 9:                                              ▷ where θ ∈ S^k
10:         N_m ← Poisson(G_m)              ▷ Group of appliances
11:                          ▷ selection sampled from a Poisson distribution
12:
13:         for n = 1 to N_m do
14:             Z_{m,n} ∼ M(θ_m)            ▷ Choose appliance type
15:                                          ▷ where Z_{m,n} ∈ {1, . . . , T}
16:             X_{m,n} ∼ P(X_{m,n}|β, Z_{m,n}) = N(β^μ_{Z_{m,n}}, β^σ_{Z_{m,n}})
17:                                              ▷ Generate Features
18:         end for
19:     end for
20:     P(G_m|π, χ, β)               ▷ Group Likelihood estimation(Eqn 2)
21:     Θ = {π, χ, β}                      ▷ Parameter estimation(Eqn 3)
22:     Return Θ
23: end procedure
```

TABLE II

BEHAVIORAL ABNORMALITY GROUP (WITH CONCERNED TIME SEGMENTS) AND APPLIANCES MAPPING FOR ENMETRIC DATASET

Group (Segment)	Definition
LA (Day)	Television-, Laptop-, Burr Coffee Grinder-
IG (Night)	Television+, Laptop+, Burr Coffee Grinder+, Microwave+
SI (Day)	All appliances through entire day
CH (Day & Night)	All appliances (+/-) at regular sleep/eat time
EU (Day & Night)	Television+, Laptop+, Burr Coffee Grinder+, Microwave+

approximate time (start and end time) of performed in-home activities. They are asked to review their own activity logs at the end of each day and requested to put BAG scores (LA, IG, SI, CH and EU) along with the frequency (1–3) of past 24 hours. Thus, higher the score of each BAG represents more abnormality of that particular group. For example, in a working day morning, an individual spends 1 hour more time in treadmill than normal (EU frequency = 1, CH frequency = 1), sleeps for 2 hours which he did not use to do (CH frequency = 3) and watches TV for 2 hours which he did not use to do more than 1 hour (SI frequency = 1, CH frequency = 2, EU frequency = 2). The final BAG scores set will be {LA, IG, SI, CH, EU} = {0, 0, 1, 6, 3 }. A trained graduate student is engaged to validate the BAG scores based on participants' activity logs and interacting with the inhabitants to confirm the anomalous nature's clarity.

A. Dataset Description

We describe two of our datasets as follows:

(1) Enmetric Dataset: Consists of one Enmetric energy monitoring system equipped with 8 appliances: iPhone charger, electric toothbrush, television, alarm clock, laptop, microwave, laundry machine and burr coffee grinder. The testbed is a single inhabitant apartment with a bedroom, a living room, a kitchen and a bathroom. We build the BAG distribution as shown in Table II. We collected one month of data with proper activity logs of normal and abnormal behavioral activities. We have total 30 days of continuous energy data with 6 anomalous days (with description) from the activity logs generating a month's data of activity and

TABLE III

BEHAVIORAL ABNORMALITY REPRESENTATION FOR CROWDS (A WEDNESDAY)

Time	Description
12.00 AM to 4.00 AM	Excess usage of refrigerator, water heater, kitchen outlets and bed light
7.00 AM to 12.00 PM	Reduced usage of kitchen outlets, bed light
1.00 PM - 5.00 PM	Excess usage of refrigerator, bed light
Group	**LA, IG, SI, CH or EU**

energy aware abnormal behavior logs.

(2) SiteSage Dataset: It consists of one Site Sage energy monitoring system equipped with 24 appliances listed in Table I. We measure the energy consumption readings of appliances at a rate of watt per second. The townhome includes different appliances which are used to perform different activities for collecting energy footprints. We consider our test bed as three floors with four rooms - Bedroom (Heater, Humidifier, Fan), Basement (Humidifier, Heater, Fan is alternatively used with previous two), Kitchen (Microwave, Refrigerator, Exhaust Fan) and Basement Closets (Washer and Dryer). We collected 11 months (313 days) of continuous energy data with daily activity and behavioral abnormality (41 abnormal days) logs for the inhabitant.

Both of the the datasets are categorized individually into normal and abnormal days based on the activity logs. Normal days reflect the participant's normal or ordinary behavior used to establish a behavioral baseline for learning the hyperparameters ($\Theta = \{\pi, \chi, \beta\}$) of our model.

IX. EVALUATION

We choose 30% of the normal behavioral datasets with uniform distribution of all segment group types and use it to train our model (learning hyperparameters). Then we apply our hierarchical group anomaly detection algorithm on the rest of the datasets including the abnormal days.

A. Performance Metrics

We have taken different approaches to calculate anomaly detection accuracy tradeoffs in our experiments. In each approach, we choose the following equation to measure the anomaly detection accuracy:

$$accuracy = \frac{\#TP + \#TN}{\#TP + \#FP + \#FN + \#TN} \quad (6)$$

where TP, TN, FP, and FN respectively represents True Positives, True Negatives, False Positives, and False Negatives. To measure accuracy of anomaly classification, we use datasets chunk by chunk to maintain its continuity. For anomaly detection, we consider the performance metrics for detecting anomalous days. For behavioral abnormality detection accuracy, we consider each day's performance metrics irrespective of order of behavioral abnormalities and take the average of all anomalous days' individual performance. For correlation analysis, we use Pearson correlation coefficients with significance on $p < 0.05$.

We calculate the entire day's BAG scores and represent the appliance usage in such a way that the caregivers can

understand the behavioral abnormality easily. We also calculate individual appliance anomaly scores for each hour which represent the deviation of energy usage from the baseline. A sample representation of energy usage is shown in Table III.

B. Abnormality Detection Performance

We implement anomaly detection algorithms in Matlab and evaluate the Pearson correlations in SPSS. We evaluate anomaly detection and behavioral abnormality detection performance based on prior mentioned performance metrics considering individual users' activity logs as ground truth. At first we consider the SiteSage datasets. We take the days of week time segment as our group of topics in hierarchical MGMM algorithm to find out the anomaly scores. We obtain 97.76% (40 days) of accuracy in anomalous days detection with 2.2% (6 days) of false positive rate. We further take hours of week time segment as our group of topics in hierarchical MGMM algorithm to find out the anomaly scores of each hour only for the concerned anomalous days. We then represent the anomalous point in terms of BAG for each anomalous hours of the anomaly days and in terms of energy usage. Finally, we take the average to measure the final abnormality assessment scores. We consider BAG score threshold $\tau = 2.8$ for the significance of specific BAG and the false positive rate is reduced to 3 count (1.1%). Finally, we achieve 87.7% behavioral abnormality prediction accuracy for anomalous days with 6.1% false alarm considering the inhabitants provided activity logs. Fig 5 and Fig 3 illustrate detailed representation of our anomaly detection pipeline with an example of four hierarchies framework. We first select the appropriate model for hierarchical MGMM (1st hierarchy). Fig 5(a) shows the scattered plot of aggregate energy usage of each day of a sample week. We can easily visualize that Wednesday (the darkest data points represent group anomalous points) is the anomalous day of the sample week (2nd *hierarchy*). Then we consider Wednesday for hourly group anomaly detection. Fig 5(b) shows the group anomaly of all the appliances for each hour of Wednesday. We can easily visualize that 17th hour is the anomalous point here which also can be multiple hours of a day (3rd hierarchy). We then calculate individual appliance anomaly scores using GLDA, calculate individual appliance anomaly scores for each hour and illustrate the result in Fig 3(a). We note that the total power consumption (WPA) and two water heaters (WH1 and WH2) show more anomalous energy consumption than others. We then represent the BAG scores of the entire day (4th hierarchy) and plot in Fig 3(b). We can visualize that IG (insomnia or overeating) and EU (Excessive Usage of appliances) scores stand over the threshold ($\tau = 2.8$). We find significant correlations among automatic detected and expert validated BAG scores ($r = 0.61*$) and daily behavioral abnormality ($r = 0.51*$) for SiteSage dataset.

Similarly, considering the Enmetric dataset, we achieve 96.67% anomalous days classification accuracy with a 3.3% false alarm rate and 85.3% of behavioral abnormality classification accuracy with 6.9% false alarm. We find significant correlations among automatic detected and expert validated

BAG scores ($r = 0.69*$) and daily behavioral abnormality ($r = 0.65*$) for Enmetric dataset. To evaluate the efficacy of our behavioral abnormality representation, we engage 10 graduate students to predict the behavioral abnormality of the detected anomalous days in the form of Table III. We achieve 83.7% and 81.5% accuracies of behavioral abnormality prediction for SiteSage and Enmetric datasets respectively.
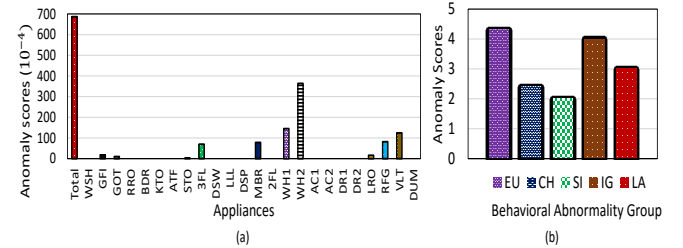

Fig. 3.   (a) Individual Appliance anomaly scores for previously detected anomalous hour (17th hour of Wednesday). Boxes with darkest points represent anomalous data points. (b) Behavioral Abnormality Group (BAG) scores (17th hour of Wednesday).

C. Anomaly Detection Performance Comparison

We apply other alternative methods such as GLDA, GMM and the K-Nearest Neighbor (KNN) algorithms [16] that have been used successfully for collective anomaly detection for long time. Fig 4 illustrates the comparison graphs of anomalous days detection based on group of appliance energy usage monitoring. We note that our approach outperforms GMM, GLDA and KNN in terms of accuracy, average precision and area under the ROC curve (AUC).
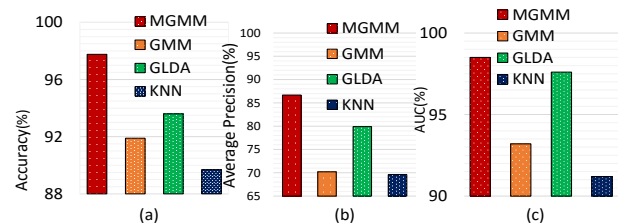

Fig. 4.   Anomalous days detection performance for SiteSage Dataset in terms of (a) Accuracy, (b) Average Precision and (c) Average Area Under the ROC Curve (AUC).

D. Importance of Hierarchical Group Anomaly Detection

To explain the importance of our method we implement [5] proposed method and calculate the hourly anomaly detection for each appliance usage for the month of November using the SiteSage dataset. We observe that each day can be considered as an anomaly in terms of any of the BAG in a huge false alarm ($\sim 76\%$ i.e., 23 days). If we apply hierarchical MGMM, we first find 7 anomalous days considering each day of week as group type, then we find the behavioral abnormality parameters distribution for anomaly scoring and finally validate the results. This reduces the false alarm to 3.3% (1 day).

X. Conclusion and Future Works

Nonintrusive load monitoring (NILM) research domain has been confined into the detection of types of appliances
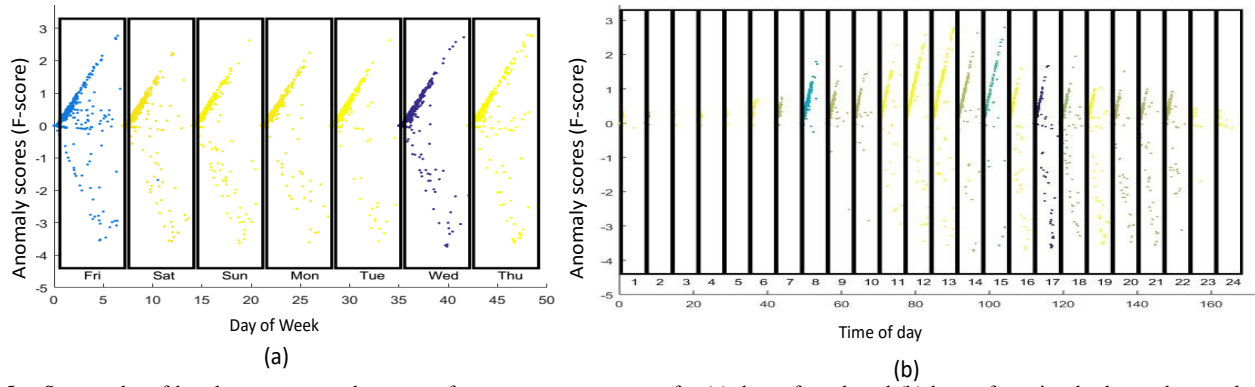
Fig. 5. Scatter plot of hourly group anomaly scores of aggregate energy usage for (a) days of week and (b) hour of previously detected anomalous day (Wednesday). Boxes with darkest points represent anomalous data points.
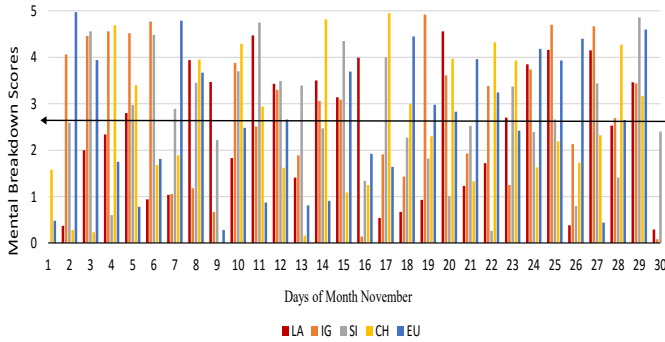

Fig. 6. Anomaly scores of each of the five BAGs for each day of November month (black straight line represents anomaly threshold $\tau = 2.8$).

people have and their behavioral patterns through energy profiling and disaggregation. In this paper, we introduce a new problem as a next step of NILM disaggregation considering, functional and behavioral health intervention programs can benefit from more information and precise behavioral data extracted from NILM. As the analysis provided in this paper enables identifying energy related behavioral patterns, it is possible to determine when prevention is necessary for an individual at risk. Our aim in this paper is to provide basic non-intrusive health monitoring technologies which can be deployed at a large scale without requiring any additional sensors to be installed in any household with multi-inhabitants. Our approach is able to identify appliance usage based activities from smart meter data, extract the pattern of usage and use those patterns to interpret the mental behavioral condition of the occupants. In addition, our meaningful representation of anomalous events and evaluation of the readability test from 10 graduate students confirm that caregivers having no knowledge about the abnormality can understand the behavioral abnormality easily through our framework. Though there remains some real world testing to be done in future, we believe the proposed preliminary approach can help mental health patients to live independently and privately in their own environment while simultaneously accessing smart services which supplement their health care, and as a result, improve their well-being. In future, we will combine individual wearable sensors (say wrist-worn IMUs) based activity abnormality with smart-

energy based behavioral abnormality to investigate insights of mental health in smart home inhabitants.

REFERENCES

[1] D. Sue, D. Wing, S. Sue, Understanding Abnormal Behavior, Eight Edition, NY: Houghton Mifflin Co; 2003.
[2] M. Hossain, S. Hossain, A. Alamri, M. Hossain, Ant-based service selection framework for a smart home monitoring environment, Multimedia Tools And Applications, 2013.
[3] MAU Alam, N. Roy, A. Gangopadhyay, E. Galik, A Smart Segmentation Technique Towards Improved Infrequent Non-Speech Gestural Activity Recognition Model, Pervasive and Mobile Computing (PMC), Special Issue on Gerontechnology, July 2016.
[4] MAU Alam, N. Roy, S. Holmes, A. Gangopadhyay, E. Galik, Automated Functional and Behavioral Health Assessment of Older Adults with Dementia, CHASE 2016.
[5] J. Alcala, O. Parson, A. Rogers, Detecting Anomalies in Activities of Daily Living of Elderly Residents via Energy Disaggregation and Cox Processes, BuildSys, 2015.
[6] MAU Alam, N. Roy, A. Misra, J. Taylor, CACE: Exploiting Behavioral Interactions for Improved Activity Recognition in Multi-Inhabitant Smart Homes, ICDCS 2016.
[7] L. Xiong, B. Poczos, J. G. Schneider, A. J. Connolly, J. VanderPlas, Hierarchical Probabilistic Models for Group Anomaly Detection. AISTATS 2011.
[8] C. Chen, D. J. Cook, Energy Outlier Detection in Smart Environments, Artificial Intelligence and Smarter Living 2011.
[9] S. Ohta, H. Nakamoto, Y. Shinagawa, and T. Tanikawa. A health monitoring system for elderly people living alone. Journal of telemedicine and telecare, 2002.
[10] J. Clement, J. Ploennigs, K. Kabitzsch, Detecting Activities of Daily Living with Smart Meters. Ambient Assisted Living: 6. AAL-Kongress 2013.
[11] MAU Alam, N. Pathak, N. Roy, Mobeacon: An iBeacon-Assisted smart-phone-Based Real Time Activity Recognition Framework, Mobiquitous 2015.
[12] F. Quilumba, L. Wei-Jen,H. Heng,D. Wang, R. Szabados, Using Smart Meter Data to Improve the Accuracy of Intraday Load Forecasting Considering Customer Behavior Similarities. IEEE Transactions On Smart Grid, 6(2), 2015.
[13] E. Hoque, R. F. Dickerson, S. M. Preum, M. A. Hanson, A. T. Barth, J. A. Stankovic, Holmes: A Comprehensive Anomaly Detection System for Daily In-home Activities. DCOSS 2015.
[14] MAU Alam, Nirmalya Roy, GeSmart: A Gestural Activity Recognition Model for Predicting Behavioral Health, SmartComp 2014.
[15] S. K. Das, N. Roy, Context-aware resource management in multi-inhabitant smart homes: A framework based on Nash H-learning, Pervasive And Mobile Computing, 2006.
[16] M. Zhao, Anomaly detection with score functions based on nearest neighbor graphs, NIPS, 2009.
[17] D. M. Blei, A. Y. Ng, M. I. Jordan, Latent Dirichlet Allocation. Journal of Machine Learning Research, 2003.
[18] G. E.Schwarz, Estimating the dimension of a model, Annals of Statistics 6 (2), 1978.
[19] M Bogdan, JK Ghosh, RW Doerge, Modifying the Schwarz Bayesian information criterion to locate multiple interacting quantitative trait loci, Genetics 167 (2), 2004.