Detecting Anomalous User Behavior in Remote Patient Monitoring

Deepti Gupta*, Maanak Gupta†, Smriti Bhatt‡, and Ali Saman Tosun§
*§Dept. of Computer Science, University of Texas at San Antonio, San Antonio, Texas 78249, USA
†Dept. of Computer Science, Tennessee Technological University, Cookeville, Tennessee 38505, USA
‡Dept. of Computing and Cyber Security, Texas A & M University-San Antonio, San Antonio, Texas 78224, USA
*deepti.mrt@gmail.com, †mgupta@tntech.edu, ‡sbhatt@tamusa.edu, §ali.tosun@utsa.edu

***Abstract*—The growth in Remote Patient Monitoring (RPM) services using wearable and non-wearable Internet of Medical Things (IoMT) promises to improve the quality of diagnosis and facilitate timely treatment for a gamut of medical conditions. At the same time, the proliferation of IoMT devices increases the potential for malicious activities that can lead to catastrophic results including theft of personal information, data breach, and compromised medical devices, putting human lives at risk. IoMT devices generate tremendous amount of data that reflect user behavior patterns including both personal and day-to-day social activities along with daily routine health monitoring. In this context, there are possibilities of anomalies generated due to various reasons including unexpected user behavior, faulty sensor, or abnormal values from malicious/compromised devices. To address this problem, there is an imminent need to develop a framework for securing the smart health care infrastructure to identify and mitigate anomalies. In this paper, we present an anomaly detection model for RPM utilizing IoMT and smart home devices. We propose Hidden Markov Model (HMM) based anomaly detection that analyzes normal user behavior in the context of RPM comprising both smart home and smart health devices, and identifies anomalous user behavior. We design a testbed with multiple IoMT devices and home sensors to collect data and use the HMM model to train using network and user behavioral data. Proposed HMM based anomaly detection model achieved over 98% accuracy in identifying the anomalies in the context of RPM.**

***Index Terms*—Anomaly Detection, Internet of Medical Things, Remote Patient Monitoring, Security, Cloud Computing, Hidden Markov Model, Behavioral Data.**

I. Introduction and Motivation

Internet of Medical Things (IoMT), also known as health care Internet of Things (IoT), is a critical data-driven application utilizing smart connected devices, relevant in context of COVID-19 like pandemic. This domain represents a connected infrastructure of medical devices, software applications, health care systems and services. There is an exponential growth in the number of IoMT devices and applications with utilization and demand across a diverse user population. According to the Grand View Research[1], IoMT is predicted to reach $534.3 Billion in market size by 2025. It is expected that the U.S. home-based healthcare market[2] will rise by about 7% annually from $103 billion in 2018 to $173 billion by 2026. With recent technological advancements, IoMT has the potential to revolutionize the future of health care industry. IoMT aims to escalate progress of the health care industry and enable people to receive timely care, enhance their treatment plans, manage medications, and lower health care costs.

In the last few years, data-driven IoMT applications using Machine Learning (ML) and AI technologies are extensively used for critical functions, such as drug discovery, Remote Patient Monitoring (RPM), predictive analytic for hospital resource optimization, interactive medicine delivery, etc., by providing the right information at the right time in the health care system. Today, RPM has became a propitious and required need for users and health practitioners where patients can be monitored remotely, such as home isolation/quarantine. This RPM ecosystem consists of smart medical devices, such as thermometer, oximeter, and wearable devices. These smart health devices collect a large amount of user health data. However, there are several issues associated with smart health devices and applications, such as security of devices and applications, privacy of user data which is shared between devices and applications among many. A recent report by the US National Institute of Standards and Technology (NIST) [1] highlights the need on securing RPM ecosystem and discusses how health care delivery organizations can implement security and privacy controls to build a secure infrastructure for RPM.

In this paper, we focus on security of RPM including user safety and identify anomalous behavior in RPM applications based on the data collected from smart devices while remotely monitoring the patient at home. Research has developed various methods to secure RPM infrastructure from anomalies, focusing on wireless sensor networks [2]–[4], patient's behavior monitoring [5]–[9], signature and correlation analysis [10], [11]. While there are several anomaly detection models developed for RPM, elderly care, and smart homes, our anomaly detection model is designed to identify anomalies in a unique RPM ecosystem which comprises the intersection of two IoT domains (smart health care and smart home). Our model aims to solve the problem of anomalous device data and user behavior which may occur due to malicious or faulty IoT devices or critical conditions of a user, for instance, when a user loses consciousness due to a heart attack and needs immediate medical attention. Proposed anomaly detection can ensure resilient RPM, and save a patient's life.

[1]https://www.healthcareitnews.com/news/asia-pacific/opportunities-pitfalls-healthcare-iot

[2]https://store.businessinsider.com/products/the-us-home-healthcare-report

978-1-6654-3875-9/21/$31.00 ©2021 IEEE

A general RPM environment includes only medical devices, which makes it challenging to identify and differentiate normal and abnormal data with user behavior. For instance, a smart oximeter reports the oxygen level of a diabetic patient as 70%, which is very lower than normal value, thus, it will identify this as an anomaly. However, it could be a false positive due to faulty oximeter. In order to accurately identify anomalies in RPM, both data from medical devices and user activity behavior need to be monitored from home sensors in a smart home. In this work we design a novel RPM ecosystem with both medical and smart home devices that can monitor user health and daily activities within home, and detect anomalies accurately based on the correlation among user behavior and health readings. Extending the above scenario before labeling the oximeter reading as an anomaly, we need to consider these questions: a) Is the user normally moving at home? b) Is the user able to do normal tasks and respond to any alerts or suggestions on smart devices (e.g., smart watch)? and c) Is the user even at home when the oximeter reported 70% oxygen level for the user? Smart home devices capture user's activities, for instance, after reported low oxygen level of the diabetic patient, if motion sensors at home detect no movement of the patient for certain time, then it will identify it accurately as an anomaly and will send an alert to health practitioner for remedial measures.

In this paper, we develop an anomaly detection model using Hidden Markov Model (HMM) [12] approach to identify anomalies in the RPM environment that comprises of smart home and medical devices. We develop the RPM use case using our hybrid approach, and deploy it in a real-world scenario to train and test the model. Using our proposed model, health practitioners can identify anomalies in RPM and trigger appropriate actions as needed in a particular situation. Our model simulation results show that our proposed model detects anomalies in RPM with high accuracy over 98% and generate low false positive rate.

The main contributions of this paper are as follows.

- We develop a novel use case of Remote Patient Monitoring (RPM) based on the intersection of two major IoT domains - smart health care and smart home.
- We design a testbed using sensors and smart health devices to implement the proposed RPM use case.
- We aggregate both network and user behavioral data collected from smart devices and convert it into temporal sequences for our anomaly detection model.
- We design an anomaly detection model using Hidden Markov Model (HMM) and test it based on aggregated training data that represents patient's normal behavior.
- We evaluate our proposed model on testing data along with sets of generated anomalies which achieved over 98% accuracy to identify anomalies.

The remainder of this paper is organized as follows. Section II presents the literature review on anomaly detection models and machine learning approaches such as HMM. Section III discusses the RPM use case and illustrate steps to develop the RPM model architecture. Section IV discusses network and behavioral data collection. Section V discusses anomalous behavior detection model using HMM approach, and deployment for RPM use case. Section VI presents the implementation details and results, followed by conclusion in Section VII.

II. Related Work

We present related work in anomaly detection models based on network and behavioral data.

A. Anomaly Detection Models for RPM

Extensive research has been done to keep RPM systems secure. Yamauchi et al. [13] introduced a method to detect attacks due to home appliances based on user behavior along with conditions at smart home and also presented a comparative study with HMM. Zhang et al. [14] proposed a medical security monitor (MedMon) that snoops on all the radio-frequency wireless communications to/from medical implantable devices to identify anomalies. Tonchev et al. [15] presented a non-intrusive sleep analyzer for real time detection of sleep anomalies based on HMM and Viterbi algorithm. The authors used non-intrusive sensors such as bed pressure sensor, microphone, and accelerometer. Pham et al. [16] used spectral coherence analysis for accelerometer data, which is collected from wearable devices to develop an anomaly detection model. Deep et al. [7] presented a survey on anomalous behavior detection for elderly case using dense-sensing networks and concluded that sensor fusion techniques could increase the efficiency of dense sensing network. Another study [17] showed that anomaly detection model for Ambient Assisted Living (AAL) focused on user's activities, such as sit down for dinner, open stoves etc., not related to user's body parameters. A study [18] presented an approach to reduce the false negative rate (FNR), i.e. disease should not get undetected. In order to protect the devices, security models are discussed in several research [19]–[23].

B. Hidden Markov Models

HMM models have been extensively used in various smart connected domains. HMM model-based behavior analysis system for assisted living is discussed in [24], where sensors are located in a smart home, each room with at least one motion detector, one temperature sensor, one light level detector and two lighting sensors status. In [25], a general framework is proposed for securing medical devices based on wireless channel monitoring and anomaly detection model using Hierarchical Hidden Markov Model (HHMM). Kotevska et al. [26] proposed coordinated intelligence approach to classify normal and abnormal behavior of resident using Markov chain at home where geographical co-located multiple sensors communicate with each other and take decisions based on the collective co-location in some capacity. Li et al. [27] constructed HMM-based anomaly detectors and presented a comparative analysis based on several transformation methods. Research [28] introduced an anomaly detection model based

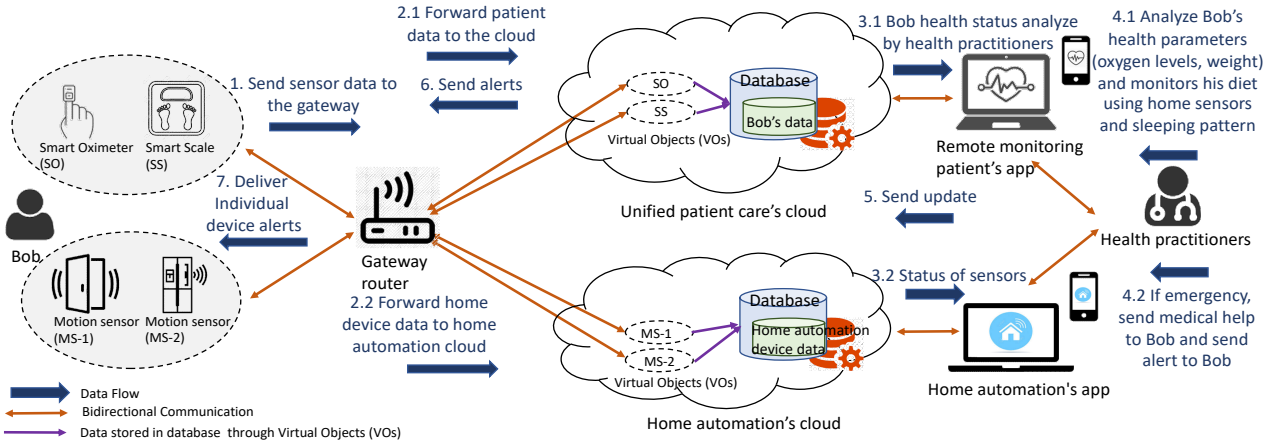


Fig. 1: A Use Case of Remote Patient Monitoring (RPM)

on user's behavior using HMM in a smart home environment. Narayanan et al. [29] proposed OBD SecureAlert model for vehicles, which detects malicious behaviors and sends alerts while a vehicle is in unsafe state. Ren et al. [30] introduced anomaly detection approach based on a dynamic Markov model where sequence data is managed by sliding window. Zhu [31] used HMM approach for detecting anomalies in a person's blood glucose levels using historical observations as a benchmark.

Extensive research has been done for anomaly detection in RPM and use of HMM in other smart domains for anomaly detection. However, as discussed earlier, an hybrid approach which uses both smart home and IoMT is still missing to detect anomalies more accurately (as suggested by our results). In addition, to our understanding and literature review HMM model has not been used in RPM. To bridge this gap, we believe our research offers a novel perspective to detect anomalies in RPM using HMM.

III. Remote Patient Monitoring (RPM)

In recent years, RPM has received significant attention due to its capability of providing health care services to patients while having the convenience of staying at home. Smart home-based health care and patient monitoring can enhance the quality of service and reduce the cost of health care.

A. RPM Use case

In this section, we present a RPM use case that is aligned with a real-world scenario. We consider a user *Bob* who is 34 years old and lives alone at his home, and has been diagnosed with *Obstructive Sleep Apnea (OSA)* disease. He is being monitored by his health practitioners continuously utilizing the RPM ecosystem enabled by smart devices deployed at his home. This RPM setup consists of set of devices including Ethernet tag manager, two wireless sensor tags, and various Bluetooth Low Energy (BLE) based smart health devices such as iHealth smart oximeter, iHealth smart scale. These specific smart health devices along with smart home devices are suggested by his health practitioners. Fridge door sensor and smart scale are used to monitor *Bob*'s diet, eating behavior, and weight. Another motion sensor is attached on bedroom door to track *Bob*'s activity, specially in night as *OSA* patient has disturbed sleep and may wake up multiple times in night. Smart oximeter keeps track of his oxygen level at regular intervals. In case *Bob* is not able to sleep due to shortage of breathe, a notification is sent to the health practitioners based on data from bedroom door sensor and oxygen level through home automation cloud service and unified patient care cloud respectively. Lets us assume that the normal range of *Bob*'s oxygen level is SpO_2 (88% - 99%). If *Bob*'s oxygen level reaches less than 80%, then the health practitioner is notified who can also send connected ambulance to *Bob*'s home. Figure 1 shows a sequential view of the RPM use case scenario where *Bob*'s health condition is continuously monitored by his health practitioners. *Bob* checks his oxygen level and weight regularly in specific time intervals. These BLE based health devices communicate with a gateway device (*Bob*'s smart phone) at object abstraction layer of Enhanced ACO architecture (EACO) [32] to allow interaction with the cloud service and applications. The generated data from these devices is transmitted through the gateway to their corresponding manufacturer's cloud platforms (e.g., iHealth cloud and Wireless sensor cloud respectively) and is stored in remote cloud. Users rely on the security mechanisms deployed by smart device manufacturers to ensure collected data privacy.

In general, anomalous behavior is defined as any violation or deviation from the normal behavior. We defined the anomalies that could occur in the RPM use case if a user deviates from his/her behavior or reflects any abnormal behavior. The following use case presents examples of abnormal behavior and possible threat scenarios.

- In RPM, the devices are used in a particular order based on user's daily routine. Any deviation could be an anomaly; for example, after checking oxygen level, a user opens the bedroom door within time window, if not

that means user needs emergency assistance.

- When user is not at home, but smart scale or oximeter are still sending data (vital readings). This could be due to a faulty smart device or possible intrusion at home.
- An attacker can compromise an IoT device(s), for example, a user is measuring his oxygen level while other sensors like smart scale and fridge door are also activated at same time.
- Bedroom door sensor is activated multiple times in a short interval of time during night and the oxygen level of user is low, i.e. user is not able to sleep during night. The condition can aggravate to a critical level and a health practitioner may need to intervene and check the patients health by sending messages to confirm his well-being.

B. System Architecture

Our RPM model architecture is categorized into 3 phases:

- **Data Collection Phase:** This is the first step in which the stream of data is collected from sensors and smart health devices to develop the anomaly detection model.
- **Model Generation Phase**: We analyze the collected data and convert into sequences to feed the HMM model.
- **Anomaly Detection Phase:** In this final phase, anomalies are detected automatically using our proposed HMM-based anomaly detection model.

In these phases, we discuss and define normal behavior of a user and devices associated with that user as patterns in sequences of observations. For tracking the user behavior and training our anomaly detection model, *opening the bedroom door, opening the fridge door, measuring weight and oxygen level* are the activities of our interest. Each activity and its relevant data is measured using corresponding smart devices in RPM ecosystem. These phases of RPM model architecture are described in the following sections.

IV. Data Collection

In RPM ecosystem, the smart connected devices generate tremendous amount of data while monitoring the patient at home. To collect this diverse data set, we set up the RPM testbed at a home by deploying IoT sensors and smart health devices (architecture shown in Figure 2). We consider that one patient lives in the smart home and performs his daily tasks, and the testbed captures data for three weeks. Data is collected from different sensors and smart health devices, which are used to build anomaly detection model.

We collect data in two different categories - (i) Network data, and (ii) Behavior data. We integrate both types of data to learn typical behaviors of a patient who uses smart medical devices and his activities are captured from home sensors. Network data refers to the TCP/UDP packets from IoT devices and smart phone. Behavioral data is collected based on the sensors' readings showing door status (open/closed) and health devices readings (e.g., 96% SpO_2). This behavioral data is fetched using API from different device manufacturer cloud data repositories (e.g., iHealth cloud, mytaglist cloud).

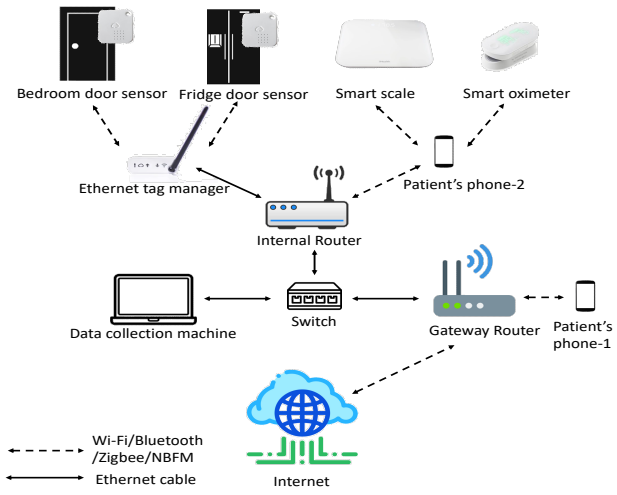


Fig. 2: Testbed Deployed for Remote Patient Monitoring

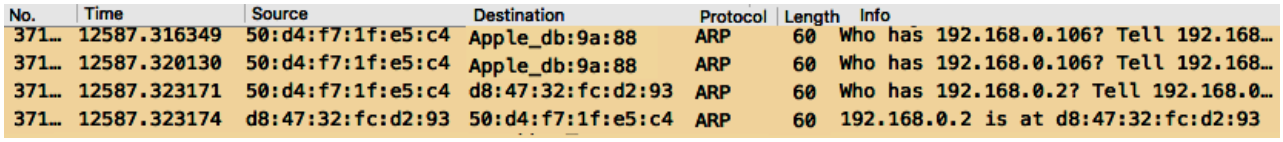
A. Network Data Collection

The varieties of data packet including audio and video flow through the gateway router of a smart home. However, we only require RPM related data from motion sensors and smart health devices. This required data cannot be separated from other non essential data inside the network, since the internal devices IPs are masked during Network Address Translation (NAT) in gateway router. Our setup consists of ARRIS SB6121 modem, two TP-Link AC1200 wireless router, NETGEAR ProSAFE Plus GS105Ev2 switch, and a data collection machine. The data collection machine is connected to the destination port of the switch and all data coming from the router will be available in it, as shown in Figure 2. The packets are collected at the switch using Wireshark[3] packet sniffer. We use the switch as a bridge between the gateway and internal router, and enable port mirroring in the switch. Port mirroring is a method of monitoring network traffic by setting up one or more source ports to send a copy of every packet received to a designated destination port. It is important to assign static IP to all devices in the testbed since IP address filtering is applicable if devices have a static IP.

As shown in Figure 2, BLE based health devices are connected to patient's phone-2 to communicate its cloud service. To avoid other traffic, phone-2 is used only for data transfer of health devices and has set a static IP during three weeks of data collection. The motion sensors communicate to their cloud through Ethernet tag manager, and these wireless motion sensors use "narrowband FM" technology[4] to communicate to Ethernet tag manager. The static IP address of this tag manager is fixed via MAC address reservation. *Bob*'s phone-1 is used to keep track his presence at his home, where phone-1 connects to the gateway router. For instance, whenever user enters his home, phone-1 is connected Wi-Fi automatically. We filter the

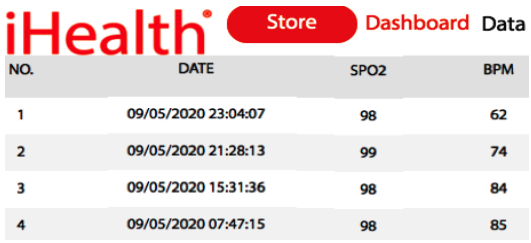
[3]https://www.wireshark.org/

[4]Long-range RF communication: Why narrowband is the de facto standard



No.	Time	Source	Destination	Protocol	Length	Info
371…	12587.316349	50:d4:f7:1f:e5:c4	Apple_db:9a:88	ARP	60	Who has 192.168.0.106? Tell 192.168…
371…	12587.320130	50:d4:f7:1f:e5:c4	Apple_db:9a:88	ARP	60	Who has 192.168.0.106? Tell 192.168…
371…	12587.323171	50:d4:f7:1f:e5:c4	d8:47:32:fc:d2:93	ARP	60	Who has 192.168.0.2? Tell 192.168.0…
371…	12587.323174	d8:47:32:fc:d2:93	50:d4:f7:1f:e5:c4	ARP	60	192.168.0.2 is at d8:47:32:fc:d2:93

Fig. 3: Snapshot of Network Data Captured using Wireshark



NO.	DATE	SPO2	BPM
1	09/05/2020 23:04:07	98	62
2	09/05/2020 21:28:13	99	74
3	09/05/2020 15:31:36	98	84
4	09/05/2020 07:47:15	98	85

Fig. 4: Snapshot of iHealth Cloud Data

required network traffic of phone-1 to know *Bob*'s presence at home, which is shown in Figure 3. This figure presents that phone-1 is connected to the gateway router, where mac address of gateway router is 50:d4:f7:1f:e5:c4 as source address and mac address of phone-1 is 20:ab:37:db:9a:88 as destination address. In this research, we define the status of phone-1 as 'in' and 'out' based on *Bob*'s presence.

B. Behavioral Data Collection

The behavioral data represents the status of bedroom door, fridge door, and vital sign readings of smart health devices. In home, patient's movement is captured through sensors, for instance, if a motion sensor is attached to fridge door which identifies if a fridge door is open or close, i.e. patient is hungry or not. Another motion sensor is attached on bedroom door which identifies about sleeping time of patient. These motion sensors are attached on the doors with 30° angle to identify door is open or not. The home sensors are connected to Ethernet based smart tag manager, and this tag manager pushes the data into home automation's cloud. To differentiate between both sensors, the status of bedroom door are defined as 'bd_open' and 'bd_close', and the status of fridge door are defined as 'fd_open' and 'fd_close'. The status of the sensors and vital sign readings of health devices are retrieved from their corresponding clouds using API calls. Figure 4 shows body parameter readings for *Bob* and from this data we derived the observation status for oximeter and smart scale. The smart health devices generate numeric reading, however, the devices' numeric readings cannot be incorporated into logical calculations. We use normal distribution to discretize continuous data by using different thresholds, and calculate its mean μ and standard deviation σ to define the range. For oximeter, three ranges (ox_1, ox_2, and ox_3) are defined, where 'ox_1' comes under range of positive second deviation $\mu+2\sigma$, which is 97% to 99.3%, and 'ox_2' comes under range of negative second deviation $\mu-2\sigma$, which is 96.9% to 94.7%, if readings fall outside the range $[\mu-2\sigma, \mu+2\sigma]$, is considered

TABLE I: List of Observations For Devices

Sensors/Devices	Observations
Bedroom door	bd_open, bd_close
Fridge door	fd_open, fd_close
Weight scale	sc_off, sc_1, sc_2, sc_3
Oximeter	ox_off, ox_1, ox_2, ox_3
Phone-2	on, off
Phone-1	in, out

TABLE II: Sample log collected from the sensors in RPM

Time	Sensors	Status	Phone-2
8:02	Phone-1	in	
8:04	Oximeter	ox_1	on
8:06	Bedroom door	bd_open	
8:07	Bedroom door	bd_close	
8:24	Scale	sc_2	on
8:32	Fridge door	fd_open	
8:33	Fridge door	fd_close	

as 'ox_3'. The fourth status of oximeter is 'ox_off'. We also define the range of smart scale in similar way. Patient's phone-2 has two status 'on' and 'off', and is connected to internal router and is only used to connect Bluetooth devices to send the data (vital readings) to iHealth cloud. Table I lists status of sensors and devices, which are considered as observations in HMM described in section V).

In our experiment, each sensor (door, fridge) sends its status every 30 seconds, while the status of IoMT devices (oximeter, scale) is measured by user at regular intervals. The user's presence is captured through his phone-1 through network traffic. To analyze, both network and behavioral data are integrated based on time-stamp, as shown in Table II.

V. Anomalous Behavior Detection Model

In this section we propose an anomaly detection model for RPM ecosystem. We define the anomalies as abnormal activities performed by the user and abnormal values triggered from IoMT devices. This proposed model is based on HMM and considers the user behavior as specific sequences of data observations to feed the HMM model. We first present the HMM model, which learns the user's normal behavior and later define the proposed anomaly detection model for RPM use case scenario (discussed in Section III) .

A. HMM for Anomaly Detection in RPM

We first collect diverse data from various IoMT devices and home sensors. Later, we analyze the collected data and convert into temporal sequences to feed HMM model. The

HMM is an augmentation of Markov chain model, which is able to detect the sequential relations among hidden states. This is a probabilistic model where a sequence of observations is generated by visible observations of internal hidden states. These hidden states are not observed directly or are not visible. The transitions between hidden states are assumed to have the form of a (first-order) Markov chain.

The specific parameters used to define the HMM are: N, M, A, B, and π, which are also mapped to the RPM use case. N refers to the total number of all possible states in this model. At *Bob*'s home, we consider his movement as a sequence of states, where sensors are activated at each state. Here, we denote an individual state from the set of states $S = S_1, S_2, S_3, ...S_N$, at time t as q_t. M is the number of unique visible observations for different sensors in a state, $V = v_1, v_2, v_3, ...v_M$ is the set of all possible observations. In our experiment, observations are simply status of different sensors, as shown in Table I. Parameter A represents the state transition matrix. It gives the probability of being able to move from one state i to another state j, as described in equation 1. If a user does not move to any other state i.e. the transition probability would be zero. Parameter B represents a probability distribution of seeing one of the observable status, given that user is in a particular state and is defined in equation 2. Finally, π is the probability of starting at a particular state, which is randomly chosen.

$$
\begin{aligned}
&A = [a_{i,j}] \\
&a_{i,j} = P(q_{t+1} = S_j | q_t = S_i) \\
&where\ q_t \text{ is the state at time t,} \\
&S_i, S_j \in S \text{(set of all possible states)} \\
&1 \leq i \\
&j \leq N, t = 1, 2, 3, ... \qquad (1)
\end{aligned}
$$

$$
\begin{aligned}
&B = [b_{i,j}] \\
&b_{i,j} = P(v_k \text{ at t} | q_t = S_i) \\
&where\ q_t \text{ is the state at time t,} \\
&v_k \in V \text{(set of all observations)} \\
&S_i, S_j \in S \text{(set of all possible states)} \\
&1 \leq i, j \leq N, t = 1, 2, 3, ... \qquad (2)
\end{aligned}
$$

B. Learning End User Behavior

In RPM environment, *Bob* uses medical IoT devices and his activities are captured through motion sensors. Figure 5 shows the timeline of *Bob*'s movement at his home, as he moves from one state to another state, where multiple sensors and devices can be activated/deactivated at same time. At T_1, oximeter is activated (observation 'ox_2') along with phone-2 ('on'), and at T_2, the status of oximeter is changed from 'ox_2' to 'ox_off', and the status of phone-2 from 'on' to 'off'. Bedroom door is opened at T_3, and close at T_4. Then, bedroom door is open (bd_open) again at T_5, followed by multiple sensors (oximeter, bedroom door, scale, phone-2) activating at T_6. At T_6, oxygen level is very low, bedroom door is also open, and smart scale is also activated. According to our model, the three possible sequences of observations will be generated from T_5 to T_6, e.g., [bd_open, ox_3], [bd_open, sc_2], and [bd_open, off]. At T_7, user is at home, phone-2 is still activated, while other sensors, devices are deactivated. At hidden state T_8, status of phone is 'off', user is out.

There are three issues in HMM which must be resolved using the following approaches. In one approach, given the sequence of observations O = O_1 O_2..... O_T and a model λ = (A, B, π) is used to evaluate $P(O|\lambda)$ probability of any

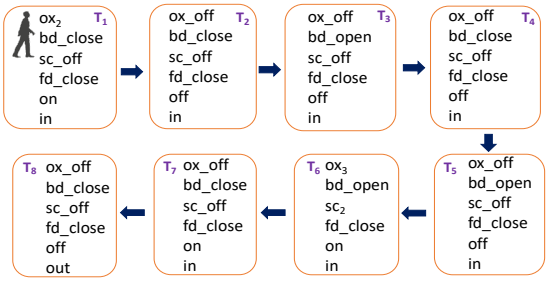


Fig. 5: Sample of Remote Patient Monitoring Events Timeline

sequence of observations using Forward-Backward algorithm [12]. Next approach is used to adjust the model parameters λ = (A, B, π) to maximize $P(O|\lambda)$ using Baum-Welch algorithm [12]. In our experiment, we collect time-series data from all the devices and sensors, and convert them into temporal sequences, which will be the training sequences of the model. The training sequences O = O_1, O_2,... ,O_K are extracted from this time-series data, where O_K is an observation at each state. In our experiment, we assume that O_K is vector as multiple sensors are activated at same time (shown in Figure 5), where $O_K = v_1^{(k)}\ v_2^{(k)}.....v_t^{(k)}$, $1 \leq k \leq K$. Observation vector (O_K) implies to generate different combinations of sequences. For example, sequence of observations extracted during state transitions from time T_1 to T_5 using Figure 5 is [{ox_2,on},{ox_off,off}, bd_open, bd_close, bd_open]. In our experiment, we extracted 420 sequences of observations from temporal data, and trained the model with these sequences using Baum-Welch algorithm to determine the HMM model parameters λ = (A, B, π). The training process repeats until the resulting probabilities converge satisfactorily. For testing, Forward-Backward algorithm is used to evaluate the new sequence of observations by determining the likelihood P(O|λ), which is the log probability of new sequence of observations. The high value of log probabilities indicates that new sequence of observations fit for the model, and low value of log probabilities implies deviation from normal behavior, and identify as anomaly.

C. Anomaly Detection Model

After generating transition, emission and initial probabilities matrices using Baum-Welch algorithm, we use this generated HMM model to detect anomalies in RPM environment. In this research, we are detecting both attack states and unsafe/anomalous states. Here is example of faulty sensors, when user is away from his home, and smart scale is still sending readings to patient's care cloud. To detect the anomalies, we generated a sliding window of hidden states as shown in Figure 6, where observation vector O_K is visible at each state. To detect anomalies in real-time, a sliding window of "m" observations, O_{window} = O_1,O_2,....O_m is introduced. The sliding window moves every time a new observation vector is available at a

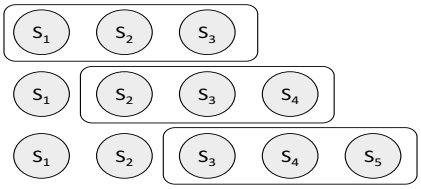


Fig. 6: Sliding window to detect anomalies using Forward-Backward Algorithm

new state and calculate the log probability of new generated sequences of observations using Forward-Backward algorithm. As we define that each observation would be a vector of different sensor values, thus, it will generate a set of log probabilities corresponding to each sequence of observations. If any log probability in this set is below a threshold (discussed in Section VI), it identifies it as an anomalous state. Any sequence of observations is considered anomalous, if it is not accepted by the proposed model with low probability.

VI. Evaluation and Results

In this section, we present the evaluation of our proposed model and also discuss the results. We generated 420 sequences of observations from collected data during the period of three weeks. This data represents *Bob*'s normal behavior in RPM environment. We divide collected data into training and testing datasets, and calculate the threshold value based on the lowest value of log probability of training sequences in various settings. The threshold value is used to identify the anomalies in RPM ecosystem. Further, we generate sets of anomalous scenarios to validate the proposed model. The details of the experiments are presented in the following subsections.

A. Analysis of Normal Dataset

For training, we use Baum-Welch algorithm and use Forward-Backward algorithm for testing. It is required to generate a threshold value to test any sequence using Forward-Backward algorithm. To validate the threshold value, we divide the experimental data into three different settings and perform these experiments. For our first experiment, 70% of data is used for training and 30% is used for testing. We train the model using 70% sequences of observations and calculate the log probabilities of training sequences, which are in the range of -5 to -14. For second experiment, we train the HMM model with the first 60% of data and observed that the log probability of training sequences ranges from -6 to -14. In third experiment, we divide 420 sequences into two parts, each set has 210 sequences, where 70% sequences are used for both training data and provide log probability ranges (-5 to -14, -6 to -13). Hence, we set the threshold value to be -14, i.e. the log probability value of any sequence lower than -14 is anomaly. The testing sequences of above defined three experiments are tested on our proposed model using their corresponding training datasets, and this model provide

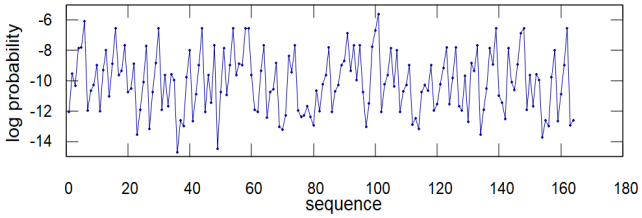


Fig. 7: Test data for Remote Patient Monitoring

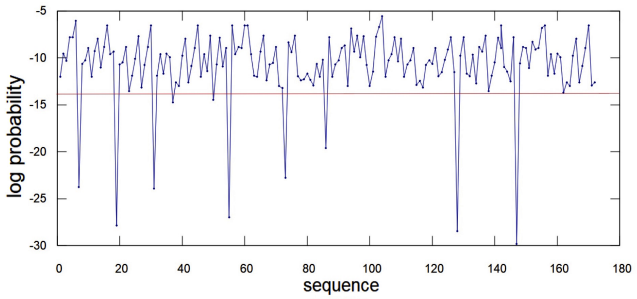


Fig. 8: Generated anomalies injected into test normal data

different sets of log probability ranges (-5 to -14.9, -7 to -14, -6.5 to -13.4, -6 to -14). The range of log probabilities of testing data of first experiment is shown in Figure 7, which also shows two false positives.

B. Abnormal Scenario Detection

To evaluate the model's performance, we manually generate a set of eight anomalous scenarios which are representative of real-world scenarios. We consider abnormal behaviors including unusual device reading, missing dependency order between devices, and compromised sensors/devices to select eight anomalous scenarios. First anomalous scenario is *User is out from his home, smart scale is still sending reading to the unified care's cloud*. It may happen due to faulty/compromised device, which can effect *Bob*'s diet. To simulate this scenario, we take a normal sequence of observation, activate the smart scale and modify the status of phone-1 from 'in' to 'out'. Another crafted scenario is *SpO_2 is very low (i.e. ox_3) and bedroom door is not activated.*, i.e. oxygen level of user is low, and he is not able to get up to open the bedroom door within a time frame. Third anomalous scenario is *when user is out, scale, oximeter and fridge door sensor activate simultaneously*, that indicates the presence of another user at home or any malicious activity can be possible. Fourth and fifth anomalous scenarios are *when user is at home, fridge door, oximeter and scale activate simultaneously*, and *oximeter reading is reported as ox_2, but ox_off is not followed by ox_2 within a time frame* respectively. Sixth anomalous scenario is *when user is at home, sc_3 is followed by ox_3*, i.e. only medical devices are activated with unusual vital readings and seventh scenario is *when user is out of his home, bd_open and fd_open are activated*. The last anomalous scenario is *bd_open followed by bd_open along with ox_3*, i.e. due to low level of oxygen, user is not able to

TABLE III: Confusion Matrix

N=172	Actual:Yes	Actual:Yes	Total
Predicted:Yes	$TP = 45$	$FP = 2$	47
Predicted:No	$FN = 1$	$TN = 124$	125
Total	46	126	

sleep. These eight generated anomalous scenarios are injected into normal testing data of first experiment and tested on proposed anomaly detection model. The threshold values of log probabilities of these eight anomalous scenarios are lower than -14, which is shown in Figure 8 along with two false positives. Hence, the proposed model detects all generated anomalies in RPM environment.

In the next experiment, we generate another set of anomalous scenarios by fixing or randomly choosing status of specific sensors/devices. For example, we change the 'ox_2' to 'ox_3', 'on' to 'off', and 'in' to 'out'. In total, we generated 38 anomalous scenarios randomly, and added eight generated anomalous scenarios to create a dataset of 46 anomalous scenarios. This dataset is tested on our trained model, which detects 45 anomalies out of 46. Moreover, this model can work with any size of dataset. The confusion matrix of the first experiment (30% of 420) along with the 46 anomalous scenario dataset is given in Table III, which shows that our proposed HMM based approach provides 98% accuracy.

VII. Conclusion

In this paper, we propose an anomaly detection model for the RPM ecosystem consisting of smart home and medical IoT devices. We design a testbed of RPM, collect time-series data, and convert data into sequences of observations which are used to train the HMM model using Baum-Welch algorithm. This proposed model detects anomalies based on a single device, and also identifies anomalies based on the combination of data collected from both IoMT and home-based devices. Our model evaluation shows that the proposed HMM based approach can detect anomalies in RPM real-world events/scenarios and randomly generated anomalies with 98% accuracy. This model can be adopted for a diverse range of anomalous use cases. For future, we plan to develop a robust anomaly detection model for multi-users behavior with geographically co-located sensors in home-based health care applications with more anomalous events.

Acknowledgement

This research is partially supported by the NSF Grant 2025682 at Tennessee Technological University, and the Chancellor Research Initiative (CRI) grant at Texas A&M University-San Antonio. The authors would like to thank Dr. Sudip Mittal for his suggestions.

References

[1] J. Cawthra *et al.*, "Securing telehealth remote patient monitoring ecosystem," *NIST SPECIAL PUBLICATION*, p. 30B, 1800.

[2] O. Salem *et al.*, "Anomaly detection scheme for medical wireless sensor networks," in *Handbook of medical and healthcare technologies*. Springer, 2013, pp. 207–222.

[3] V. Vippalapalli *et al.*, "Internet of things (IoT) based smart health care system," in *2016 IEEE International Conference on Signal Processing, Communication, Power and Embedded System*, pp. 1229–1233.

[4] X. Luo *et al.*, "Design and implementation of a distributed fall detection system based on wireless sensor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, p. 118.

[5] L. Rachakonda *et al.*, "iLog: an intelligent device for automatic food intake monitoring and stress detection in the IoMT," *IEEE Transactions on Consumer Electronics*, vol. 66, no. 2, pp. 115–124, 2020.

[6] M. Bi *et al.*, "Anomaly detection model of user behavior based on principal component analysis," *Journal of Ambient Intelligence and Humanized Computing*, vol. 7, no. 4, pp. 547–554, 2016.

[7] S. Deep *et al.*, "A survey on anomalous behavior detection for elderly care using dense-sensing networks," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 352–370, 2019.

[8] K. N. Swaroop *et al.*, "A health monitoring system for vital signs using IoT," *Internet of Things*, vol. 5, pp. 116–129, 2019.

[9] D. Gupta *et al.*, "Future smart connected communities to fight covid-19 outbreak," *Internet of Things*, p. 100342, 2020.

[10] R. Jurdak *et al.*, "Wireless sensor network anomalies: Diagnosis and detection strategies," in *Intelligence-Based Systems Engineering*. Springer.

[11] C. Fu, Q. Zeng, and X. Du, "Hawatcher: Semantics-aware anomaly detection for appified smart homes."

[12] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, 1989.

[13] M. Yamauchi *et al.*, "Anomaly Detection in Smart Home Operation From User Behaviors and Home Conditions," *IEEE Transactions on Consumer Electronics*, vol. 66, no. 2, pp. 183–192, 2020.

[14] M. Zhang, A. Raghunathan, and N. K. Jha, "MedMon: Securing medical devices through wireless monitoring and anomaly detection," *IEEE Transactions on Biomedical circuits and Systems*, vol. 7, no. 6, 2013.

[15] K. Tonchev *et al.*, "Non-intrusive sleep analyzer for real time detection of sleep anomalies," in *2016 39th International Conference on Telecommunications and Signal Processing (TSP)*. IEEE, 2016, pp. 400–404.

[16] T. T. Pham *et al.*, "Wearable healthcare systems: A single channel accelerometer based anomaly detector for studies of gait freezing in parkinson's disease," in *IEEE Int. Conf. on Communications*, 2017.

[17] P. Parvin *et al.*, "Anomaly Detection in the Elderly Daily Behavior," in *IEEE 2018 14th International Conference on Intelligent Environments*.

[18] A. Ukil *et al.*, "IoT healthcare analytics: The importance of anomaly detection," in *2016 IEEE 30th international conference on advanced information networking and applications (AINA)*, 2016, pp. 994–997.

[19] D. Gupta *et al.*, "Access Control Model for Google Cloud IoT," in *Proc. of IEEE 6th Intl Conf. on Big Data Security on Cloud*, 2020.

[20] M. Gupta *et al.*, "Secure V2V and V2I communication in intelligent transportation using cloudlets," *IEEE Trans. on Services Comput.*, 2020.

[21] M. Gupta and R. Sandhu, "Towards activity-centric access control for smart collaborative ecosystems," in *Proceedings of the ACM Symposium on Access Control Models and Technologies*, 2021, pp. 155–164.

[22] D. Gupta *et al.*, "Learner's Dilemma: IoT Devices Training Strategies in Collaborative Deep Learning," in *2020 IEEE 6th World Forum on Internet of Things (WF-IoT)*, pp. 1–6.

[23] Ö. Aslan *et al.*, "Intelligent behavior-based malware detection system on cloud computing environment," *IEEE Access*, 2021.

[24] D. N. Monekosso *et al.*, "Behavior analysis for assisted living," *IEEE Transactions on Automation science and Engineering*, 2010.

[25] W. Kang *et al.*, "Detecting and predicting of abnormal behavior using hierarchical markov model in smart home network," in *IEEE Int. Conf. on Industrial Engineering and Engineering Management*, pp. 410–414.

[26] O. Kotevska *et al.*, "Kensor: Coordinated Intelligence from Co-Located Sensors," in *2019 IEEE International Conference on Big Data*.

[27] J. Li *et al.*, "Multivariate time series anomaly detection: A framework of Hidden Markov Models," vol. 60, pp. 229–240, 2017.

[28] S. Ramapatruni *et al.*, "Anomaly detection models for smart home security," in *IEEE 5th Intl Conference on Big Data Security on Cloud*.

[29] S. N. Narayanan *et al.*, "OBD_SecureAlert: An anomaly detection system for vehicles," in *IEEE Int. Conf. on Smart Computing*, 2016.

[30] H. Ren, Z. Ye, and Z. Li, "Anomaly detection based on a dynamic Markov model," *Information Sciences*, vol. 411, pp. 52–65, 2017.

[31] Y. Zhu, "Automatic detection of anomalies in blood glucose using a machine learning approach," *Journal of Communications and Networks*.

[32] S. Bhatt *et al.*, "An access control framework for cloud-enabled wearable Internet of Things," in *2017 IEEE 3rd International Conference on Collaboration and Internet Computing*, 2017, pp. 328–338.