AutoTag: Recurrent Variational Autoencoder for Unsupervised Apnea Detection with RFID Tags

Chao Yang*, Xuyu Wang* and Shiwen Mao
Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849-5201, USA
Email: {czy0017, xzw0029}@auburn.edu (∗: co-first authors), smao@ieee.org

***Abstract*—With the growth of smart healthcare in the Internet of Things (IoT), breathing monitoring and apnea detection are of increasing importance. In this paper, we propose AutoTag, a recurrent variational autoencoder model for breathing and apnea detection with commodity RFID Tags. The AutoTag system consists of signal extraction, calibration, and respiration monitoring modules. We propose a novel method to mitigate the frequency hopping offset with realtime calibration for FCC complaint RFID systems, and a new recurrent variational autoencoder method for apnea and breathing detection. Experimental results demonstrate the effectiveness of the proposed AutoTag system in two different environments.**

***Index Terms*—Radio-frequency identification (RFID); Vital sign monitoring; Deep learning; Apnea**

I. INTRODUCTION

With the aging population worldwide, healthcare has become an important problem [1]–[3]. As a key component of healthcare, detection and monitoring of vital signs are performed in traditional healthcare systems with dedicated equipment, such as capnography [4]. In addition, detecting an abnormality may require considerable efforts and experience. For example, many breathing disorders during sleep are hard to detect and diagnose because of the lacking of a suitable monitoring system. One of these disorders is obstructive sleep apnea, which can cause long-term damage on human health such as heart disease, stroke, and high blood pressure. Therefore, autonomous, unobstructive, and low-cost vital sign monitoring systems with the abnormality identification functionality are highly desirable, which can help a person to detect sleep disorders and reduce the danger of, e.g., sudden infant death syndrome (SIDS) for sleeping infants [5].

Different wireless signals have been used to detect vital signs. The idea is to capture the small signal caused by the rise and fall of the chest (or heart beats) from the received wireless signal. For example, Radar-based systems have been developed to monitor human respiration, such as ultra-wideband radar [6] and frequency modulated continuous wave (FMCW) radar [7]. However, a Radar-based system requires expensive and complicated hardware, and may operate on a wide spectrum, which may not be proper for many application scenarios. Other WiFi based techniques have the advantage of low-cost and easy deployment, which can monitor human respiration and heartbeat by analyzing the received signal strength (RSS) [8] or channel state information (CSI) [9]–[12]. Although WiFi based systems can effectively monitor human respiration, the WiFi signals are easily affected by changes in the environment, such as movements of a person nearby, thus leading to a relatively lower accuracy of breathing estimation and apnea detection. In a recent work [3], we developed SonarBeat, which is a smartphone app that exploits ultrasound signals for respiration monitoring.

RFID sensing systems have drawn increasing attention recently, which have been employed for object tracking [13], drone relays [14], orientation estimation [15], and recently, for breathing monitoring [16]. These works mainly exploit the RFID phase information, which is collected from the low level data with an RFID reader. For example, Tagoram system leverages phase information for real-time tracking of RFID tags with a differential augmented hologram technique [13]. RFly uses drones as relays for battery-free networks using RFID phase information [14]. As related to our work, Tagyro calibrates phase values from all channels to one channel for orientation estimation, where the system requires to measure the phase offset offline [15]. In addition, Tagbreathe monitors breathing signals by grouping the signals with the same channel index and using the calculated displacement in each channel [16]. However, this method does not work very well for US RFID systems, which operate in the frequency range from 902.5 MHz to 927.5 MHz with 50 channels as required by the FCC. This is because channel hoping among 50 different frequencies will cause a considerably larger latency, which make it much harder to obtain a breathing signal. Moreover, Tagbreathe does not consider apnea detection.

In this paper, we leverage phase information to monitor breathing and apnea with multiple attached RFID tags. A novel technique is proposed to mitigate the frequency hopping offset for FCC compliant RFID systems. We first model the RFID phase information as a function of distance, channel frequency, and phase offset. FCC requires channel hopping over 50 different frequencies every 0.2 s. Such an RFID system can hardly be used for vital sign monitoring, because of the large frequency hopping offset and latency. In this paper, we propose a novel method to map the phase data from each channel to one selected channel with realtime calibration, which is different from the Tagyro method that requires offline calibration. Moreover, the proposed scheme is a general method for mitigating the impact of channel hopping in different frequency bands. This method also has a much lower delay because it does not require grouping the signals with the same channel index.

In addition, we propose a deep learning based approach for apnea detection. A recurrent variational autoencoder model has been employed for sequence modeling [17] and human motion synthesis [18]. Motivated with the above works, we propose a modified recurrent variational autoencoder model for apnea detection, where Kullback Leibler (KL) divergence is used to measure the similarity between the reconstructed signal and the original signal. This method belongs to unsupervised learning, which does not require collecting labels for breathing and apnea signals, thus greatly reducing the workload. Moreover, this method achieves a better performance than the state-of-the-art energy threshold based method for apnea detection when the patient makes small movements. This is because the proposed method can learn the periodic characteristics of the breathing signal in an offline phase, which helps to easily distinguish the non-period signals from small body movements.

In particular, we design the AutoTag system, a recurrent variational **Auto**encoder model for unsupervised apnea detection with RFID **Tag**s. The AutoTag system includes signal extraction, calibration, and respiration monitoring modules. First, phase information is extracted from an RFID reader as in the signal extraction module. Then, we implement the calibration module including frequency hopping offset mitigation, movement detection, direct current (DC) removal, tag selection, downsampling, and filtering. For the respiration monitoring module, we propose a modified recurrent variational autoencoder approach for apnea and breathing signal detection. We then exploit a modified peak detection method for breathing rate estimation. We implement the AutoTag system with commodity RFID devices and evaluate its performance with four persons over a period of three months in two indoor environments, including a computer laboratory and a long corridor.

The main contributions of this paper are summarized in the following.

- To the best of our knowledge, this is the first work to leverage a modified recurrent variational autoencoder model for apnea detection.
- We design the AutoTag system including signal extraction, calibration, and respiration monitoring modules. We propose a novel method to mitigate the channel hopping offset with realtime calibration. Moreover, a modified peak detection is proposed for breathing rate estimation.
- We implement AutoTag with commodity RFID devices and evaluate its performance in two indoor environments. The experimental results verify the effectiveness of the proposed AutoTag system.

We first introduce the preliminaries in Section II. We then present the AutoTag system design in Section III and validate its performance in Section IV. Section V summaries this paper.

II. Preliminaries

Following FCC regulations, Ultra High Frequency (UHF) RFID readers adopt channel hopping to avoid co-channel interference. The UHF RFID readers operate in the 902.5 MHz to 927.5 MHz band, which is divided into 50 non-overlapping channels, each used for 0.2 second. We first provide an analysis on the causes of the phase offset due to channel hopping. From the RFID reader manual [19], the phase information φ can be written as

$$\varphi = \mathrm{mod}\left(\frac{2\pi d}{\lambda} + \alpha_T + \alpha_R + \alpha_{Tag}, 2\pi\right), \qquad (1)$$

where d is the round-trip distance between the tag and reader, λ is the wavelength, and α_T, α_R, and α_{Tag} are the phase offsets caused by the transmitter circuit, the receiver circuit, and the tag's reflection characteristics, respectively.

For commodity RFID Reader Impinj R420, we find that the phase offset between its two adjacent channels at the same distance d is not identical, which means that the phase offset is actually caused by the phase offsets α_T, α_R, and α_{Tag}. Because these three offsets are independent to the distance d, we can combine them into a single phase offset α_i for each channel i. Thus, the phase information $\varphi(f_i, d)$ for channel frequency f_i under round-trip distance d can be written as

$$\varphi(f_i, d) = \mathrm{mod}\left(\frac{2\pi f_i d}{c} + \alpha_i, 2\pi\right), \qquad (2)$$

where c is the the speed of light.

III. The AutoTag System

A. AutoTag System Architecture

The AutoTag system is designed to monitor human respiration and detect apnea with multiple attached RFID tags. As shown in (1), the phase information collected by an RFID reader is a function of the distance d between the corresponding tag and the reader antenna. When the tag is attached to the human body, this distance d changes periodically following the chest movement when the person is breathing. As a result, we can leverage the RFID phase information to recover the periodical signal of breathing and to detect apnea using multiple tags. However, there are many challenges for the system to accurately monitor human respiration and effectively detect apnea, such as the phase offset caused by channel hopping, tag differences, and disturbance from other movements nearby. To address these challenges, we design the AutoTag system with three modules, including signal extraction, calibration, and respiration monitoring, as shown in Fig. 1.

B. Signal Extraction

The signal extraction module is responsible for collecting low level data from tags. When human breaths, both the chest and abdomen movements generate the breathing signal. To increase the robustness of the system, we attach three passive RFID tags on the upper body. The reader is equipped with a directional antenna, which transmits RF signals to the tags and extracts the low-level data from the backscatter of these tags, which includes phase information, RSSI, Doppler shift, and time stamp. In addition to phase data, RSSI and Doppler shift may also carry the breathing signal. However, the resolution of RSSI is very low, and the signal-to-noise ratio (SNR) of

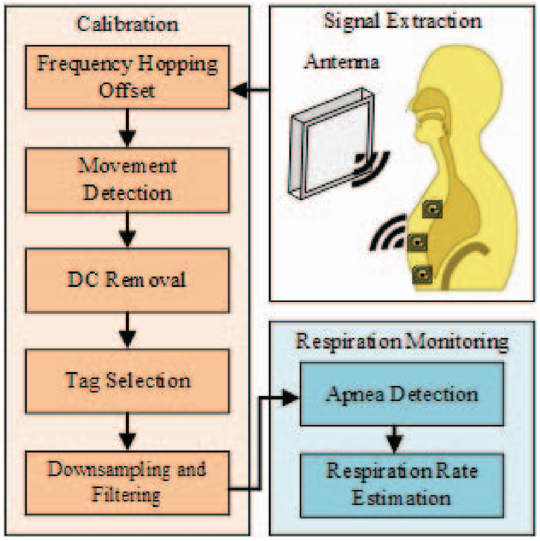

Fig. 1. The AutoTag system design.

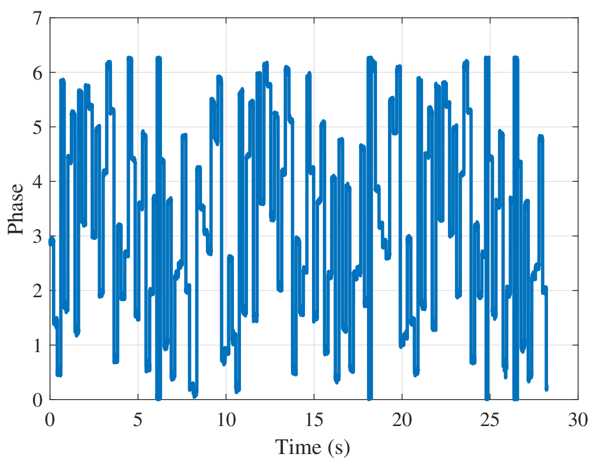

Fig. 2. Measured raw phase data.

Doppler shift is also very low because of the low velocity of human chest fluctuation. Thus both are not very useful for breathing and apnea detection. As a result, AutoTag uses phase information for respiration monitoring.

C. Calibration

Fig. 2 plots the raw phase data received from one tag for a period of 28 seconds. We can see that when the system hops among different channels (0.2 second per channel), the phase information exhibits big variations. Moreover, the phase data changes discontinuously from one channel to the next channel. It is extremely challenging to extract the human respiration signal directly from the received phase data; we need to calibrate the raw phase as follows in order to effectively extract the breathing signal.

1) Frequency Hopping Offset Mitigation: First, the phase information should be unwrapped to remove the large jump caused by the modulo operation, which constraints the phase information to be within 0 to 2π. As a result, a slight change of the real phase may cause a large variation in the received phase data. For example, when the real phase changes from 0.1π to -0.1π, the received phase value will jump from 0.1π to 1.9π due to modulo operation. Since the sample frequency of the RFID reader is larger than 100 Hz, the interval between two adjacent samples is less than 0.01 s. We assume that the change of two adjacent phase values is no larger than π in such a short time, so we can add $\pm 2\pi$ to restore the correct phase value when the variation is larger than π. Because channel hopping will also cause large jumps between two consecutive phases, the unwrapping will only be employed for continuous phase values when the same channel is used. After unwrapping, the phase values from each channel become stable.

Next, we will splice all the phase values from different channels into a single phase value, by effectively removing the frequency hopping offset. Our idea is to map the phase data from each channel to one chosen channel with realtime calibration. We choose the first channel as the target channel, and the phase information from all other channels will be mapped to the first channel. The phase from the target channel, i.e., channel 1, at distance d can be written as

$$\varphi(f_1, d) = \text{mod}\left(\frac{2\pi f_1 d}{c} + \alpha_1, 2\pi\right), \quad (3)$$

where f_1 and α_1 are the frequency and the initial phase offset of the target channel, respectively. We then map phase $\varphi(f_i, d)$ from each channel f_i to the target channel f_1, as

$$\varphi(f_1, d) = \frac{\varphi(f_i, d) f_1}{f_i} - \frac{\alpha_i f_1}{f_i} + \alpha_1, \ i = 2, 3, ..., 50. \quad (4)$$

Because the frequency for each channel i is known, we can multiply the phase $\varphi(f_i, d)$ from each channel i with a coefficient f_1/f_i to remove the frequency shift.

The next value we need to calibrate is $-\alpha_i f_1/f_i + \alpha_1$ in each channel i. The reader has a high sampling rate over 500 Hz. Thus we assume that the distance and the environment is basically the same when the channel hops from one channel to another. Based on this assumption, the last phase value from the previous channel should be equal to the first phase value from the current channel. Since the received signal is affected by thermal noise, we firstly filter the signal in each channel with a Hampel filter using a 20-sample sliding window and a threshold of 0.01. After filtering, we consider the mean of the last six data in the former channel and the mean of the first six data in the current channel, to obtain the frequency hopping offset as $-\alpha_i f_1/f_i + \alpha_1$. After mitigating the frequency hopping offset, the phase value in the current channel can be expressed by the target phase value $\varphi(f_1, d)$. Fig. 3 shows the calibrated phase data after removing the frequency hopping offset. It is noticed that the calibrated phase data exhibits an obvious breathing signal, which is in great contrast to the raw phase data shown in Fig. 2. The channel hopping effect has been effectively removed.

2) Movement Detection: After dealing with the frequency hopping offset, AutoTag next detects whether the subject is in a stationary state. The respiration signal is an extremely

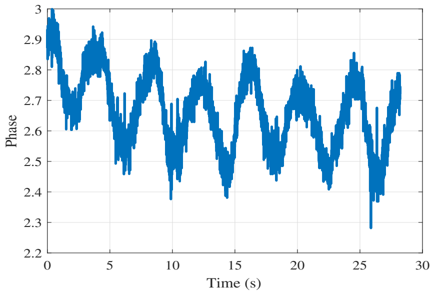

Fig. 3. Calibrated phase data by removing the frequency hopping offset.

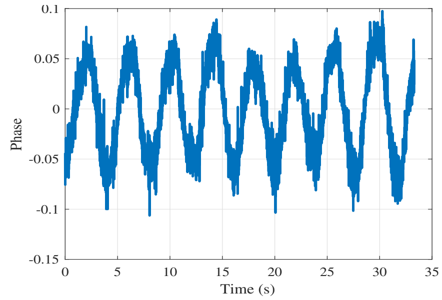

Fig. 4. Calibrated phase data after DC removal.

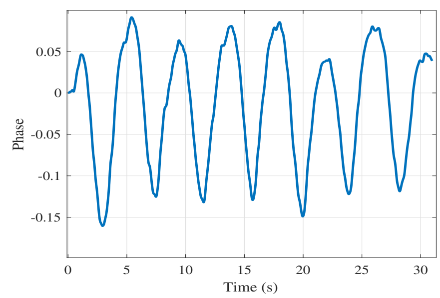

Fig. 5. Recovered breathing signal.

weak signal comparing to other movements of human body. To eliminate the influence caused by large movements, only signals extracted in the stationary state will be used to monitor human respiration and detect apnea.

We propose a threshold based method to decide whether the person is stationary or not. Let T denote the sum of mean absolute deviations of phase values for all RFID tags in a sliding window, which is given by

$$T = \frac{1}{|W|} \sum_{j=1}^{N} \sum_{k \in W} |\varphi^j(k) - \mathbb{E}(\varphi^j(k))|, \tag{5}$$

where W is the index set of all the packets in the sliding window, $|W|$ is the length of the sliding window, N is the number of RFID tags, and $\varphi^j(k)$ is the phase value for packet k from tag j. If the subject is moving, the phase value will have larger variations (due to large variations in d). Thus, the threshold-based method can detect large movements. In our experiments, when the window size is larger than 6 s, the T value of human breathing signal within a window is smaller than 0.9; thus we set 0.9 as the threshold for movement detection.

3) DC Removal: Although large movements of the body can be detected and excluded using the above threshold based method, there are still some small movements that remain. Such small movements will also cause phase shifts, which is not related to and interfere with the human respiration signal. Furthermore, the initial offset of the first channel is random, leading to a different DC components in the signal for every new experiment.

To remove all unrelated factors in the phase signal, the signal should be detrended before respiration estimation and apnea detection. To this end, we leverage a Hampel filter with a windows size of 2000 and a threshold of 0.001 to obtain the DC component of the entire signal. Then, the finally calibrated signal is obtained by subtracting the trend from the filtered signal. Fig. 4 illustrates the calibrated signal after DC removal. We can notice that the mean value of the signal is almost zero, and the tendency of the signal is also removed.

4) Tag Selection: In our experiments, we find that the sensitivity of the tags is different, because the corresponding propagation environment for each tag is different. Moveover, the angle of the tag to the reader antenna is also another reason to cause the difference in sensitivity. We calculate the signal strength by using mean absolute deviation within a certain window size of 6 s. To obtain the most sensitive signal, only the tag that has the strongest signal strength will be selected for the following processes.

5) Downsampling and Filtering: Since the sampling frequency of RFID radar is very high, which is about 600 Hz when only one antenna works, the signal should be downsampled to reduce the complexity of the following training algorithm. In our experiments, we downsample the signal by a factor of 10. In addition, there are many false peaks caused by thermal noise, making it hard to detect the peaks corresponding to the inhale and exhale movements. Since the normal human breathing rate is lower than 0.5 Hz, we apply a low-pass filter with cutoff frequency of 0.5 Hz on the downsampled signal to reduce the high frequency noise. The filtered signal is plotted in Fig. 5. We can see that the signal is sufficiently smooth, and can be useful for the following apnea detection and breathing rate estimation processes.

D. Breathing Signal Detection

1) Apnea Detection Based on Recurrent Variational Autoencoder: For apnea detection, our basic idea is to leverage a variational autoencoder for measuring the difference between the input signal sequence and the reconstructed signal sequence within a window, which is an unsupervised learning approach [20], [21]. When the difference is larger than a given threshold, this signal sequence is considered to be a breathing signal; otherwise, it is considered to be a detection of apnea. Note that we have used the energy based threshold for movement detection, and thus small signals can be detected in this phase.

First, we model the variational autoencoder method to obtain the reconstructed input data. This model maximizes the marginal likelihood $p_\theta(x) = \int p_\theta(x|z)p(z)dz$, where x, z, and θ are the observed variables, the latent random variables, and the parameter set, respectively; $p(z)$ is the prior over the latent random variables z; and $p_\theta(x|z)$ is the conditional probability, which is an observation model under the parameter set. Note that $p_\theta(x)$ is intractable for the above model because of the integral operation. Although Monte Carlo sampling methods can solve the model, it has a high computational cost even for small-sized data samples. The variational autoencoder model employs the variational approximation $q_\phi(z|x)$ instead of the true posterior $p_\theta(z|x)$. Generally, the variational autoencoder method considers $q_\phi(z|x)$ with parameters set ϕ as the encoder and $p_\theta(x|z)$ with parameters set θ as the decoder. Based on Jensen's inequality, the variational autoencoder model can optimize sets ϕ and θ by maximizing a lower bound on the log-likelihood, denoted by L, which is formulated as [20]

$$\max L = -D_{KL}(q_\phi(z|x)||p(z)) + \mathbb{E}_{z \in q_\phi(z|x)}\left[p_\theta(x|z)\right], \quad (6)$$

where D_{KL} means the KL divergence. The first term in (6) is the regularization over the latent variable z. The second term can reflect the autoencoder, where the latent variable z can be sampled from $q_\phi(z|x)$, and the reconstructed x can be obtained from $p_\theta(x|z)$.

For effective training, the variational autoencoder model adopts the reparametrization trick, where the latent vector z can be computed by the mean vector $\mu_\phi(x)$ and the variance vector $\sigma_\phi^2(x)$, as

$$z = \mu_\phi(x) + \sigma_\phi(x) \odot \epsilon, \quad (7)$$

where $\epsilon \in N(0,1)$, and $\odot$ is the element-wise product operation. Then, the lower bound on the log-likelihood L can be approximately by

$$L \approx \frac{1}{2}\sum_{j=1}^{J}(1+\log(\sigma_j^2(x)) - \mu_j^2(x) - \sigma_j^2(x)) + \frac{1}{M}\sum_{l=1}^{M}\log p_\theta(x|z_l), \quad (8)$$

where M is the number of sampling z, and J is the dimensionality of z.

We treat the breathing signal in a time window as a time sequence, and employ a long short-term memory (LSTM) network to process the input breathing data. LSTM is a type of recurrent neural network (RNN) that can effectively capture the long-range dependency in data. In our AutoTag system, an LSTM network fist encodes the breathing data sequence in a window. Its output is then used to estimate the mean vector $\mu_\phi(x)$ and the variance vector $\sigma_\phi^2(x)$ using two linear modules. Then, the sampling z can be fed to another LSTM network for decoding the estimated mean vector $\mu_\phi(z)$ and variance vector $\sigma_\phi^2(z)$, thus obtaining the reconstructed breathing signal in the same window.

After obtaining the reconstructed breathing signal, we propose a KL divergence method for apnea detection. First, both the reconstructed signal and the original signal in a window should be normalized, to guarantee that both the breathing signal and the apnea signal are at the same amplitude level. Then, KL divergence is used to compute the similarity between the reconstructed signal and the original signal, where a proper threshold λ is set for determining whether the measured signal in the window is apnea or not. If the KL divergence is larger than threshold λ, the proposed AutoTag system considers the signal is an apnea signal. Otherwise, the signal is a normal breathing signal when KL divergence is smaller than threshold

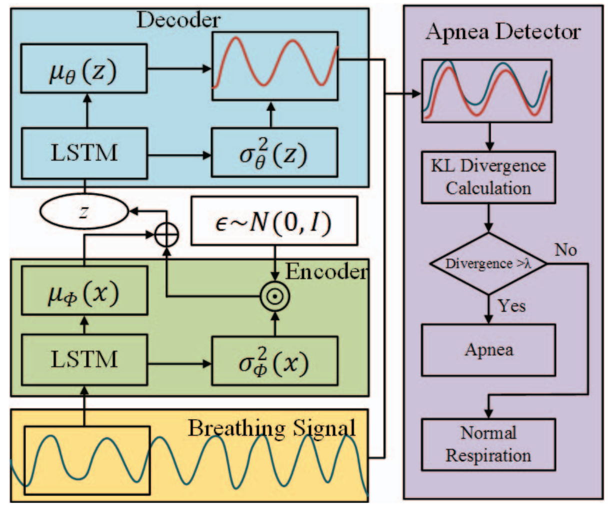

Fig. 6. Recurrent variational autoencoder for unsupervised apnea detection.

λ. This proposed method has two main advantages. First, the AutoTag system leverages a recurrent variational autoencoder, which is an unsupervised learning approach. Thus, this method does not require for collecting labels for two types of signals, thus reducing the work load. Second, this method is greater than the energy based threshold method for apnea detection when the person makes small movements. This is because the proposed method can learn the periodic characteristics of breathing signals in an offline phase, and can easily recognize the non-period signal caused by other small movements.

2) Breathing Rate Estimation Based on Peak Detection: We employ peak detection to measure the average intervals between two adjacent peaks. Although most noise has been removed in the previous steps, there are still some false peaks, which are hard to distinguish from the peaks of inhaling or exhaling. To avoid the effect of these false peaks, the peak detection algorithm should operate with a sliding window. Only when the median of the sliding window becomes the maximum in the window, the median can be identified as a peak. Then the period of the breathing signal will be estimated by averaging all peak-to-peak intervals, which is represented as t. The respiration rate can be estimated as $60/t$ breaths per minute (bpm).

IV. Experimental Study

A. Implementation and Test Configuration

In AutoTag, we use a commodity RFID reader Impinj R420, equipped with a directional antenna, to receive phase information from ALN-9740 tags. Frequency hopping is among 50 channels in the band from 902.5 MHz to 927.5 MHz, which is FCC-compliant. We also employ an MSI laptop to implement a user interface, which is also responsible for signal processing. The software used in our implementation is based on an RFID library, which allows the laptop to communicate with the RFID reader through a Low-Level Reader Protocol (LLRP). The RFID reader can extract RSSI, phase information, Doppler shift, and time stamp from received tag responses. In addition,

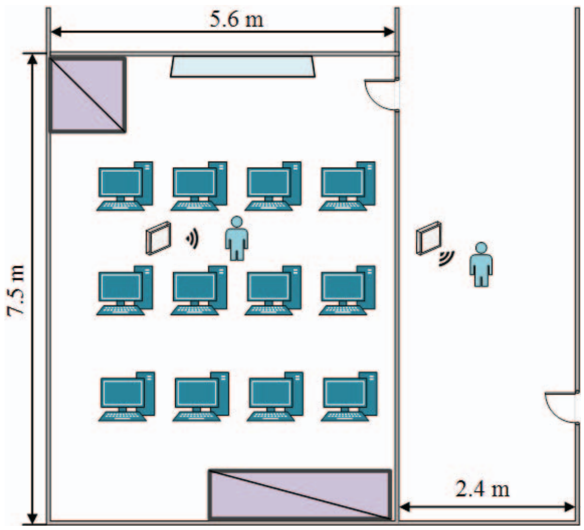

Fig. 7. Experimental setup for performance validation.

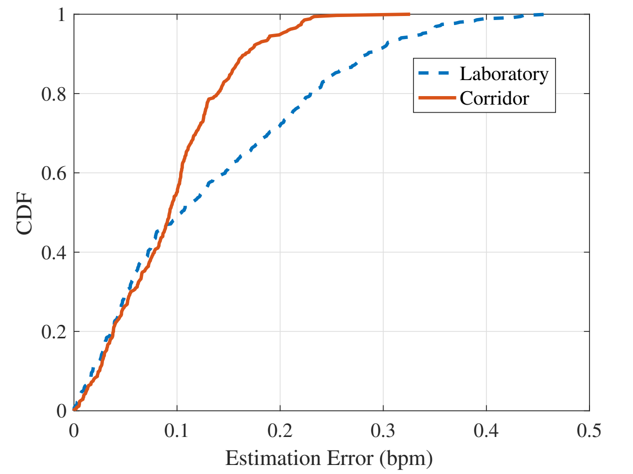

Fig. 8. CDFs of breathing rate estimation in the computer laboratory and corridor scenarios.

TABLE I
TN AND TP RATES USING THE PROPOSED AUTOTAG METHOD AND THE ENERGY BASED METHOD IN THE STATIONARY ENVIRONMENT

	Energy based method	AutoTag
TN rate	0.95688	0.94694
TP rate	0.88119	0.92371

TABLE II
TN AND TP RATES USING THE PROPOSED AUTOTAG METHOD AND THE ENERGY BASED METHOD WHEN THERE ARE SMALL BODY MOVEMENTS

	Energy based method	AutoTag
TN rate	0.95443	0.91786
TP rate	0.56116	0.92328

we implement the recurrent variational autoencoder model using the Tensorflow backend on the MSI laptop with an Intel(R) Core(TM) i7-6820HK CPU, and a Nvidia GTX1080 GPU.

We conduct extensive experiments with four volunteers and present in this section the experimental results in two different environments. The test scenarios include a 7.5 m $\times$ 5.6 m laboratory which is croweded with computers and desks, and a 20 m $\times$ 2.4 m corridor with no obstacles. The multipath effect in the laboratory is larger than that in the corridor, because the desks and other furniture can cause more reflections in the received RF signal. All volunteers are required to perform three different movements including breathing normally while sitting in the chair, holding breath while sitting in the chair, and moving randomly.

For breathing rate estimation, we adopt the cumulative distribution function (CDF) of estimation errors as performance metric. For apnea detection, we consider two metrics to evaluate the performance of the proposed AutoTag system, including the True Negative (TN) rate and True Positive (TP) rate. The TN rate is defined as the success rate that the breathing signal is correctly detected, while the TP rate is defined as the success rate that apnea is correctly detected.

B. Experimental Results and Discussions

For breathing rate estimation, we test the accuracy of the estimated respiration rate in the two environments. The NEULOG Respiration Belt Sensor is used to record the ground truth. Fig. 8 presents the CDFs of breathing estimation errors in the computer laboratory and corridor scenarios. We can see that the maximum error in the computer laboratory is 0.1 bpm larger than that in the corridor. Moreover, over 50% estimation errors tested in corridor case is smaller than the errors tested in the laboratory. This indicates that the multipath effect has certain influence on breathing rate estimation in the AutoTag system. However, we also find that the median errors in both laboratory and corridor cases are both about 0.1 bpm, which means that the influence of multipath effect is not very serious. Since the RFID reader uses a directional antenna, the backscatter signal in the light of sight (LOS) path is much stronger than other multipath components. Thus, the multipath effect is negligible, leading to accurate breathing estimation in both the corridor and laboratory environments.

We also compare the proposed AutoTag system with the energy based method presented in [22] in the two scenarios. In the first scenario, the volunteer sits stationarily without any other movements, while in the other scenario, the volunteers can move slightly, such as moving legs, twisting neck, and shaking hands. Table I presents the TN and TP rates using the proposed AutoTag method and the energy based method in the stationary environment. We can see that when the user is completely stationary, the TN rates for the two methods are both above 94%. Moreover, the TP rates using the energy based method and the proposed AutoTag method are 88% and 92%, respectively. This shows both methods can achieve considerably high success rates when the subject is completely stationary.

However, quite different results are obtained for the second slight-movement scenario as shown in Table II. The TP rate decreases to 56% using the energy based method, while the TP rate obtained by AutoTag still remains very high at 92%. This is because the energy of the apnea signal is increased a lot under the disturbance of small movements. Thus, it is hard to set a proper energy threshold for apnea detection. On the other hand, our method is based on matching the shapes of the reconstructed signal and the original signal, rather than on the energy of the signal. Thus, the energy disturbance in this test can hardly affect the performance of the proposed AutoTag system. In addition, in Table II, we find the TN rates using the proposed AutoTag system and the energy method are both above 91%, which are slightly lower than that in the stationary environment. This is because the shape of breathing signal is corrupted a little by the small body movements. The proposed AutoTag method can effectively detect apnea and breathing signals in both experiments.

V. Conclusions

We presented the AutoTag system in this paper, for breathing and apnea detection using a recurrent variational autoencoder model with commodity RFID Tags. We designed the AutoTag system design, including signal extraction, calibration and respiration monitoring modules. A novel technique was proposed to effectively mitigate the frequency hopping offset in phase data for FCC compliant RFID systems. Experimental results in two indoor environments validated the effectiveness of the proposed AutoTag system.

Acknowledgment

This work is supported in part by the US NSF under Grant CNS-1702957, and through the Wireless Engineering Research and Education Center (WEREC) at Auburn University.

References

[1] O. Boric-Lubeke and V. Lubecke, "Wireless house calls: Using communications technology for health care and monitoring," *IEEE Microwave Mag.*, vol. 3, no. 3, pp. 43–48, Apr. 2002.

[2] X. Wang, X. Wang, and S. Mao, "RF sensing for Internet of Things: A general deep learning framework," *IEEE Communications*, vol. 56, no. 8, Aug. 2018.

[3] X. Wang, R. Huang, and S. Mao, "SonarBeat: Sonar phase for breathing beat monitoring with smartphones," in *Proc. ICCCN 2017*, Vancouver, Canada, July/Aug. 2017, pp. 1–8.

[4] M. L. R. Mogue and B. Rantala, "Capnometers," *J. Clinical Monitoring*, vol. 4, no. 2, pp. 115–121, Apr. 1988.

[5] C. Hunt and F. Hauck, "Sudden infant death syndrome," *Can. Med. Assoc. J*, vol. 174, no. 13, pp. 1309–1310, Apr. 2006.

[6] J. Salmi and A. F. Molisch, "Propagation parameter estimation, modeling and measurements for ultrawideband MIMO radar," *IEEE Trans. Microw. Theory Technol.*, vol. 59, no. 11, pp. 4257–4267, Nov. 2011.

[7] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. Miller, "Smart homes that monitor breathing and heart rate," in *Proc. ACM CHI'15*, Seoul, Korea, Apr. 2015, pp. 837–846.

[8] H. Abdelnasser, K. A. Harras, and M. Youssef, "Ubibreathe: A ubiquitous non-invasive WiFi-based breathing estimator," in *Proc. IEEE MobiHoc'15*, Hangzhou, China, June 2015, pp. 277–286.

[9] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng, "Tracking vital signs during sleep leveraging off-the-shelf WiFi," in *Proc. ACM Mobihoc'15*, Hangzhou, China, June 2015, pp. 267–276.

[10] X. Wang, C. Yang, and S. Mao, "PhaseBeat: Exploiting CSI phase data for vital sign monitoring with commodity WiFi devices," in *Proc. IEEE ICDCS 2017*, Atlanta, GA, June 2017, pp. 1–10.

[11] ——, "TensorBeat: Tensor decomposition for monitoring multi-person breathing beats with commodity WiFi," *ACM Trans. Intelligent Systems and Technology*, vol. 9, no. 1, pp. 8:1–8:27, Sept. 2017.

[12] ——, "ResBeat: Resilient breathing beats monitoring with online bimodal CSI data," in *Proc. IEEE GLOBECOM 2017*, Singapore, Dec. 2017, pp. 1–6.

[13] L. Yang, Y. Chen, X.-Y. Li, C. Xiao, M. Li, and Y. Liu, "Tagoram: Real-time tracking of mobile RFID tags to high precision using COTS devices," in *Proc. ACM MobiCom'14*, Maui, HI, Sept. 2014, pp. 237–248.

[14] Y. Ma, N. Selby, and F. Adib, "Drone relays for battery-free networks," in *Proc. ACM SIGCOMM'17*, Los Angeles, CA, Aug. 2017, pp. 335–347.

[15] T. Wei and X. Zhang, "Gyro in the air: tracking 3D orientation of batteryless internet-of-things," in *Proc. ACM MobiCom'16*, New York City, NY, Oct. 2016, pp. 55–68.

[16] Y. Hou, Y. Wang, and Y. Zheng, "Tagbreathe: Monitor breathing with commodity RFID systems," in *Proc. IEEE ICDCS'17*, Atlanta, GA, June 2017, pp. 404–413.

[17] J. Chung, K. Kastner, L. Dinh, K. Goel, A. C. Courville, and Y. Bengio, "A recurrent latent variable model for sequential data," in *Proc. NIPS'15*, Montreal, Canada, Dec. 2015, pp. 2980–2988.

[18] I. Habibie, et al., "A recurrent variational autoencoder for human motion synthesis," *IEEE Computer Graphics and Applications*, vol. 37, p. 4, 2017.

[19] Impinj, "Low level user data support," 2013, note, Impinj Speedway Revolution Reader Application.

[20] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint, arXiv:1312.6114*, 2013.

[21] O. Fabius and J. R. van Amersfoort, "Variational recurrent auto-encoders," *arXiv preprint, arXiv:1412.6581*, 2014.

[22] L. Almazaydeh, K. Elleithy, M. Faezipour, and A. Abushakra, "Apnea detection based on respiratory signal classification," *Elsevier Procedia Computer Science*, vol. 21, pp. 310–316, 2013.