Check for updates

Beyond Respiration: Contactless Sleep Sound-Activity Recognition Using RF Signals

CHEN LIU, University of Victoria, Canada. Northwest University, China
JIE XIONG, University of Massachusetts Amherst, USA
LIN CAI, University of Victoria, Canada
LIN FENG, XIAOJIANG CHEN, and DINGYI FANG, Northwest University, China. Shaanxi International Research Center for Passive Internet of Things, China

Sleep sound-activities including snore, cough and somniloquy are closely related to sleep quality, sleep disorder and even illnesses. To obtain the information of these activities, current solutions either require the user to wear various sensors/devices, or use the camera/microphone to record the image/sound data. However, many people are reluctant to wear sensors/devices during sleep. The video-based and audio-based approaches raise privacy concerns. In this work, we propose a novel system TagSleep to address the issues mentioned above. For the first time, we propose the concept of two-layer sensing. We employ the respiration sensing information as the basic first-layer information, which is applied to further obtain rich second-layer sensing information including snore, cough and somniloquy. Specifically, without attaching any device to the human body, by just deploying low-cost and flexible RFID tags near to the user, we can accurately obtain the respiration information. What's more interesting, the user's cough, snore and somniloquy all affect his/her respiration, so the fine-grained respiration changes can be used to infer these sleep sound-activities without recording the sound data. We design and implement our system with just three RFID tags and one RFID reader. We evaluate the performance of TagSleep with 30 users (13 males and 17 females) for a period of 2 months. TagSleep is able to achieve higher than 96.58% sensing accuracy in recognizing snore, cough and somniloquy under various sleep postures. TagSleep also boosts the sleep posture recognition accuracy to 98.94%.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing systems and tools**.

Additional Key Words and Phrases: Contactless Sensing, Sleep sound-activity, Sleep Postures, Machine Learning, RFIDs

ACM Reference Format:
Chen Liu, Jie Xiong, Lin Cai, Lin Feng, Xiaojiang Chen, and Dingyi Fang. 2019. Beyond Respiration: Contactless Sleep Sound-Activity Recognition Using RF Signals. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 96 (September 2019), 22 pages. https://doi.org/10.1145/3351254

1 INTRODUCTION

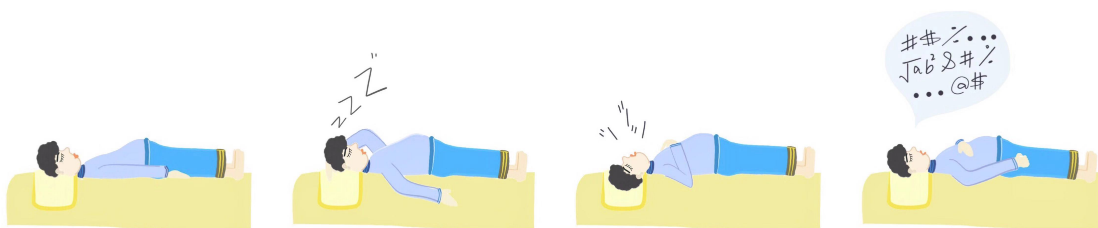
Sleep quality is an important factor affecting an individual's health. Snore, cough, and somniloquy (see Fig. 1) are common sleep activities closely related to sleep quality. It was reported that one in every three healthy people aged thirty and above snore during sleep, and this number rises to 40% for middle-aged people [53]. Even more severe, multiple studies revealed a positive correlation between sleep sound-activities and health problems, such as sleep

Authors' addresses: Chen Liu, liuchen@nwu.edu.cn, University of Victoria, Victoria, Canada. Northwest University, Xi'an, China; Jie Xiong, jxiong@cs.umass.edu, University of Massachusetts Amherst, Massachusetts, USA; Lin Cai, cai@ece.uvic.ca, University of Victoria, Victoria, Canada; Lin Feng; Xiaojiang Chen; Dingyi Fang, Northwest University, Xi'an, China. Shaanxi International Research Center for Passive Internet of Things, Xi'an, China.

2474-9567/2019/9-ART96 $15.00
https://doi.org/10.1145/3351254

disorders [3], heart attack [54], sudden infant death syndrome (SIDS) [31], sleep apnea-hypopnea [12], night coughing [7], and confusional arousals [19]. Therefore, continuous monitoring of these sleep sound-activities[1] is critical to collect vital data for sleep analysis and early detection of many diseases.

Traditionally, sleep-sound activities can be detected and identified by the Laboratory-based Polysomnography (PSG) in hospital, where a dozen of sensors are attached to a patient's body [30]. Although PSG can achieve high accuracy and provide rich information for sleep sound-activity analysis, it requires a well-trained technician to assist the user to wear sensors properly, which is expensive, intrusive and discomfortable for long-term monitoring. Some portable systems have been designed for in-home monitoring [15, 42]. However, they still require the microphone and sensors attached to the user's nose and throat, which may interrupt user's sleep behavior and even disrupt sleep.



(a) Respiration. (b) Snore. (c) Cough. (d) Somniloquy.

Fig. 1. Sleep sound-activities. [@Daily.J]

Recently, some works use the data recorded by the camera or built-in microphone of the smartphone or smartwatch to detect these sleep sound-activities without wearing sensors [8, 23, 32, 51, 58]. However, video-based solutions [5, 33] may cause privacy concern and require line-of-sight view and good lighting conditions. The performance of audio-based solutions degrades in noisy environment and privacy issue still exists as somniloquy may contain personal information. More importantly, solely using the sound signal without the knowledge of chest/abdomen movement is insufficient to further detect and alarm many diseases [45]. Furthermore, besides sleep sound-activity, sleep posture is another important metric for sleep quality assessment [14] and related to health problems such as apnea and back pain [41], especially for elderly [37] and pregnant women [43]. However, audio-base solutions cannot recognize sleep posture.

In this paper, we design a two-layer RFID sleep monitoring system, TagSleep, to detect and recognize sleep-sound activities by using RF signals alone without using a microphone (i.e., without compromising privacy) or any sensor attached to the human body (i.e., in a contactless manner). We employ commercial COTS RFID hardware to sense these sleep sound-activities through chest/abdomen movement detection. Specifically, with one RFID tag and one reader deployed on the two sides of the bed, besides the direct path signal between the reader and tag, there are reflected paths from the human body. Thus, the signal at the receiver is a combination of the reflected signal and the LOS signal. The chest movement (around 5 mm) during respiration will cause the path length of the reflected signal change. This path length change induces signal variation at the receiver in terms of both amplitude and phase. Our design is based on the following key *observation*: during snore, cough, and somniloquy, the amount of chest displacements and the frequency of respiration change. The effect of each activity on the signal variation is unique when we combine sufficient features. Therefore, we consider the respiration information measured by the wireless signals as the first-layer information. Based on this, we can further accurately infer the second-layer information such as snore, cough and somniloquy by carefully processing the phase changes with machine learning algorithms. However, it is non-trivial to realize TagSleep due to the following *challenges*:

[1]For simplicity, we define snore, cough and somniloquy as sleep sound-activities in this paper.

- *How to obtain fine-grained respiration information?* To infer the sleep sound-activities, we need to obtain not just the respiration rate change but also the subtle respiration depth variation. This is much more challenging than just obtaining the respiration rate.
- *How to deploy just a few RFID tags to achieve high accuracy throughout the night?* A user may sleep at one location initially and move to another location overnight. Due to the well-known blind zone issue [72], one tag is not able to cover all the locations on the bed. Employing as few as possible tags to cover the whole bed for monitoring is non-trivial.
- *How to select the features from the RFID phase measurements?* To infer the sleep sound-activities, we need to obtain not just the respiration rate change but also other features to quantify the correlation between the sound-activities and the changing pattern of RF signals. This is more challenging than just calculating the periodical respiration rate.
- *How to achieve robust sleep sound-activity recognition across different sleep postures?* Even for the same activity, the phase change pattern varies when the user has different sleep postures. Without knowing the sleep posture, recognizing sleep sound-activity is challenging.

To deal with the above challenges, first, we adopt the phase variation rather than the amplitude variation because the phase readings have a much finer resolution (0.0015 radians). We propose an enhancement scheme by using the reflection from the nearby wall/furniture to enlarge the phase change caused by the respiration. Second, we design an effective tag deployment layout with 3 tags by considering the geometric relationship among the user, RFID reader and tags. Note that the 3 tags are not required to be precisely located at specific locations. Roughly deploying them in three regions suffices our sensing purpose. Third, by jointly considering the time-domain, frequency-domain and sample entropy features, TagSleep extracts a set of effective features that vary significantly among different sleep sound-activities while remaining stable within a given activity. Last, we design a sleep posture recognition method with 3 tags. By recognizing sleep posture first, TagSleep achieves reliable sound-activity recognition performance under different postures.
Contributions: This paper makes the following main contributions.

- To our best knowledge, this is the first paper that proposes and deploys two-layer sensing. The proposed two-layer sensing concept can be applied to enable many new sensing applications.
- We propose a method of employing the surrounding wall/furniture to obtain a 'clearer' phase change for subtle periodic movement such as respiration-related activity. Furthermore, we design an effective deployment layout to enable a robust sleep sound-activity recognition across different sleep locations.
- We implement TagSleep with RFID device and evaluate the performance with extensive real-world experiments involving 30 users (13 males and 17 females) for a period of 2 months. The results demonstrate that with RF signals alone, TagSleep is able to achieve higher than 96.58% accuracy in recognizing snore, cough and somniloquy in person-dependent classification, and even 92.27% accuracy for person-independent classification. TagSleep also boosts the sleep posture recognition accuracy significantly to 98.94%.

2 RELATED WORK

Recent years have witnessed a growing interest in sleep monitoring for healthcare. Laboratory-based Polysomnography (PSG) is used for sleep monitoring [30] in hospitals, where a dozen of sensors are attached to a patient's body. Although this approach can achieve high accuracy, it typically requires a well-trained professional to assist users to wear dedicated devices which are expensive, intrusive and discomfortable for long-term monitoring. Continuous in-home sleep monitoring has received a lot of research interests [22, 23, 25, 38, 50]. These approaches can be broadly grouped into three categories: Audio/video-based, Sensor-based and RF-based approaches.

2.1 Audio/video-based Approaches

The existing works [6, 17, 32, 51, 58] recognized sleep sound-activities, such as snore and cough, mainly using sound data recorded by the bulit-in microphone in the *smartphone* (e.g., iSleep [23], Sleep hunter [22]), *smartwatch* (e.g., SleepGuard [8], ubiSleep [49]) and *headband* (e.g., Sleep Profiler [44]). Other works used a camera to monitor sleep from the video stream [5, 33]. Although their recognition accuracy is high, using sound data/visual data may raise privacy concerns. More importantly, solely using the sound signal without chest/abdomen movement and the sleep posture information is not sufficient to detect sleep-related diseases [45].

2.2 Sensor-based Approaches

Some other works used sensors including infrared, sonar, pressure sensors and accelerometers for sleep monitoring. (i) The infrared (Yang *et al.* [66]) and sonar (ApneaApp [45]) based approaches can only detect the body movements during sleep, which cannot detect sleep sound-activities and sleep postures. (ii) The built-in pressure sensors in the bed pad [56] are employed to detect the respiration, snore and body movements. However, the pressure sensor-based approaches cannot provide sleep posture and other important sound-activities, like cough and somniloquy. Besides, the built-in sensors in the bed pad may be tore easily by body movements. (iii) For accelerometer-base approaches, most sensors are built in the *smartphone* (e.g., Sleep Cycle [13], Sleepbot [57] and S plus [44]), *smartwatch* (e.g., SleepGuard [8] and Sleepmonitor [59], SleepHunter [22], FitBit [18] and Jawbone [29]) and *bed pad* (e.g., Hoque *et al.* [24] and Beddit [6]). These works can detect the body movements and even sleep postures (e.g., SleepGurad and Sleepmonitor), but cannot recognize sleep sound-activities. Also, as mentioned earlier, some people may not like to wear smartwatch/wearable devices during sleep, especially the elderly [20].

2.3 RF-based Approaches

Various wireless technologies including Radar, FMCW, Millimeter wave and Wi-Fi have been proposed for sleep monitoring in a contactless manner. Among them, the *Radar* (e.g., Dopplersleep [50], SleepSense [34]), *FMCW* (e.g., EZ-Sleep [25], Zhao *et al.* [75], Vital-Radio [2] and DeepBreath [70]) and *Millimeter wave*-based schemes (e.g., mmVital [67, 68]) are able to achieve high accuracy of respiration rate, heartbeat rate, body movements and even sleep posture detection. However, these approaches required either dedicated or customized hardware platforms. Since Wi-Fi devices are ubiquitous, some recent works employed the fine-grained CSI (Channel State Information) information for monitoring respiration rate (e.g., FullBreathe [71], UbiBreathe [1], WiBreathe [52], Niu *et al.* [46], Liu *et al.* [36] and PhaseBeat [62]) and sleep posture (e.g., Wi-Sleep [38, 39]). However, these approaches still may not recognize sleep sound-activities. Recent works embedded dozens of RFID tags into the bed cloth [26] or under a bed sheet [55] to detect respiration rate, body movements and sleep postures. Again, these approaches cannot recognize snore, cough and somniloquy. TagSleep differs from the existing systems by considering the respiration as the basic first-layer information to obtain rich second-layer sensing information including snore, cough, somniloquy and sleep posture. Indeed, the idea of two-layer sensing of TagSleep can be extended to other RF-based techniques as well. TagSleep extends the scope of current RF-based respiration monitoring, and it is the first work that can recognize sleep sound-activities by solely using RF signals.

3 PRELIMINARY

When a user lies between the antenna and the tag (Fig. 2(a)), there is an additional reflection path from the user (denoted as dynamic path H_d) besides the LOS path and other reflection paths from the wall and furniture (denoted as static path H_s). Thus, the phase value measured at the reader is the superposition of multiple signals as $H = H_s + H_d$.

As shown in Fig. 2(b) and Fig. 2(c), the received signal phase θ can be notated as $\theta = \beta - \alpha$ or $\theta = \beta + \alpha$ according to different values of Fresnel phase ρ (denoted as the phase difference between H_s and H_d). Here, α is

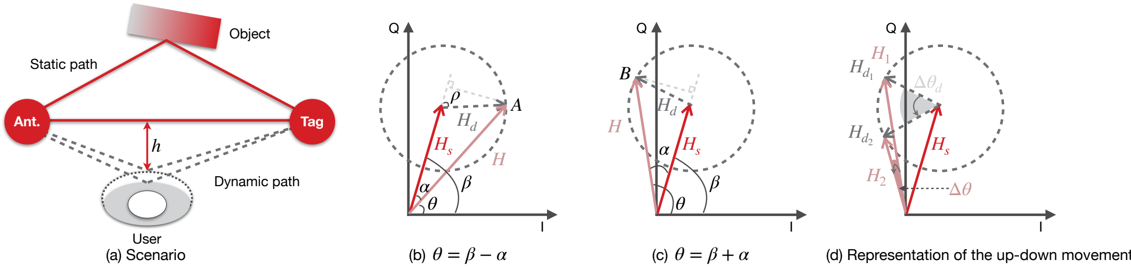


Fig. 2. The representation of the phase change caused by chest movements.

the phase difference between H and H_s, and β is the phase of H_s. Taking the case $\theta = \beta + \alpha$ as an example, θ can be approximated [71] as

$$\theta \approx \beta + \frac{|H_d| \sin \rho}{\sqrt{|H_s|^2 + |H_d|^2 + 2|H_s||H_d| \cos \rho}}. \tag{1}$$

To explain the relationship between the chest movement and the phase of the received signal, for simplicity, we denote two states: the *up state* and the *down state* to describe the chest movements. As shown in Fig. 2(d), when the chest moves from the *up state* to *down state*, the phase of the dynamic component (gray dot line) changes from H_{d_1} to H_{d_2}. Accordingly, the phase of the received superposition signal (brown line) changes from H_1 to H_2. Let θ_1 and θ_2 be the measured phases of the received signal for the *up state* and the *down state*, respectively. Thus, the phase change of the received signal can be calculated as $\Delta\theta = \theta_2 - \theta_1$.

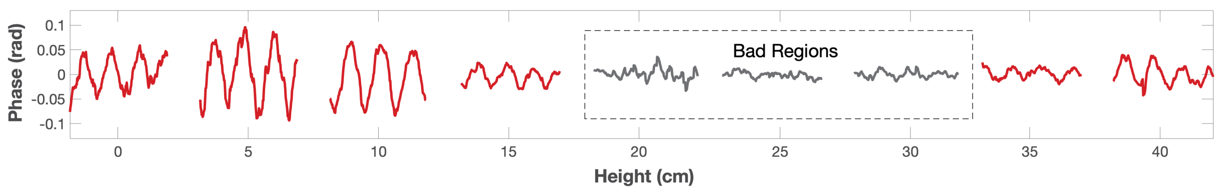


Fig. 3. The phase change pattern with mean removed caused by respiration under the varied height of tag-antenna pair. The tag and the reader antenna (a directional antenna) are placed on the two sides of the bed with a distance of 1.8 m. We let a user lie in the middle of the bed and breathe normally, and vary the height *h* from 0 cm to 40 cm with a step of 5 cm.

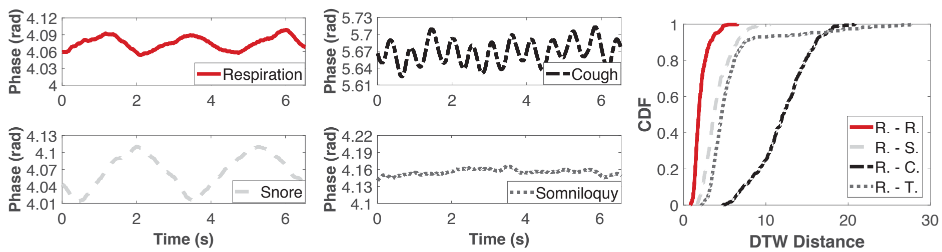
To better understand how the vertical distance h between the tag-antenna pair (LOS path) and the chest affects the received phase change pattern during this up-down movement, we conduct a microbenchmark experiment. Fig. 3 shows the observed phase change pattern under different h. Obviously, the 'good' regions and 'bad' regions are presented alternatively, which is consistent with the results reported in existing Wi-Fi based respiration monitoring works and can be explained by Fresnel zone model [40, 60, 71, 72]. As we can see, a good height is between 0 cm to 10 cm in our scenario, and we pick the height of around 5 cm in our experiment.

4 OVERVIEW

4.1 Basis of TagSleep

Traditionally, snore, cough and somniloquy can be identified by the sound signals. In addition to generating sounds, these activities also cause specific changes in respiration patterns. During the process of these activities, the amount of chest movements and the frequency of respiration change, resulting in a unique signal variation for each activity. For instance, snoring is the abnormal breathing sound [4] and it is often accompanied by a deep inspiratory breath hold [48]. A cough is a sudden and often repetitively occurring [11], and the frequency of chest movement is usually faster compared with normal respiration. The talking procedure is achieved by specific control of respiration [27]. During somniloquy, the chest displacement is significantly reduced, leading to a much weaker periodical variation pattern. Somniloquy can involve complicated dialogues or monologues, complete gibberish or mumbling. In this work, we mainly focus on the somniloquy lasting for a few seconds[2].

To understand how the sound-activity affects the phase change pattern, we conducted a microbenchmark experiment where the tag and the reader antenna are placed on the two sides of the bed. The tag-antenna link is parallel to and on top of the human target's chest. The details of the employed hardware, system implementation and data collection are introduced in Sec. 7.1. We make the following observation:



(a) Phase of resp. and snore. (b) Phase of cough and somniloquy. (c) DTW distance of activities.

Fig. 4. The phase change patterns of different activities are distinguishable. The 'R. - R.' refers to the distance of the respiration itself, which can be considered as a baseline. 'R. - S.'/'R. - C.'/'R. - T.' denote the distance between the respiration and snore/cough/somniloquy (talking), respectively.

Snore, cough and somniloquy affect the user's respiration in different ways. Fig. 4(a) and Fig. 4(b) show the phase sequences from the four different activities within 6 seconds. Note that, to make a fair comparison, the activity samples that we plot are under the same sleep posture (i.e., supine). Compared to normal respiration, snore induces larger phase variations, while cough increases the frequency of the phase variation. The phase variation is smaller for somniloquy. These results are consistent with the symptoms analyzed above. To further quantify the differences of these sleep sound-activities, we adopt the dynamic time warping (DTW) distance which is a common metric for evaluating the similarity/difference between two time series [61, 73]. Normally, a larger DTW distance indicates a lower similarity of two time series. Specifically, we calculate the DTW distance between the phase sequences of normal respiration and each sleep sound-activity, respectively. Fig. 4(c) illustrates the cumulative distribution function (CDF) of the DTW distances with 380 samples for each activity. As we can see, compared to normal respiration, the average DTW distances of the snore, cough and somniloquy are increased by 102.6%, 152.8% and 561.4%, indicating a large difference to normal respiration. Therefore, the phase change pattern can be used as a reliable and robust primitive for snore, cough and somniloquy recognition. That is why TagSleep can employ the respiration as the first-layer information to infer the second-layer sensing.

4.2 System Overview

TagSleep is composed of four modules that operate in a pipelined manner as shown in Fig. 5:

- Tag deployment: An effective layout of multiple tags design, which enables TagSleep to recognize sleep postures and enhance the phase change pattern.
- Pre-processing: Since the raw measured phase readings contain the noise imposed by the hardware and the environment, TagSleep needs to filter out the noise. Then, TagSleep removes periods with body motions and keeps the stable periods for the following sleep posture and activity recognition.

[2]Since the time duration of shouts or screaming is too short, we cannot capture these types of somniloquy by chest movement detection alone. We cannot capture the very short cough either, like throat clearing.

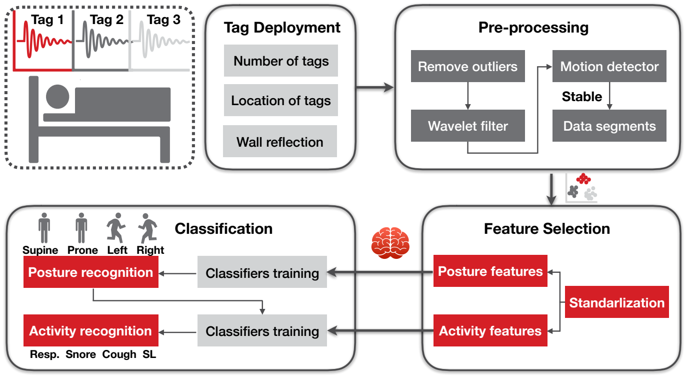


Fig. 5. System architecture of TagSleep.

- Feature selection: A feature extraction method which takes the measured phase sequences of the 3 tags as input and returns a series of effective features that uniquely correlated to the user's sleep sound-activities and postures.
- Classification: A sleep posture and sound-activity recognition subsystem, which employs the above-extracted features to recognize the sleep posture first. With the known posture, TagSleep can further recognize user's sleep sound-activity effectively, including snore, cough and somniloquy (SL).

5 TOPOLOGY FOR MULTIPLE TAGS DEPLOYMENT

In this section, we first analyze the geometric relationship among the bed, user, reader and tags to guide the deployment topology. Then we propose a phase change enhancement method with extra reflection to obtain a 'clearer' pattern for sleep sound-activity recognition.

5.1 Deployment Topology

5.1.1 Safe Distance Between Adjacent Tags. With multiple tags, when two RFID tags are placed close to each other, mutual coupling will change the radiation pattern of the tag antenna, affecting the received phase change pattern [16, 63]. To investigate how the mutual coupling between tags affects the signals, we conducted an experiment where we varied the distance between two tags from 0 cm to 20 cm and ran our system. We observed that, when two tags are very close, the phase change pattern becomes less obvious and the noise level is much higher. As the distance increased, the measured phase change matches the respiration cycles very well and presents a very 'clear' pattern. We find that when the distance between two tags is larger than a threshold, e.g., 15 cm for the tags used in our experiments, the coupling effect is small enough and can be ignored.

5.1.2 Effective Number of Tags and Its Deployment. Intuitively, we can deploy many tags around the bed as long as the distance between each adjacent pair is above the threshold, e.g., 15 cm. To ensure high performance and low cost, we would like to investigate the topology of the deployed tags and minimize the number of tags. To do this, we carefully consider the physical phenomena of the sleep sound-activities together with the geometric relationship among the bed, user, reader and tags.

As shown in Fig. 6, many organs are involved in the respiration, including the chest, abdomen, nose, trachea, lungs, rib cage, muscles and diaphragm [64]. Among these organs, both the chest and abdomen present the periodic movement corresponding to the inhalation and exhalation. Thus, those regions that can sense the chest and abdomen movements are the key regions for placing the tags. From our experiments, we find that using 2 tags deployed in the region #1 and region #2 (see Fig. 6), respectively, is sufficient to cover both the chest and abdomen areas when the user lies on his/her back (supine). Given the user may sleep on his/her side, to detect

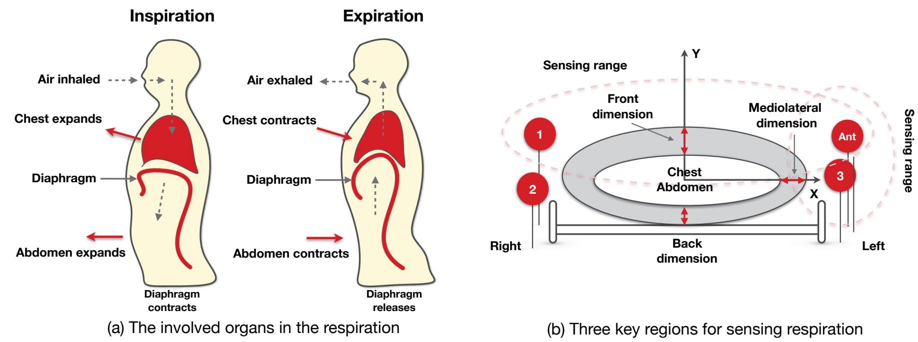


Fig. 6. Illustration of body movement during inspiration/expiration and the effective regions for tag deployment.

the posture and sense respiration under various sleep postures, we need to deploy another tag at the same side as the reader, i.e., region #3. The overall layout of the 3 tags is illustrated in Fig. 7.

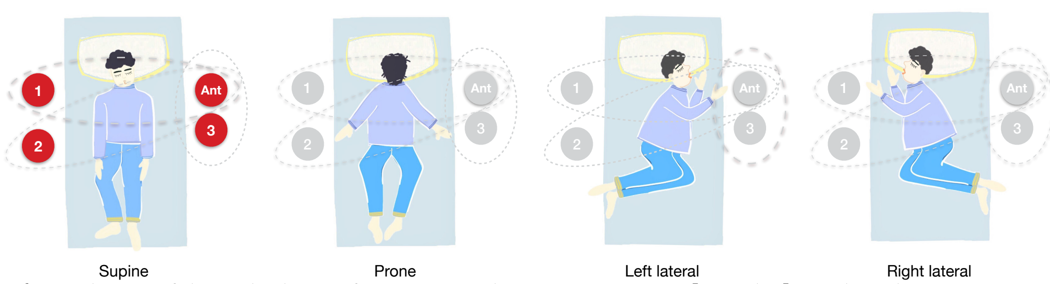


Fig. 7. Effective layout of the multiple tags for sensing under various postures. [@Daily.J] TagSleep does not require a precise geometry for tag placement. Roughly deploying 3 tags in the 3 regions, respectively, suffices our sensing purpose.

5.2 Enhance the Phase Change Pattern

For those users who have relatively smaller chest/abdomen movements due to their smaller lung capacity, it is challenging to obtain a fine-grained 'clear' phase change pattern for respiration-related sensing. We propose a phase change enhancement method to guarantee a 'clear' phase change pattern.

5.2.1 Phase Change Pattern Analysis. We first derive the factors affecting the sensing capability (e.g., a larger $\Delta\theta$ caused by the chest movement means a high sensing capability). Then we explain how to control those factors to enhance the sensing capability.

To analyze the factors affecting $\Delta\theta$, we derive the detailed expression of $\Delta\theta$ as

$$\Delta\theta \approx \frac{|H_{d_2}|\sin\rho_2}{\sqrt{|H_s|^2 + |H_{d_2}|^2 + 2|H_s||H_{d_2}|\cos\rho_2}} - \frac{|H_{d_1}|\sin\rho_1}{\sqrt{|H_s|^2 + |H_{d_1}|^2 + 2|H_s||H_{d_1}|\cos\rho_1}}. \quad (2)$$

Eq. 2 is obtained by placing the received phase expression (i.e., Eq. 1) into the equation $\Delta\theta = \theta_2 - \theta_1$ (introduced in Sec. 3). To further simplify Eq. 2, according to [46], for the respiration-related activity, the amplitude of the dynamic path component can be considered as a constant (i.e., $|H_{d_1}| = |H_{d_2}| = |H_d|$), and $|H_d|$ in the denominator of Eq. 2 can be ignored since it is much smaller than $|H_s|$. Thus, $\Delta\theta$ can be approximately simplified as

$$\Delta\theta \approx \frac{|H_d|}{|H_s|}(\sin\rho_2 - \sin\rho_1). \quad (3)$$

From Eq. 3, the received phase change is determined by $|H_d|$, $|H_s|$ and ρ. Note that the signal reflected from the dynamic path caused by the chest movements is much weaker, i.e., $|H_d| \ll |H_s|$. Thus, $\Delta\theta$ is mainly determined by the Fresnel phase ρ, therefore, we can change ρ to obtain a larger $\Delta\theta$. Recall that ρ is defined by the phase difference between H_s and H_d. To change ρ, we can either change H_s or H_d. In fact, the user will change his/her position or posture during the whole night, so, it is difficult to control the dynamic component (H_d). Instead, we adjust the static component (H_s) to change ρ for obtaining a larger $\Delta\theta$.

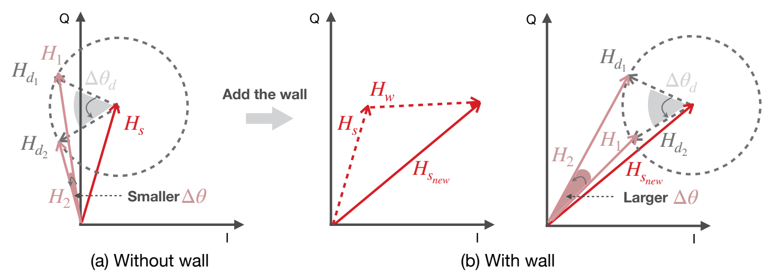


Fig. 8. Phasor representation when extra reflection path from the wall behind the tag.

5.2.2 Phase Change Pattern Enhancement with Multipath. A straightforward way of changing the static component is adding an extra reflection from other static object. TagSleep fully utilizes the natural layout of the bedroom to create extra reflection from the wall/furniture behind the tag. To see it clearly, Fig. 8(a) illustrates an example of the phasor representation (introduced in Sec. 3) for the up-down movement with a small $\Delta\theta$. Then, the extra reflection signal H_w from the wall behind the tag creates a new static component $H_{s_{new}}$ (Fig. 8(b)). By doing this, the Fresnel phase (i.e., both ρ_1 and ρ_2) changes accordingly, so $\Delta\theta$ changes from a small value to a larger one. Therefore, the phase change pattern associated with the chest movement can be enhanced by the extra reflection. It is worth noting that the value of $\Delta\theta$ changes periodically, which indicates that the 'good' locations and 'bad' locations for the tag are presented alternatively (see Fig. 9).

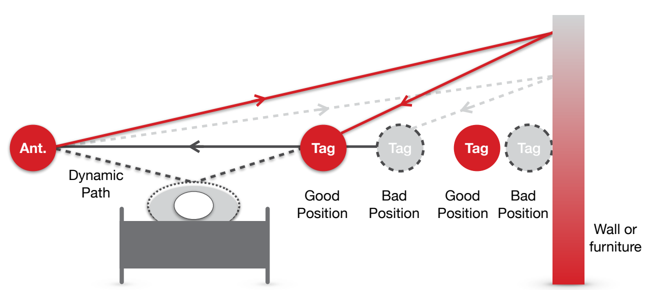


Fig. 9. Extra reflection path from the wall with the 'good' locations and 'bad' locations alternatively.

We further conducted another benchmark experiment to validate the signal enhancement. For simplicity, we fixed the tag-antenna distance as 1.8 m and only changed the distance L between the tag and the wall behind from 2 cm to 130 cm. As shown in Fig. 10, the 'good' and 'bad' locations for the enhancement are alternative, which is consistent with the above analysis. Specifically, when L is between 3 cm to 10 cm or at 100 cm, the variance of the phase change pattern is larger (e.g., around 0.2 radians), indicating a very 'clear' pattern. We pick the tag-wall/furniture distance within the effective range of $[3cm, 10cm]$ in the following experiment. Note that not every user but the users who have relatively smaller chest/abdomen movements need the signal enhancement from the nearby wall/furniture for a better performance.

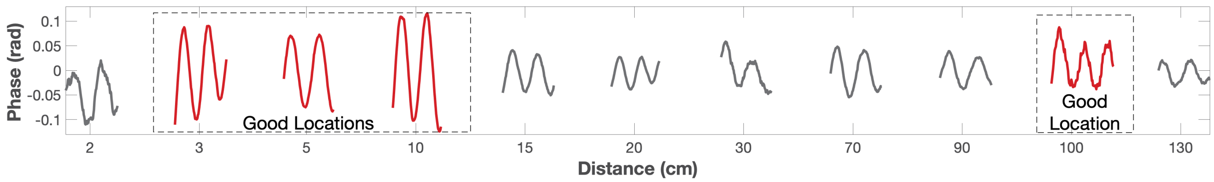


Fig. 10. Phase measurements with mean removed from normal respiration under different tag-wall distances. We attached a tag on a flat paper surface supported by a tripod on one side of the bed, while using another tripod to support the antenna on the other side of the bed. The vertical distance between the user's chest and the LOS path around 5 cm based on the analysis in Sec. 3. During this experiment, the user lied on her back.

6 DATA PROCESSING

In this section, we introduce the data processing in TagSleep, including the phase pre-processing, feature extraction and the sleep posture/activity classification.

6.1 Phase Pre-processing

6.1.1 Remove the Outliers. The raw measured phase readings from the RFID hardware contain the noise imposed by the hardware and the environment. Thus, it is difficult to observe a 'clear' phase change pattern caused by the respiration as it is overwhelmed by the noise. To solve this problem, first, we remove outliers by a simple smooth method [9]. In this method, two adjacent phase values are compared and the later one is dropped if the difference between these two values exceeds a threshold. The rationale behind this scheme is that the adjacent phase values should exhibit the property of continuity and thus we can employ a threshold to remove the large fluctuation (phase outlier). Note that occasionally dropping one sample does not affect the respiration detection since the RFID sampling rate is around 20 to 30 samples per second. To avoid the situation when the first value in the phase sequence is an outlier, we compare the first value with the average value of this phase sequence. If they exceed a threshold, we drop the first value and then the second value will become the new first value. Second, we apply the wavelet filter to remove high-frequency noise in the measured phase sequence. Unlike other low-pass filters, which would smooth the sharp transitions caused by sleep sound-activities as well, the wavelet filter can preserve these sharp transitions well.

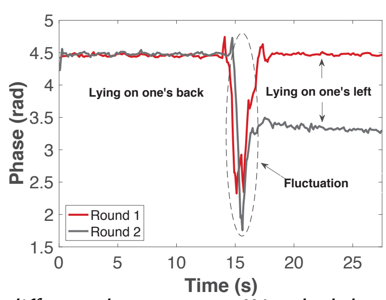


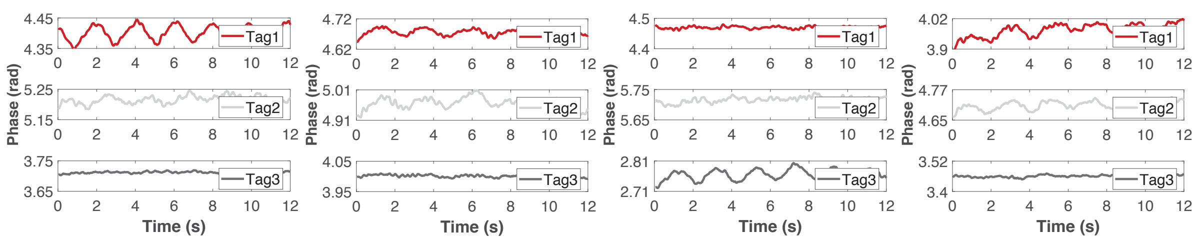
Fig. 11. The phase change of tag #1 under different sleep postures. We asked the user to lie on her back for about 10 seconds, then turn left and keep the left lateral for another 10 seconds, and repeated this process for several rounds. We randomly pick the measured phases at tag #1 obtained in two rounds.

6.1.2 Motion Detector. In TagSleep, we only process the stable periods for sleep posture and sound-activity recognition. The periods containing body movements (e.g., turning over) are discarded. For example, as shown in Fig. 11, a significant fluctuation in the measured phase of tag #1 is observed when the user turns over. Note that such fluctuation can be observed at the other two tags too. Thus, if the measured phases of all 3 tags exceed a threshold, a motion is detected.

6.1.3 Data Segments. TagSleep employs a window as a unit for feature extraction. The time duration of the window should be long enough to be discriminative and short enough to provide high accurate labeling results. From our dataset, a window of 10 seconds is enough to capture snore, cough and somniloquy. However, it is possible that an activity may start at the end of a window and end at the beginning of the next one. To avoid this situation, two adjacent windows are overlapped for 6 seconds, since the short duration activities, such as the cough and somniloquy, usually last for no more than 6 seconds. For each window, TagSleep estimates a sleep posture and the final estimated result for this stable period can be calculated by majority voting based on the result of each window. For sleep sound-activity recognition, TagSleep estimates an activity from every window.

6.2 Feature Extraction

6.2.1 Features of Sleep Posture. From the results observed in Fig. 11, it is challenging to recognize sleep postures[3] directly through the measured phase value. Alternatively, TagSleep employs the clearness of respiration patterns from the 3 tags for sleep posture recognition. The intuition underlying our design is that the chest/abdomen displacement caused by respiration in each dimension is different, resulting in different clearnesses of respiration patterns (see Fig. 12).



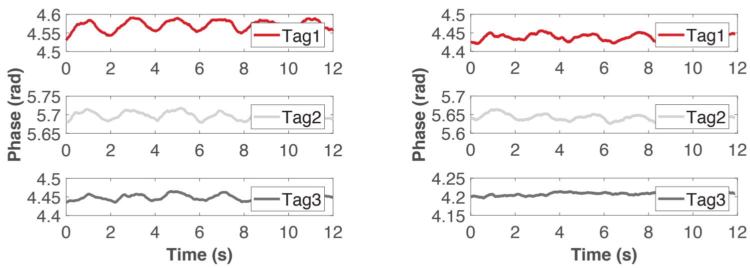
(a) Supine. (b) Prone. (c) Left lateral. (d) Right lateral.

Fig. 12. Phase change patterns under different sleep postures. For supine posture in (a), the phase change patterns from tag #1 and tag #2 are 'clear' and 'less clear', respectively. For prone in (b), tag #1 and tag #2 are 'less clear'. For left lateral in (c), only tag #3 is 'clear'. For right lateral in (d), tag #1 and tag #2 are both 'less clear'.

To represent the clearness, we define the *periodicity level* and *sensitivity level* as *clearness metrics*. Specifically, the *periodicity level* of phase change pattern can be calculated as $P_r = a/RMSE$ [39], where $RMSE$ refers to root-mean-square error and a is the amplitude of the identified sinusoidal wave. Note that a larger P_r means a higher periodicity level. We use the dominant frequency ratio (DFR) together with the variance to describe the sensitivity level. Under the condition that a measured phase sequence has a high periodicity level and a high DFR, generally, the larger its variance is, the higher its sensitivity level is.

However, we cannot solely use the clearness of the observed respiration pattern for reliable and robust sleep posture recognition. First, the respiration patterns under prone and right lateral postures are similar for tag #1 and tag #2 as shown in Fig. 12(b) and Fig. 12(d), respectively. Second, even for the same sleep posture, a minor change of the limbs or locations can lead to a significant variation of the phase change pattern (see Fig. 13). In fact, recognizing sleep postures is more challenging than detecting respiration under various postures. To solve the above intra-class variability, TagSleep extracts a set of effective features that vary significantly among different sleep postures while staying stable within a particular posture. Table 1 summarizes the features used in TagSleep.

[3]For simplicity, in this work, we focus on recognizing four common sleep postures including supine, prone, left lateral and right lateral. For each posture, the body orientation is not limited, instead it can vary in different limb locations.



(a) Right lateral #1. (b) Right lateral #2.

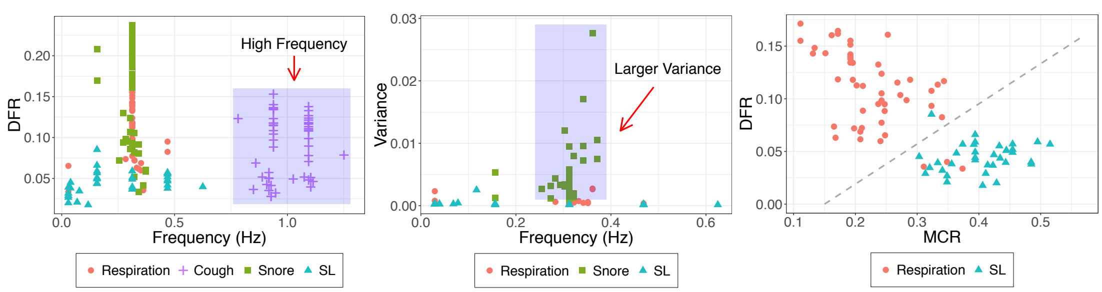
Fig. 13. Phase change patterns of the same sleep posture are different.

Note that we standardize all the features to make the gradient descent procedure in classifiers faster and more stable, especially for classifiers like support vector machines and neural networks [21].

Table 1. Features used in TagSleep.

Time-domain features:
max-min, variance, standard deviation, Mean Crossing Rate (MCR), above α -mean ratio (AMR), kurtosis, periodicity level, max, min, mean
Frequency-domain features:
frequency, Dominant Frequency Ratio (DFR), Left to Right Ratio (LRR), full width at half maximum (FWHM), energy
Nonlinear:
sample entropy

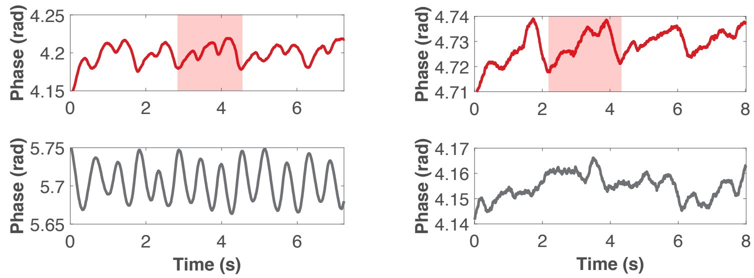
6.2.2 Features of Sleep-activity. The respiration rate is widely used in respiration monitoring systems. However, precise respiration rate calculation cannot ensure the accuracy for sleep sound-activity recognition. The reason is that even if the number of peaks and valleys are identified correctly, their locations and fluctuations in the measured phase sequence may introduce errors in the sleep sound-activity recognition. For an effective feature representation, we study the physical phenomena of different sleep sound-activities and four features including the frequency, DFR, variance, and MCR are extracted as shown in Fig. 14.



(a) Features in frequency domain. (b) Variance. (c) DFR and MCR.

Fig. 14. Four features for distinguishing different sleep sound-activities. In (a), the frequency representing the speed of chest movements can be used to distinguish cough from other activities. Then, in (b), the variance (representing the chest displacement) of snoring is larger than that of the other activities. Finally, in (c), DFR and Mean Crossing Rate (MCR) [34] can be used to distinguish somniloquy from normal respiration.

Although most activities can be recognized using the above four features, some activities may present different patterns even from the same user (see Fig. 15). To solve this problem, recall that the sound-activities and the chest movement are highly related, we therefore use extra features from the sound signal processing field [23, 32, 58] as listed in Table 1.

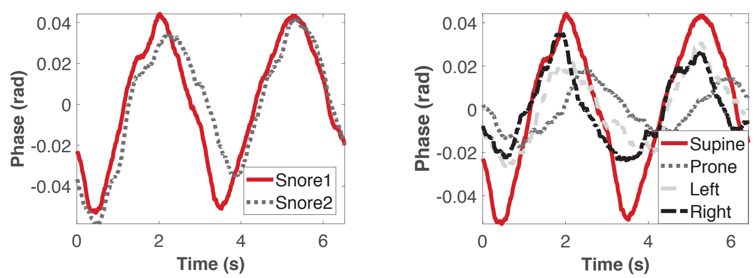


(a) Cough (b) Somniloquy

Fig. 15. Different patterns for the same sleep sound-activity. In (a), the bottom cough (i.e., a strong cough) has a larger frequency (i.e., quicker chest movements), while the top cough (i.e., a light cough) presents both its own frequency and the respiration frequency. The same phenomenon is observed in the somniloquy in (b).

6.3 Classification

Since the same activity may present different patterns across sleep postures. As illustrated in Fig. 16(b), the snore activity varies significantly under four sleep postures, while it keeps relatively stable for a fixed posture (see Fig. 16(a)). The same property is observed from other activities. This implies that we should recognize sleep activity with the knowledge of sleep posture. In fact, even under the person-dependent scenario, it is challenging to find a one-fits-all model that can recognize sound-activities with high accuracy. Therefore, TagSleep recognizes sleep posture and activity, separately. By doing this, we can reduce not only the amount of data needed for training but also improve the classification accuracy on the test data.



(a) Under supine posture. (b) Under different sleep postures.

Fig. 16. The stability and variation of snore activity.

TagSleep uses the extracted features from the 3 tags as inputs to train the classifiers for sleep postures and sound-activities, separately. Specifically, for sleep sound-activity recognition, TagSleep builds up four models according to the four sleep postures. Various classifiers that widely used in the existing RF-based activity recognition [10, 73, 74] can be used in TagSleep including Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Tree, Naive Bayes, Adaboost, Back Propagation Neural Network (BPnet) and Random Forest (RF). The performances of these classifiers are evaluated in Sec. 7. It is worth noting that recent works have shown that deep neural networks are powerful in extracting features [47, 69]. However, they require a huge number of training samples than what we have, while collecting more data is difficult since it will induce prohibitively high overhead and inconvenience to the user.

7 EVALUATION

In this section, we first introduce the implementation of TagSleep. Then we evaluate the performance of TagSleep under various conditions and compare it with the state-of-the-art methods.

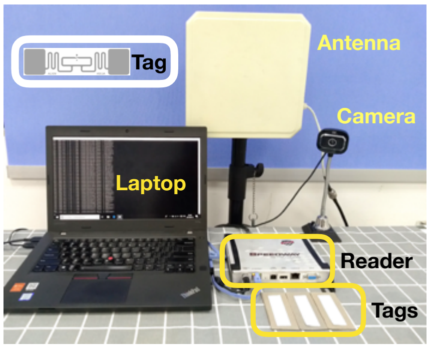


Fig. 17. Hardware used in TagSleep.

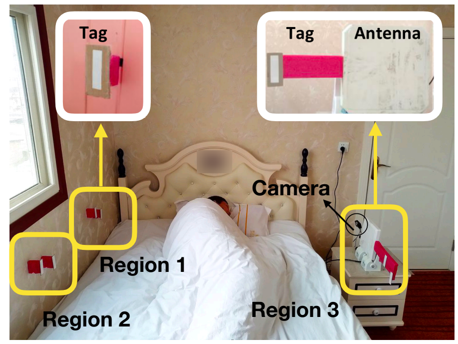


Fig. 18. TagSleep deployment in a bedroom.

7.1 Implementation

Hardware implementation: TagSleep consists of an Impinj R420 RFID reader, three low-cost Alien tags and a directional antenna with 8 dBi gain that powered by the reader as shown in Fig. 17. The reader operates at a fixed frequency of 924.375 MHz. We use the Octane SDK [28] together with the reader to obtain the backscattered responses from the RFID tags at a sampling rate around 20 to 30 samples per second. Note that we use the camera in Fig. 17 to record the video and sound only for obtaining the ground truth and we labeled all the data manually.

Software implementation: We use a laptop with 2.5 GHz CPU and 8 GB memory for data collection and processing. The RFID reader is connected to the laptop through an Ethernet cable and the low-level reader protocol (LLRP) is used for communications. The proposed method is programmed in *C#* and MATLAB R2017b.

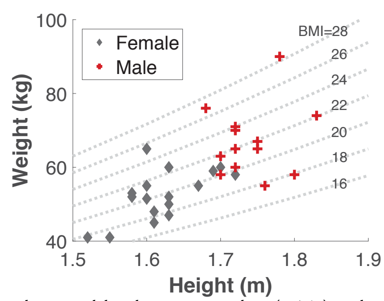


Fig. 19. The heights, weights and body mass index (BMI) values of all the participants.

Environment/system setup and participants: We evaluated TagSleep with 30 users including 13 males and 17 females and collected 2 months of sleep data in total. Their ages are in the range of 18 to 60 years, with the height varying from 152 to 183 cm and weight from 41 kg to 90 kg as shown in Fig. 19, representing a wide range of users. Among all 30 users, there are 9 users presenting snore, cough or somniloquy during sleep. Specifically, all the 9 users snored during sleep, while 5 and 6 users among them coughed and talked, respectively. All experiments with human targets are approved by our IRB (Institutional Review Board).

The experimental settings vary in different bedrooms and different bed sizes. Specifically, we have 8 different bedrooms with the size ranging from $9m^2$ to $40m^2$. The beds also vary in size, including three twin, one full, one queen, two Cal king and one king-size beds. In all settings, the beds are against the wall or at a maximum distance of around 0.6 m to the wall. In this case, we can attach the tags to the nearby wall and place the antenna on the nightstand. The distance [4] between the tag (i.e., the tag in region #1) and the antenna is about 1.8 m, 1.5 m, 1.8 m, 2 m, 2.7 m for the twin, full, queen, Cal king and king-size beds, respectively. Fig. 18 shows one example of TagSleep deployment. All the tags and antenna are of the same height and keep around 5 cm above the user's chest. To avoid the situation when the user blocks the signal at certain locations during natural movements of sleep, we deployed 2 tags within each region. TagSleep then selects one tag from each region. The selected tags have the continuously better measurements for data processing. Note that the 2 tags within each region do not require to be accurately located at specific locations. Considering the layout of a common bedroom, we attach tags in region #1 and region #2 to the wall near the bed. To prevent the signals from being absorbed by the wall, we first attach the tag to the foam with a thickness of about 3 to 10 cm. Then we attach the foam to the wall[5]. For the deployment of tags in region #3, we also use the foam with a length of 30 cm to attach to the antenna. This tag-antenna pair is placed on the nightstand for convenience.

Performance metrics: We employ four widely used metrics in machine learning field including the accuracy, F1 score, precision and recall [65] to evaluate the performance of TagSleep.

Table 2. Performance evaluation of different classifiers for sleep posture recognition.

	Classifier						
Metric (%)	SVM	KNN	Decision Tree	Naive Bayes	Adaboost	BPnet	**Random Forest**
Accuracy	96.16	93.34	97.22	88.42	96.20	98.05	98.94
Precision	97.29	94.57	97.37	89.51	96.64	97.83	98.81
Recall	96.28	93.44	97.26	88.45	96.34	98.13	98.97
F1	96.59	93.72	97.25	87.32	96.29	97.93	98.86

7.2 Overall Performance

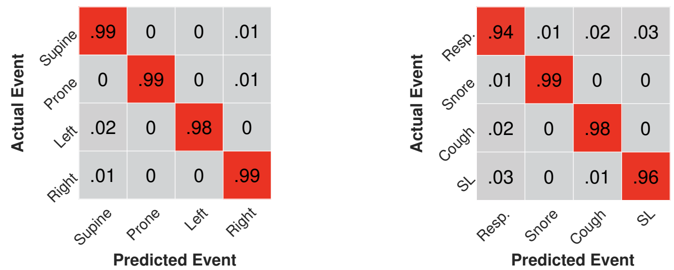
In the overall evaluation, we intend to evaluate the performance of TagSleep with different classifiers. For each user, we randomly divided his/her data into different subsets and applied the commonly used 10-fold cross-validation to evaluate our approach.

7.2.1 Evaluation of Sleep Posture Recognition. The results of the 7 widely used classifiers[6] for sleep posture recognition are shown in Table 2. For a particular classifier, the value of each metric shown in the table is the average value of all 30 users. For each cross validation, we make sure that the training/testing dataset is the same among different classifiers. As we can see, all 7 classifiers can achieve satisfying performance. 'Random Forest' performed slightly better (i.e., above 98%) than other classifiers. This is mainly because 'Random Forest' is an ensemble method which uses multiple learning algorithms (i.e., building a forest of decision tress) to obtain better performance than what could be obtained from any of the constituent learning algorithms alone. Moreover, the performance of other classifiers could be further increased by the fine-tuning of parameters or using more training data. For simplicity, in the following, we adopt 'Random Forest' for sleep posture recognition. Furthermore, to see the results clearly for each sleep posture, Fig. 20(a) illustrates the confusion matrix of sleep posture recognition. Obviously, the accuracy of each posture is higher (i.e., above 98%) and there is no obvious difference for the four postures in terms of recognition capability.

[4] In our experiment, the maximum distance between the tag and the antenna that can sense the respiration is 4 m. Note that for other types of tags and antennas, the distance may be slightly different.

[5] Note that if the wall or furniture is non-reflective to radio signals, such as drywall, we can add a metal sheet behind the foam.

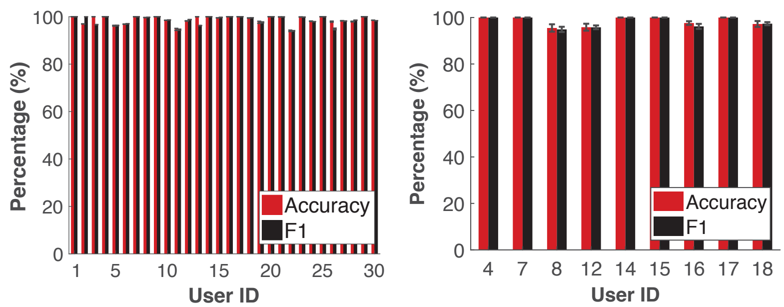
[6] Since this paper do not focus on how to apply advanced machine learning algorithms or combination strategies to achieve the best performance, we directly employed the build-in functions provided by Matlab for 7 widely used classifiers. For each classifier, we only tried several different parameters and chose better configurations based on our dataset, such as Naive Bayes with default parameters, SVM with RBF kernel, KNN with "K=5", Decision tree with CART algorithm, AdaBoost with 100 learning cycles, a two-layer BPnet with training function "traingdx", and Random Forest with 50 decision trees. Note that how to further fine tune these parameters is left for future work.



(a) Sleep posture recognition. (b) Sleep sound-activity recognition.

Fig. 20. Confusion matrix. The diagonal of each of these matrices shows the classification accuracy and the off-diagonal grid points show the confusion error.

In addition, Fig. 21(a) illustrates the sleep posture recognition performance of TagSleep in terms of accuracy and F1 score for each individual user, respectively. As we can see, the accuracy and F1 score of most users exceed 98% with a very small variance shown in the black error bar. The performances of user 2, 5, 6 and 11 are slightly lower (i.e., around 96%) than others. The reason is that these four users have relatively smaller chest/abdomen movements than other users due to their smaller lung capacity, producing smaller phase change patterns.



(a) Sleep posture recognition. (b) Sleep sound-activity recognition.

Fig. 21. Performance of TagSleep for different users. (Random Forest)

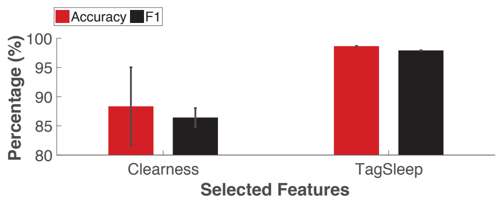
Table 3. Performance evaluation of different classifiers for sleep sound-activity recognition.

	Classifier						
Metric (%)	SVM	KNN	Decision Tree	Naive Bayes	Adaboost	BPnet	**Random Forest**
Accuracy	91.79	82.93	93.47	90.20	86.70	90.69	96.58
Precision	93.21	81.85	92.88	90.46	83.69	90.60	97.27
Recall	91.79	82.93	93.47	90.20	86.70	90.69	96.58
F1	91.44	81.28	92.21	89.05	84.26	89.46	96.32

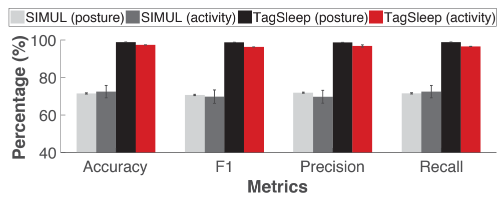
7.2.2 Evaluation of Sleep Sound-activity Recognition. Table 3 compares the 7 widely used classifiers for sleep sound-activity recognition. For a particular classifier, the value of each metric is the average value of all 9 users who present snore, cough or somniloquy during sleep. Obviously, the accuracy of 'Random Forest' is above 96%, which slightly outperforms the other 6 classifiers. Again, the performance of the other classifiers could be increased by the fine-tuning of parameters. In the following evaluation, we adopt 'Random Forest' for sleep sound-activity recognition. It is worth noting that the overall accuracy of recognizing snore, cough and somniloquy is slightly lower compared with sleep posture recognition (i.e., 96% V.S. 98%). The reason is that

TagSleep recognizes sleep sound-activity based on the recognition result of the sleep posture, and the accumulated error from the posture recognition will bring some side-effects to activity recognition.

Fig. 20(b) shows the confusion matrix. It is observed that TagSleep can recognize snore and cough activities very well, while only in some cases, it cannot distinguish somniloquy from normal respiration with a confusion error of 0.03. It is reasonable since light sleep talking could be overwhelmed by the user's respiration. Fig. 21(b) illustrates the accuracy and F1 score of sleep sound-activity recognition for each user, respectively. It is observed that the recognition accuracy is around 98% for majority users, while the accuracy decreases slightly (i.e., about 94%) for user 8, 12, 16 and 18. The reason is that these four users talked frequently during sleep compared with other users, which would bring the misclassification between somniloquy and normal respiration.



(a) Comparison of sleep posture recognition.



(b) Comparison of sleep sound-activity recognition.

Fig. 22. Effectiveness of TagSleep.

7.2.3 Sleep Posture Recognition W and W/o Feature Extraction.

Recall that it is challenging to recognize sleep posture with clearness metrics of respiration patterns alone as discussed in Sec. 6.2.1. As shown in Fig. 22(a), with the clearness features alone, the average accuracy of sleep posture recognition is just around 85%, while the accuracy is improved to 98% when using a set of effective features in TagSleep. The variation of the accuracy when using clearness metrics is around 13%, which indicates an unreliable performance. In contrast, the variation for TagSleep is significantly small. The similar comparison results are observed in the F1 score. These experimental results indicate that TagSleep is able to extract effective features to accurately recognize sleep posture.

7.2.4 Sleep Sound-activity Recognition W and W/o Posture Recognition.

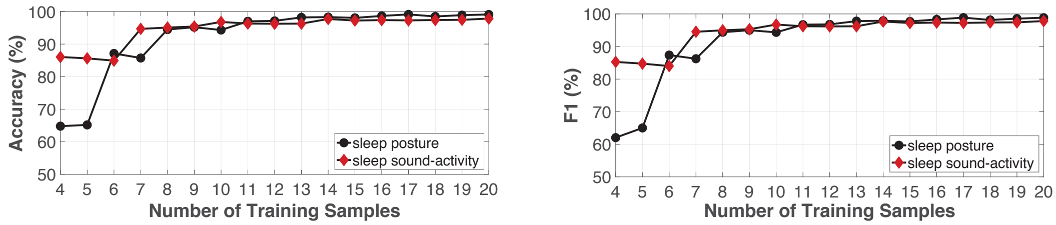
We further evaluate the effectiveness of the idea that recognizing sleep posture is the first step, and sleep sound-activity can be recognized based on the posture information. To do this, we use all the extracted features from the 3 tags as the input, but with different recognition procedure. Specifically, we recognize sleep posture and sound-activity simultaneously by using Random Forest classifier as a baseline. Then we follow the two-step recognition using the same classifier and compare it with the baseline results. Fig. 22(b) illustrates the comparison results and note that 'SIMUL' refers to recognizing posture and activity simultaneously. Obviously, the average accuracy of sleep sound-activities increases from 71% to above 96% when we employ the two-step recognition. These results further emphasize that a universal model is unlikely to deliver good performance and confirm the effectiveness of TagSleep's design.

7.3 Impact of Various Factors

Next, we intend to address the following questions: (i) how many training data are required of TagSleep in person-dependent scenario to achieve high performance; (ii) will the performance decrease in person-independent scenario (i.e., without training data from the testing user).

7.3.1 Impact of Training Set Size.

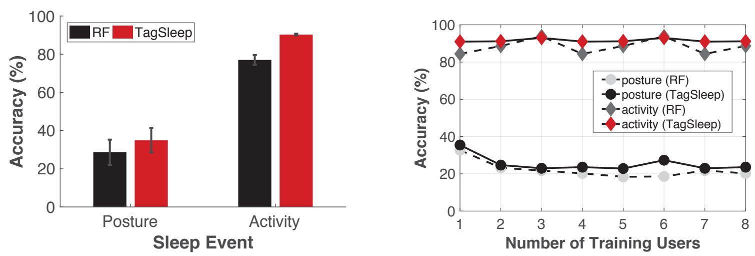
Fig. 23 shows the average performance in terms of the accuracy and F1 score when different numbers of training samples are included in TagSleep. As we can see that the performance of both sleep posture and sleep sound-activity recognition is improved when the number of training samples is increased. More importantly, our system can reach a satisfying performance with a very small number of training samples.



(a) Accuracy of TagSleep.
(b) F1 score of TagSleep.

Fig. 23. Impact of training set size.

For example, TagSleep achieves around 98% and 97% accuracy for sleep posture and sound-activity recognition, respectively, when only 14 training samples are included (see Fig. 23(a)). Note that one training sample refers to one data segment introduced in Sec. 6.1.3. The results indicate that TagSleep has robust performance by extracting a set of effective features.



(a) Accuracy of TagSleep.
(b) Accuracy V.S. Number of users.

Fig. 24. Impact of the untrained user.

7.3.2 Impact of Untrained User. We also evaluate TagSleep using the person-independent classification, where the model is tested on untrained users. Specifically, we pick a given number of randomly selected users for training (i.e., 29 users for sleep posture while 8 users for sleep sound-activity), and the data from the remaining one user are used for testing. This process iterates until all users are tested. To enable the generalization ability of TagSleep for untrained users, we further employ plurality voting from the total 7 classifiers. As shown in Fig. 24(a), the accuracy of sleep sound-activity recognition drops by over 20% when the testing user is not included in the training set (see RF classifier). However, with the help of plurality voting, the drop in accuracy of TagSleep is much smaller (i.e., less than 5% on average) compared with the person-dependent scenario.

Although plurality voting can increase the accuracy of both sleep posture and sound-activity recognition for unseen users, we observe that the average accuracy of posture recognition is still lower compared with that of activity. The reason is that the diversity of sleep postures is higher than that of sound-activities across different people. Specifically, most errors happen for distinguishing supine, left and right lateral postures. Thus, a simple one-time calibration is needed when a new user uses TagSleep. Also, it is worth mentioning that this observation is counterintuitive since TagSleep recognizes sleep sound-activities with the posture information. This is because most activities occurred under the supine posture in our dataset. Thus, although the sleep posture recognition error rate is high, most errors do not affect the accuracy on the activity recognition. Furthermore, we change the number of users in the training set, and randomly choose one user from the remaining users as the testing set. We repeat this process until all users are trained/tested for at least once. Fig. 24(b) illustrates that the performance is quite stable when the number of users increases, which indicates that TagSleep is capable to capture the underlying features of sleep activities across people even with a small number of training users.

7.4 In Comparison to the State-of-the-art Systems

We further compare TagSleep with the state-of-the-art sleep monitoring systems. Since most existing technologies cannot identify both sleep sound-activity and sleep posture simultaneously, we compare the posture and activity recognition performance, separately.

Table 4. Comparison of sleep posture recognition.

Method	Supine	Prone	Left	Right
Wi-Sleep Ext.	91%	-	85.7%	88.9%
SleepGuard	97.3%	96.4%	98.0%	99.5%
TagSheet	89%	91%	100%	100%
TagSleep	99.12%	99.67%	98.39%	98.86%

Table 5. Comparison of sleep sound-activity recognition.

Method	Snore	Cough	Somniloquy
iSleep	96.7%	100%	-
SleepHunter	78%	66%	89%
SleepGuard	96.9%	88.9%	91.3%
TagSleep	98.28%	98.33%	96.08%

For the sleep posture recognition, we summarize the results reported in each system as shown in Table 4. Obviously, TagSleep outperforms commercial Wi-Fi-based system (Wi-Sleep Ext. [39]) with an average 10% improvement of the classification accuracy. Besides, TagSleep can recognize prone posture very well, which Wi-Sleep Ext. cannot support. It is also observed that TagSleep achieves comparable accuracy with wearable sensor-based system (SleepGuard [8]). Furthermore, compared with RFID-based system (TagSheet [35]) which employs 540 tags taped under a bed sheet, our TagSleep with only 3 tags improves the accuracy of supine and prone postures by 10% and 8%, respectively.

For sleep sound-activity recognition, we compare TagSleep against iSleep [23], SleepHunter [22] and Sleep-Guard [8], where the recorded human voice is employed to recognize these activities. Again, Table 5 summarizes the reported results for each system. As we can see, compared with iSleep which uses the built-in microphone of the smartphone, TagSleep can achieve comparable results with RF signals alone. For cough activity, TagSleep outperforms SleepHunter and SleepGuard with the average accuracy improvement of 30% and 10%, respectively. TagSleep can also achieve better performance for the snore and somniloquy activities compared with these two systems. This is because both SleepHunter and SleepGuard employ the built-in microphone of the smartwatch which may be sensitive to the location of the smartwatch and the background noise of the environment. Thus, TagSleep can achieve the comparable and even better results compared with the audio-based approaches and provides a new method to recognize the snore, cough and somniloquy by using RF signals alone.

7.5 Discussions

Several limitations and opportunities for further improvement are discussed as follows:

Tag deployment: In our current experiments, the tags are attached to the nearby wall/furniture. When there is no wall or furniture on the two sides of the bed, we can use the tripods to support the tags. In future work, we will explore different tag deployment strategies which may relax the requirement of having walls or furniture near to the bed.

Multi-user scenario: It is a common scenario when two users (e.g., a couple) sleep on the same bed. It is still challenging for TagSleep to recognize the sleep posture and sleep sound-activity for each user, respectively. We leave this interesting problem as our future work.

Identifying sleep stages and other sleep-activities: TagSleep has the potential for sleep stage identification by combing all the information together (e.g., sleep sound-activities, sleep postures, body movements, sleep duration, etc.), which is left for future work. Furthermore, how to recognize other sleep-activities (e.g., bruxism) and some small movements (e.g., the arm, hand or leg movements) remains an open issue.

Target and environment diversities: To achieve high recognition accuracy, TagSleep relies on the training data collected with a particular user in a particular environment. How to apply advanced machine learning technologies to further improve the accuracy across users and environments is left for future work.

8 CONCLUSION

In this paper, we introduce a new method to recognize sleep sound-activities including snore, cough and somniloquy by using RF signals alone, without requiring the user to wear any sensors or a microphone to record the sound. For the first time, we propose the concept of two-layer sensing. We employ the respiration as the basic first-layer sensing information to further obtain the second-layer information such as sleep sound-activities and sleep postures. Comprehensive real-world experiments show that TagSleep can achieve high accuracy for both sleep sound-activity and sleep posture recognition. We believe this two-layer sensing concept can be applied to enable a large range of new sensing applications.

ACKNOWLEDGMENTS

This research is supported by the National Natural Science Foundation of China (61602382), Natural Sciences and Engineering Research Council of Canada (NSERC), Canada Foundation for Innovation (CFI), BC Knowledge Development Fund (BCKDF), Shaanxi Science and Technology Project (2018SF-369), Shaanxi International Cooperation (2019KWZ-05), and the Science and Technology Innovation Team Supported Project of Shaanxi Province (2018TD-O26).

REFERENCES

[1] Heba Abdelnasser, Khaled A Harras, and Moustafa Youssef. 2015. UbiBreathe: A ubiquitous non-invasive WiFi-based breathing estimator. In *MobiHoc 2015*. ACM, 277–286.

[2] Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2015. Smart homes that monitor breathing and heart rate. In *CHI 2015*. ACM, 837–846.

[3] Adriano M Alencar, Diego Greatti Vaz da Silva, Carolina Beatriz Oliveira, André P Vieira, Henrique T Moriya, and Geraldo Lorenzi-Filho. 2013. Dynamics of snoring sounds and its connection with obstructive sleep apnea. *Physica A: Statistical Mechanics and its Applications* 392, 1 (2013), 271–277.

[4] Ali Azarbarzin and Zahra MK Moussavi. 2011. Automatic and unsupervised snore sound extraction from respiratory sound signals. *IEEE Transactions on Biomedical Engineering* 58, 5 (2011), 1156–1162.

[5] Marek Bartula, Timo Tigges, and Jens Muehlsteff. 2013. Camera-based system for contactless monitoring of respiration. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2672–2675.

[6] Beddit. 2019. *Beddit 3 sleep monitor*. Retrieved May 10, 2019 from https://www.beddit.com

[7] Daniel J Buysse, Charles F Reynolds III, Timothy H Monk, Susan R Berman, and David J Kupfer. 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research* 28, 2 (1989), 193–213.

[8] Liqiong Chang, Jiaqi Lu, Ju Wang, Xiaojiang Chen, Dingyi Fang, Zhanyong Tang, Petteri Nurmi, and Zheng Wang. 2018. SleepGuard: capturing rich sleep information using smartwatch sensing data. *Ubicomp* 2, 3 (2018), 98.

[9] Liqiong Chang, Jie Xiong, Ju Wang, Xiaojiang Chen, Yu Wang, Zhanyong Tang, and Dingyi Fang. 2018. RF-Copybook: A Millimeter Level Calligraphy Copybook based on commodity RFID. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 128.

[10] X. Chen, L. Chen, C. Feng, D. Fang, J. Xiong, and Z. Wang. 2019. Sensing Our World Using Wireless Signals. *IEEE Internet Computing* 23, 3 (2019), 38–45.

[11] Kian Fan Chung and Ian D Pavord. 2008. Prevalence, pathogenesis, and causes of chronic cough. *The Lancet* 371, 9621 (2008), 1364–1374.

[12] Jana R Cooke and Sonia Ancoli-Israel. 2011. Normal and abnormal sleep in the elderly. In *Handbook of clinical neurology*. Vol. 98. Elsevier, 653–665.

[13] Sleep Cycle. 2019. *Sleep Cycle alarm clock*. Retrieved May 10, 2019 from https://www.sleepcycle.com

[14] Gustavo Desouzart, Ernesto Filgueiras, and Rui Matos. 2015. Relationship between postural reeducation technique during sleep and relaxation technique in sleep quality. *Procedia Manufacturing* 3 (2015), 6093–6100.

[15] Peter A Deutsch, Michael S Simmons, and Jeanne M Wallace. 2006. Cost-effectiveness of split-night polysomnography and home studies in the evaluation of obstructive sleep apnea syndrome. *Journal of Clinical Sleep Medicine* 2, 02 (2006), 145–153.

[16] Han Ding, Jinsong Han, Chen Qian, Fu Xiao, Ge Wang, Nan Yang, Wei Xi, and Jian Xiao. 2018. Trio: utilizing tag interference for refined localization of passive RFID. In *Infocom 2018*. IEEE, 828–836.

[17] WD Duckitt, SK Tuomi, and TR Niesler. 2006. Automatic detection, segmentation and assessment of snoring from ambient acoustic data. *Physiological measurement* 27, 10 (2006), 1047.

[18] Fitbit. 2019. *Fitbit*. Retrieved May 10, 2019 from https://www.fitbit.com

[19] National Sleep Foundation. 2018. *Sleep Talking*. https://sleepfoundation.org/sleep-disorders-problems/sleep-talking

[20] Hulya Gokalp and Malcolm Clarke. 2013. Monitoring activities of daily living of the elderly and the potential for its use in telecare and telehealth: a review. *TELEMEDICINE and e-HEALTH* 19, 12 (2013), 910–923.

[21] Joel Grus. 2015. *Data science from scratch: first principles with python*. " O'Reilly Media, Inc.".

[22] Weixi Gu, Longfei Shangguan, Zheng Yang, and Yunhao Liu. 2016. Sleep hunter: Towards fine grained sleep stage tracking with smartphones. *IEEE Transactions on Mobile Computing* 15, 6 (2016), 1514–1527.

[23] Tian Hao, Guoliang Xing, and Gang Zhou. 2013. iSleep: unobtrusive sleep quality monitoring using smartphones. In *Sensys 2013*. ACM.

[24] Enamul Hoque, Robert F Dickerson, and John A Stankovic. 2010. Monitoring body positions and movements during sleep using wisps. In *Wireless Health 2010*. ACM, 44–53.

[25] Chen-Yu Hsu, Aayush Ahuja, Shichao Yue, Rumen Hristov, Zachary Kabelac, and Dina Katabi. 2017. Zero-Effort In-Home Sleep and Insomnia Monitoring using Radio Signals. *Ubicomp* 1, 3 (2017), 59.

[26] Xiaoxuan Hu, Kagome Naya, Peng Li, Toshiaki Miyazaki, and Kun Wang. 2017. Non-invasive sleep monitoring based on RFID. In *IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*. IEEE.

[27] Chenyu Huang, Huangxun Chen, Lin Yang, and Qian Zhang. 2018. BreathLive: Liveness Detection for Heart Sound Authentication with Deep Breathing. *Ubicomp* 2, 1 (2018), 12.

[28] IMPINJ. 2019. *Octane sdk for impinj*. Retrieved May 10, 2019 from https://support.impinj.com/hc/en-us/articles/202755268-Octane-SDK

[29] Jawbone. 2018. *Jawbone*. Retrieved Jul. 31, 2018 from https://jawbone.com/

[30] Anastasi Kosmadopoulos, Charli Sargent, David Darwent, Xuan Zhou, and Gregory D Roach. 2014. Alternatives to polysomnography (PSG): a validation of wrist actigraphy and a partial-PSG system. *Behavior research methods* 46, 4 (2014), 1032–1041.

[31] Henry F Krous, J Bruce Beckwith, Roger W Byard, Torleiv O Rognum, Thomas Bajanowski, Tracey Corey, Ernest Cutz, Randy Hanzlick, Thomas G Keens, and Edwin A Mitchell. 2004. Sudden infant death syndrome and unclassified sudden infant deaths: a definitional and diagnostic approach. *Pediatrics* 114, 1 (2004), 234–238.

[32] Asaf Levartovsky, Eliran Dafna, Yaniv Zigel, and Ariel Tarasiuk. 2016. Breathing and snoring sound characteristics during sleep in adults. *Journal of Clinical Sleep Medicine* 12, 03 (2016), 375–384.

[33] Wen-Hung Liao and Chien-Ming Yang. 2008. Video-based activity and movement pattern analysis in overnight sleep studies. In *2008 19th International Conference on Pattern Recognition*. IEEE, 1–4.

[34] Feng Lin, Yan Zhuang, Chen Song, Aosen Wang, Yiran Li, Changzhan Gu, Changzhi Li, and Wenyao Xu. 2017. SleepSense: A noncontact and cost-effective sleep monitoring system. *IEEE transactions on biomedical circuits and systems* 11, 1 (2017), 189–202.

[35] Jia Liu, Xingyu Chen, Shigang Chen, Xiulong Liu, Yanyan Wang, and Lijun Chen. 2019. TagSheet: Sleeping Posture Recognition with an Unobtrusive Passive Tag Matrix. In *Infocom 2019*. IEEE.

[36] Jian Liu, Yan Wang, Yingying Chen, Jie Yang, Xu Chen, and Jerry Cheng. 2015. Tracking vital signs during sleep leveraging off-the-shelf wifi. In *MobiHoc 2015*. ACM, 267–276.

[37] Jason J Liu, Wenyao Xu, Ming-Chun Huang, Nabil Alshurafa, Majid Sarrafzadeh, Nitin Raut, and Behrooz Yadegar. 2013. A dense pressure sensitive bedsheet design for unobtrusive sleep posture monitoring. In *PerCom 2013*. IEEE, 207–215.

[38] Xuefeng Liu, Jiannong Cao, Shaojie Tang, and Jiaqi Wen. 2014. Wi-Sleep: Contactless sleep monitoring via WiFi signals. In *Real-Time Systems Symposium (RTSS), 2014 IEEE*. IEEE, 346–355.

[39] Xuefeng Liu, Jiannong Cao, Shaojie Tang, Jiaqi Wen, and Peng Guo. 2016. Contactless respiration monitoring via off-the-shelf WiFi devices. *IEEE Transactions on Mobile Computing* 15, 10 (2016), 2466–2479.

[40] Junyi Ma, Yuxiang Wang, Hao Wang, Yasha Wang, and Daqing Zhang. 2016. When can we detect human respiration with commodity wifi devices?. In *Ubicomp 2016*. ACM, 325–328.

[41] Raul Marin, Tamara Cyhan, and Wendy Miklos. 2006. Sleep disturbance in patients with chronic low back pain. *American Journal of Physical Medicine & Rehabilitation* 85, 5 (2006), 430–435.

[42] Juan F Masa, Jaime Corral, Ricardo Pereira, Joaquin Duran-Cantolla, Marta Cabello, Luis Hernández-Blasco, Carmen Monasterio, Alberto Alonso, Eusebi Chiner, Manuela Rubio, et al. 2011. Effectiveness of home respiratory polygraphy for the diagnosis of sleep apnoea and hypopnoea syndrome. *Thorax* 66, 7 (2011), 567–573.

[43] Jodi A Mindell, Rae Ann Cook, and Janeta Nikolovski. 2015. Sleep patterns and sleep disturbances across pregnancy. *Sleep medicine* 16, 4 (2015), 483–488.

[44] Advanced Brain Monitoring. 2019. *Sleep Profiler*. Retrieved May 10, 2019 from http://www.advancedbrainmonitoring.com/sleep-profiler/

[45] Rajalakshmi Nandakumar, Shyamnath Gollakota, and Nathaniel Watson. 2015. Contactless sleep apnea detection on smartphones. In *MobiSys 2015*. ACM, 45–57.

[46] Kai Niu, Fusang Zhang, Jie Xiong, Xiang Li, Enze Yi, and Daqing Zhang. 2018. Boosting fine-grained activity sensing by embracing wireless multipath effects. In *CoNext 2018*. ACM, 139–151.

[47] Vardan Papyan, Yaniv Romano, Jeremias Sulam, and Michael Elad. 2018. Theoretical foundations of deep learning via sparse representations: A multilayer sparse model and its connection to convolutional neural networks. *IEEE Signal Processing Magazine* 35, 4 (2018), 72–89.

[48] Anders N Pedersen, Stine Korreman, Håkan Nyström, and Lena Specht. 2004. Breathing adapted radiotherapy of breast cancer: reduction of cardiac and pulmonary doses using voluntary inspiration breath-hold. *Radiotherapy and oncology* 72, 1 (2004), 53–60.
[49] Nuno Pombo and Nuno M Garcia. 2016. ubiSleep: An ubiquitous sensor system for sleep monitoring. In *2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. IEEE, 1–4.
[50] Tauhidur Rahman, Alexander T Adams, Ruth Vinisha Ravichandran, Mi Zhang, Shwetak N Patel, Julie A Kientz, and Tanzeem Choudhury. 2015. Dopplesleep: A contactless unobtrusive sleep sensing system using short-range doppler radar. In *UbiComp 2015*. ACM, 39–50.
[51] Tauhidur Rahman, Alexander Travis Adams, Mi Zhang, Erin Cherry, Bobby Zhou, Huaishu Peng, and Tanzeem Choudhury. 2014. BodyBeat: a mobile system for sensing non-speech body sounds. In *MobiSys 2014*, Vol. 14. 2–13.
[52] Ruth Ravichandran, Elliot Saba, Ke-Yu Chen, Mayank Goel, Sidhant Gupta, and Shwetak N Patel. 2015. WiBreathe: Estimating respiration rate using wireless signals in natural settings in the home. In *PerCom 2015*. IEEE, 131–139.
[53] O Resta, MP Foschino-Barbaro, G Legari, S Talamo, P Bonfitto, A Palumbo, A Minenna, R Giorgino, and G De Pergola. 2001. Sleep-related breathing disorders, loud snoring and excessive daytime sleepiness in obese subjects. *International journal of obesity* 25, 5 (2001), 669.
[54] Harvard Medical School. 2014. *Why snoring could be a sign your heart is at risk*. Retrieved August, 2014 from https://www.health.harvard.edu/heart-health/why-snoring-could-be-a-sign-your-heart-is-at-risk
[55] Shuyu Shi, Yaxiong Xie, Mo Li, Alex X Liu, and Jun Zhao. 2019. Synthesizing Wider WiFi Bandwidth for Respiration Rate Monitoring in Dynamic Environments. In *Infocom 2019*. IEEE.
[56] Jae Hyuk Shin, Young Joon Chee, Do-Un Jeong, and Kwang Suk Park. 2010. Nonconstrained sleep monitoring system and algorithms using air-mattress with balancing tube method. *IEEE transactions on information technology in biomedicine* 14, 1 (2010), 147–156.
[57] Sleepbot. 2018. *Sleepbot*. Retrieved Jul. 31, 2018 from https://mysleepbot.com
[58] Xiao Sun, Zongqing Lu, Wenjie Hu, and Guohong Cao. 2015. SymDetector: detecting sound-related respiratory symptoms using smartphones. In *Ubicomp 2015*. ACM, 97–108.
[59] Xiao Sun, Li Qiu, Yibo Wu, Yeming Tang, and Guohong Cao. 2017. Sleepmonitor: Monitoring respiratory rate and body position during sleep using smartwatch. *Ubicomp* 1, 3 (2017), 104.
[60] Hao Wang, Daqing Zhang, Junyi Ma, Yasha Wang, Yuxiang Wang, Dan Wu, Tao Gu, and Bing Xie. 2016. Human respiration detection with commodity wifi devices: do user location and body orientation matter?. In *Ubicomp 2016*. ACM, 25–36.
[61] Jue Wang and Dina Katabi. 2013. Dude, where's my card?: RFID positioning that works with multipath and non-line of sight. In *SigComm 2013*, Vol. 43. ACM, 51–62.
[62] Xuyu Wang, Chao Yang, and Shiwen Mao. 2017. PhaseBeat: Exploiting CSI phase data for vital sign monitoring with commodity WiFi devices. In *ICDCS 2017*. IEEE, 1230–1239.
[63] Teng Wei and Xinyu Zhang. 2016. Gyro in the air: tracking 3D orientation of batteryless internet-of-things. In *MobiCom 2016*. ACM, 55–68.
[64] Wikipedia. 2019. *Breathing- Wikipedia*. Retrieved Feb. 24, 2019 from https://en.wikipedia.org/wiki/Breathing
[65] Wikipedia. 2019. *Precision and recall*. Retrieved March 16, 2019 from https://en.wikipedia.org/wiki/Precision_and_recall
[66] Yong Ju Yang, Jae Woo Shin, Seung Jin Jang, Hyun Sook Lee, and Young Ro Yoon. 2007. Research of body movement during sleep with an infrared triangulation distance sensor, wavelets and neuro-fuzzy reasoning. In *World Congress on Medical Physics and Biomedical Engineering 2006*. Springer, 760–763.
[67] Zhicheng Yang, Parth H Pathak, Yunze Zeng, Xixi Liran, and Prasant Mohapatra. 2016. Monitoring vital signs using millimeter wave. In *MobiHoc 2016*. ACM, 211–220.
[68] Zhicheng Yang, Parth H Pathak, Yunze Zeng, Xixi Liran, and Prasant Mohapatra. 2017. Vital sign and sleep monitoring using millimeter wave. *ACM Transactions on Sensor Networks (TOSN)* 13, 2 (2017), 14.
[69] Shuochao Yao, Shaohan Hu, Yiran Zhao, Aston Zhang, and Tarek Abdelzaher. 2017. Deepsense: A unified deep learning framework for time-series mobile sensing data processing. In *WWW 2017*. International World Wide Web Conferences Steering Committee, 351–360.
[70] Shichao Yue, Hao He, Hao Wang, Hariharan Rahul, and Dina Katabi. 2018. Extracting Multi-Person Respiration from Entangled RF Signals. *Ubicomp* 2, 2 (2018), 86.
[71] Youwei Zeng, Dan Wu, Ruiyang Gao, Tao Gu, and Daqing Zhang. 2018. FullBreathe: Full Human Respiration Detection Exploiting Complementarity of CSI Phase and Amplitude of WiFi Signals. *Ubicomp* 2, 3 (2018), 148.
[72] Fusang Zhang, Daqing Zhang, Jie Xiong, Hao Wang, Kai Niu, Beihong Jin, and Yuxiang Wang. 2018. From Fresnel Diffraction Model to Fine-grained Human Respiration Sensing with Commodity Wi-Fi Devices. *Ubicomp* 2, 1 (2018), 53.
[73] Jie Zhang, Zhanyong Tang, Meng Li, Dingyi Fang, Petteri Tapio Nurmi, and Zheng Wang. 2018. CrossSense: towards cross-site and large-scale WiFi sensing. In *MobiCom 2018*. ACM.
[74] Mingmin Zhao, Fadel Adib, and Dina Katabi. 2016. Emotion recognition using wireless signals. In *MobiCom 2016*. ACM, 95–108.
[75] Mingmin Zhao, Shichao Yue, Dina Katabi, Tommi S Jaakkola, and Matt T Bianchi. 2017. Learning sleep stages from radio signals: a conditional adversarial architecture. In *International Conference on Machine Learning*. 4100–4109.