CoughLoc: Location-Aware Indoor Acoustic Sensing for Non-Intrusive Cough Detection

[1]Zheng Sun, [1]Aveek Purohit, [2]Kathleen Yang, [1]Neha Pattan, [1]Dan Siewiorek, [1]Asim Smailagic, [1]Ian Lane, [1]Pei Zhang
[1]Carnegie Mellon University, PA, USA [2]Samsung Communication Lab, San Jose, CA, USA
{zhengs, apurohit, peizhang}@cmu.edu

ABSTRACT

Pervasive medical monitoring has become an ideal alternative to nursing care for elderly people and patients in hospitals. Existing systems using single body-worn sensors are often intrusive and less reliable. By contrast, ubiquitous acoustic sensing techniques can support non-intrusive and robust medical monitoring. In this paper, we describe CoughLoc, a ubiquitous acoustic sensing system for continuous cough detection using a wireless sensor network. We show how knowledge of sound source locations can be leveraged to improve the detection accuracy of sound events caused by mobile users. Experiments in indoor environments show our system achieves over 90% cough detection performance under quiet backgrounds, and 1.6 times higher performance compared to a baseline approach with no location information.

1. INTRODUCTION

With recent developments in ubiquitous system technologies, automatic medical monitoring of patients via multimodal sensors is now a feasible alternative to supervision by healthcare professionals [1]. One such approach is acoustic monitoring, which enables the most critical patient condition information to be available to caregivers in real time. This includes the frequency and intensity of coughs, sneezes, nose blowing, throat clearing and other medical symptoms, in addition to the monitoring of specific conditions, such as sleep apnea [2]. Ambient acoustic monitoring can be performed remotely without a significant invasion of privacy and rapid medical response can be provided when needed.

Cough is the most common symptom for respiratory disorders, including chronic lung disease, pneumonia, tuberculosis and influenza [3]. Continuous monitoring of objective cough frequency and severity can greatly assist physicians in early diagnosis of patient illnesses and the assessment of treatment efficiency [2].

Currently, the assessment of cough mainly relies on cough severity scores, through which patients subjectively describe their perceived symptom. However, these scores are loosely related to objective cough counts, and often affected by mood, vigilance and patient expectations [2]. The state-of-the-art automatic cough monitoring systems are ambulatory cough monitors [3, 4]. These systems use microphones placed either on patients' clothing or directly against their body to monitor mobile users. Recordings are made over a period of time (e.g. a day), after which analysis is performed using statistical machine learning techniques. The main drawback of these systems is their intrinsic intrusiveness and limited reliability. The use of body-worn devices are usually not comfortable for some patients. Furthermore, if the devices were damaged or not worn by the patients, the systems would not function.

Given these issues in the automatic cough monitoring, this paper presents CoughLoc, a distributed acoustic sensing system for non-intrusive cough monitoring. Compared to previous work [3, 4], CoughLoc employs distant microphones to preserve non-intrusiveness and to provide convenience to patients. Using multiple distributed sensor nodes and microphones in wards or rooms, CoughLoc also achieves robustness against single node failure and is capable of monitoring coughing patients that move inside a building.

In the design of such a system, however, several research challenges remain: 1) Existence of various indoor noises and echos, such as human speech, barking animals, and TV sounds, degrades audio quality. Ambient monitoring drastically increases the influence of the noises and echos compared to head-mounted microphones. 2) Resource-constrained hardware limits processing capability and bandwidth. 3) Non-stop, high-frequency audio acquisition commonly generates large amount of raw audio data per second. Given the bandwidth and CPU limitations, balancing tradeoffs between transmitting and processing the data is non-trivial.

To overcome inherent noises in ambient monitoring systems, we design a collaborative sensing system that uses sound source location information to improve detection performance of sound events. In addition, we address the limited resource problem by using an inter-node task partitioning strategy. This strategy uses the location information to dynamically assign computationally intensive tasks to a subset of sensor nodes, thus reduces computation and data transmission on other nodes.

We implemented the system and evaluated it in indoor settings. Experimental results show that our system achieves over 90% accuracy under quiet environments, which are competitive with previous work using body-worn sensors. Under noisy environments, the system achieves over 80% average accuracy under various noise backgrounds, which is 1.6 times higher than a baseline approach with no location information.

The contributions of this paper are as follows:

1. The design and implementation of a collaborative acoustic sensing system that monitors mobile coughing patients. The system leverages information of source locations to improve sound event detection performance and is robust against ambient noisy environments.

2. An efficient task partitioning strategy that dynamically balances complex acoustic processing loads on distributed sensors based on different sound source locations over time.

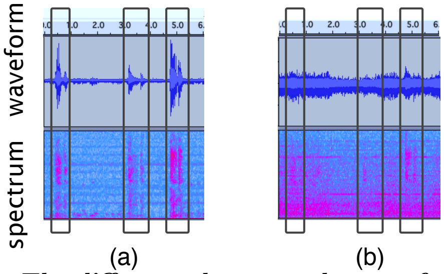


Figure 1: The difference between the waveforms and spectrums of three coughs (indicated by the rectangles on the graph) recorded when (a) close to the cough source and (b) close to a running vacuum cleaner.

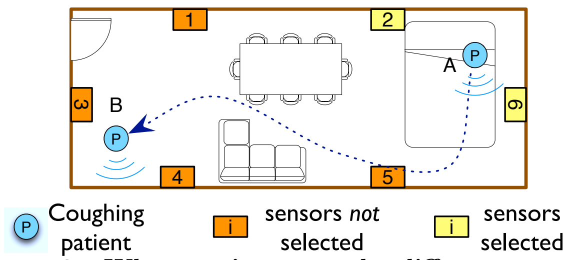


Figure 2: When patients cough, different sensors would be selected to compute audio features based on their proximity to the cough locations.

3. The evaluation of the CoughLoc system for cough monitoring through indoor experiments.

The rest of the paper is organized as follows. Section 2 discusses the benefits of sound source location information for indoor acoustic sensing, and how CoughLoc leverages this information to improve detection accuracy. Section 3 describes the location-based task partitioning strategy. Section 4 presents evaluation results. Section 5 relates our work to other projects. Finally, Section 6 concludes the work and summarizes our contributions.

2. LOCATION-AWARE SOUND EVENT DETECTION

Indoor acoustic sensing faces challenges in overcoming low audio quality, especially when sound sources are far away from microphones. In a ward, for example, when a patient moves and coughs at different locations, sensor nodes will experience significantly different signal-to-noise ratios (SNRs) due to their different distances to the coughing patient and to ambient noise sources.

To show this effect, Figure 1 compares waveforms and spectrums of three coughs recorded with different distances from the same cough source. In this example, a vacuum cleaner is also placed 5m away from the cough source. In Figure 1(a), the microphone is 1m from the cough source and 4m from the vacuum cleaner, while in Figure 1(b) the microphone is 4m from the cough source and 1m from the vacuum cleaner. Due to higher SNR values, the waveforms and spectrums in Figure 1(a) are easy to distinguish, whereas those in Figure 1(b) have been highly contaminated by noise from the vacuum cleaner.

Unlike radio signals, however, estimation of local SNR values for audio signals is computationally demanding, since both the interesting sounds (such as cough, sneeze, or human speech) and non-interesting sounds (such as background ambient noise) are unknown. Thus, SNR estimation usually requires sensor nodes to either conduct local sound classification or to send audio samples back to a central server for a network-level data fusion [5, 6]. These approaches are highly demanding for the resource-limited systems.

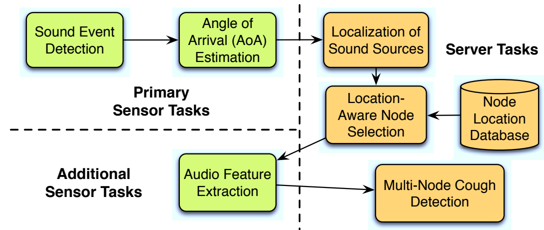


Figure 3: Tasks of the CoughLoc system. Brown: tasks on the CoughLoc Server. Green: tasks on CoughLoc sensor nodes.

However, since collaborative acoustic sensing systems can leverage multiple distributed sensors, these systems can follow the physical locations of sound events and apply a *location-aware data filtering*, such that *only* audio likely to have high SNR will be used for classification. Since SNR of the audio signal generally degrades as the distance to the sound source increases, this filtering can be achieved by dynamically choosing sensor nodes with the shortest distance to the sound source to extract audio features. Consequently, compared to traditional SNR estimation, sensor selection can be done with small computational overhead.

Figure 2 illustrates such a process. When a patient coughs at location A, the system will require Sensor 2 and 6 to extract audio features, due to their close proximity to the patient; likewise, when a patient moves and coughs at location B, Sensor 3 and 4 will be selected. Because of the mobility of coughing patients, different spatially distributed nodes can be selected separately, resulting in reduced overall computation time and bandwidth requirements.

3. SYSTEM DESCRIPTION

CoughLoc consists of a set of distributed acoustic *sensor nodes* and a *server*. Each sensor node has a microprocessor, a radio for communication, and a pair of miniature microphones that continuously capture ambient sound. The sensor nodes are mounted on walls to reduce disturbance to patients and obtain good line-of-sight. The server is an off-the-shelf desktop computer. When a sound event occurs, the system first uses directional information from multiple nodes to localize the sound source, then selects a subset of the nodes to extract audio features, which are transmitted to and gathered at the server for sound classification.

3.1 Task Partitioning Strategy

In order to perform computationally-intensive sensing and classification tasks on the resource-limited CoughLoc nodes, we designed a task partitioning strategy to dynamically assign tasks to different sensor nodes. Figure 3 shows the flow of the partitioned tasks:

- ***Sound Event Detection*** and ***Angle-of-Arrival (AoA) Estimation.*** Each node detects a sound event from ambient noise, and estimates the direction of the incoming acoustic signal by computing angles of signal arrivals.
- ***Localization*** and ***Location-Aware Node Selection.*** The server gathers AoA values from sensor nodes, and uses triangulation to localize the sound source. Later, a subset of sensor nodes is selected to request audio features.
- ***Audio Feature Extraction.*** In this task, audio features are extracted by the selected nodes from the sound event.

- ***Multi-Node Cough Detection.*** The server collects audio features from the selected nodes, and uses Gaussian mixture models [7] to classify the sound event into cough or non-cough.

The tasks are modularized into three task sets, as shown in Figure 3, and separated by dotted lines. The *Server Tasks* contain centralized operations on the CoughLoc server. The *Primary Sensor Tasks* contain operations on every sensor node. The *Additional Sensor Tasks* contain operations only on the selected sensor nodes.

Introducing task sets into the system design has three advantages:

1. Programming the *Primary Sensor Tasks* and the *Additional Sensor Tasks* together on each node enables dynamic task switch based on changing environmental noise conditions over time.
2. Eliminate the need of additional dedicated hardware (e.g. DSPs), and simplify programming.
3. The server assigns the *Additional Sensor Tasks* to only a subset of nodes, reducing computations and data transmissions on others.

Table 1 summarizes average processing time and bandwidth requirements for processing one audio frame (i.e. 16ms). The calculation is based on extracting the 13-dimension Mel Frequency Cepstral Coefficients (MFCC) features [5] from each audio frame. After task partitioning, sensor nodes *not* selected by the server has 33% less processing time and 98% less bandwidth requirement.

Table 1: Processing time and bandwidth requirement for processing one 16ms audio frame (‡selected nodes only)

	Processing Time	Bandwidth
Sound Event Detection	1.1ms	0 bit
AoA Estimation	9.3ms	<10 bits
Feature Extraction‡	5.2ms	416 bits

3.2 Sound Event Detection

We use a frame-based admission control approach [8] to detect possible sound events. Each sensor node first segments continuous audio samples into audio frames, then the root mean square value (RMS) of each frame is computed. To handle hardware variation, we estimate each microphone's circuit noise by recording ambient sound under quiet conditions. The minimum RMS value $RMS_{circuit}$ of the recorded ambient sound is used to quantify circuit noise, and the average RMS value $RMS_{average}$ is used to represent general ambient sound level. Then for each incoming frame, if its RMS value RMS satisfies

$$RMS - RMS_{circuit} < thd \cdot (RMS_{average} - RMS_{circuit}), \quad (1)$$

where *thd* is a threshold indicating ambient sound level relative to the quiet conditions, the frame is considered as ambient noise and discarded; otherwise, a sound event is detected, and the node saves the current frame into buffer. The node repeats this process for the next frame, until either an incoming frame satisfies (1) and thus is discarded, or the buffer becomes full. During the experiments we conservatively set *thd* to 1.10 based on empirical measurements to reduce probability of missing distant sounds. However, an adaptive threshold could also be applied to further reduce computation.

After detecting sound events, each sensor node aggregates consecutive events to reduce computation. Here we use a temporal proximity measure Δt to quantify the adjacency of the events, such that all sound events detected within Δt are considered to have come from the same sound source. We note that the determination of the value of Δt depends highly on types of the sound sources. For example, with stationary sound sources generating continuous sound, such as a microwave or a TV, a larger Δt value could be applied to reduce localization computation times; however, if the sound source generates abrupt but short sounds, such as a coughing patient, a smaller Δt value must be used to avoid missing events. Since cough sounds generally have a duration less than 1 second, we conservatively set Δt to 48ms (i.e. three audio frames).

3.3 Localization of Sound Sources

To improve the audio SNR, the CoughLoc server requests audio features from sensor nodes that are physically close to the sound source. We use sensor proximity to the sound source to avoid computationally expensive SNR estimation. This requires robust localization of the sound source, which consists of two steps: 1) Angle of acoustic signal arrival (AoA) estimation, and 2) AoA-based triangulation.

AoA Estimation. After detecting one sound event, each sensor node calculates the angle from which the signals arrive by estimating the time difference of signal arrival (TDoA) between the two microphones on each node.

AoA-based Triangulation. After estimating AoA values, the CoughLoc server gathers the estimates from sensor nodes and uses every two estimates to triangulate a location candidate of the sound source. Considering the skewness of the distribution of localization errors, the server determines the source location as the median value of all location candidates.

3.4 Location-Aware Node Selection

Given locations of sensor nodes and estimated locations of sound sources, the server selects sensor nodes by comparing their distances to the sound source. Because of variation of microphone sensitivity and effects from indoor echoes and noises, being physically close to the sound source suggests but not guarantees a higher detection performance. In order to take these varying factors into account, the CoughLoc server select a sensor node if it is within d meters to the sound source. If none of the nodes meets this criterion, the closest node is selected implicitly.

The value of d is determined by both localization accuracy and degradation speed of the detection performance. For example, if the estimated sound source location were l meters away from the actual sound source, the maximal distance from the selected sensor nodes to the actual sound source would be $l + d$ meters. Through evaluations in indoor settings, we found CoughLoc could localize the sound source to within 1m at 50th percentile. Therefore, in the experiments we set d as 2m such that the average distance from the selected nodes to the actual sound source would be within 3m, which can still maintain 70% detection accuracy under all tested noise conditions. We would like to note, however, that this d value is fairly conservative given the current localization accuracy, which can be further improved by using more accurate localization algorithms, such as the Approximate Maximum Likelihood algorithm used in [9].

3.5 Feature Extraction and Multi-Node Cough Detection

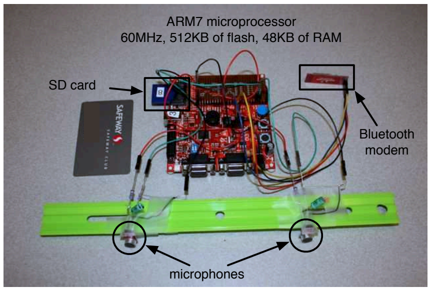


Figure 4: A CoughLoc sensor node with a credit card placed aside. The two microphones are fixed on a ruler to maintain constant inter-microphone distance.

We formulate the cough detection problem as a binary sound classification problem. The goal is to determine if each audio frame has come from a cough sound or a non-cough sound. To this end, we need an audio feature that is discriminative enough to represent the uniqueness of cough sounds. Although cough sounds can also be used for patient identification, in this work, we restrict our focus on general cough detection as a proof of concept. Interested readers can be referred to [3], [4], and [10]. Generally, coughs feature broad and flat power spectrum ranging from 0Hz to 15Hz, with varying duration from 0.1s to 1s. To keep our detection algorithm comparable with related work on cough detection, we choose to use a relatively complex but discriminative audio feature, the Mel Frequency Cepstral Coefficients (MFCC) [5], which is widely used in cough detection and other audio processing research [3, 4, 8].

Each selected node first calculates the 13-dimension MFCC features as well as their 1st- and 2nd-order derivatives, then transmits all the features and the derivatives to the server. Two Gaussian mixture models (GMM) [7], each having 4 Gaussian distributions, are trained on the server to represent cough and non-cough events, respectively. After receiving MFCC features, the CoughLoc server first classifies sound events based on audio features from each node, and then makes a final classification using a weighted majority vote as shown below.

$$S = K \cdot \sum_{i=1}^{K} w_i \frac{a_i}{\sum_{i=1}^{K} w_i}, \tag{2}$$

where K is the number of the selected sensor nodes, a_i the per-node GMM classification result (1 for cough, 0 for non-cough), and w_i the corresponding likelihood for the ith node. If $S > K/2$, a cough, otherwise a non-cough, is reported by the server.

4. SYSTEM IMPLEMENTATION AND EVALUATION

In order to evaluate the system, we built the sensor nodes and deployed them in indoor environments for testing. Each node features an ARM7-based 60MHz CPU, 48KB memory, and a Bluetooth modem. Each node also has two microphones that are placed 15cm apart. Figure 4 shows a CoughLoc sensor node. Since the nodes are deployed indoors, we assume that the nodes are connected to a power source. The extraction of MFCC features and the training of GMMs are implemented using the HTK toolkit [11].

4.1 Experimental Traces and Setup

To ensure repeatability of the experiments, we first recorded 12 minutes of clear human coughs using commercial headsets. To fully reproduce the coughs, the recording was performed in a sound isolated room. This cough audio was then played back using a loudspeaker during the experiments.

The experiments were done in an 8×6 m^2 room, where ambient sound was captured concurrently by 8 CoughLoc nodes distributed throughout the room. We placed the loudspeaker at eight uniformly distributed locations in the room and replayed the recorded coughs. We tested three representative types of noises to model various indoor noises: 1) human speech-like sound (television); 2) continuous low and steady hum (microwave oven); 3) continuous loud and grating sound (vacuum cleaner). All the noises were generated using corresponding devices. In the TV case, an ABC World News program was played using a Sony HDTV, containing human speech and music. For each sensor node, audio was captured using the two microphones at 10bit, 16KHz and saved on a local SD card.

Throughout the experiments, 16 microphones on the 8 sensor nodes collected a total of 40 hours of training data ($\sim$35,400 coughs) and 5 hours of test data ($\sim$4,700 coughs).

4.2 Performance of Localizing Sound Sources

Sound source localization is a crucial task of CoughLoc. Its accuracy affects the correct selection of sensor nodes that tend to have better audio quality. This section presents the evaluation of the localization performance, and discusses possible sources of localization errors.

4.2.1 *Performance of Localizing Sound Sources*

As the omni-directional microphones on each CoughLoc node features a limited acquisition range [5], we expect the localization error to be small to provide robust and correct node selection. Given limited inter-microphone distance, supported sampling rates, and errors existing in AoA estimation, localization performance is highly limited.

In this section, we first evaluate the performance of source localization relating to the sampling rate to explore the limit of hardware capability. We then compare different approaches to deal with AoA estimation errors by computing mean and median of pair-wise source location candidates. Figure 5 shows the cumulative distribution functions of errors (CDF), which indicate the probability of a measurement occurring below a certain error. We tested the localization performance under 16KHz, 8KHz, and 4KHz sampling rates by downsampling the raw audio signals. As shown in the figure, the accuracy drops gradually as the sampling rate decreases. Specifically, under 8KHz, the average localization error is 1m higher than that under 16KHz. Under 4KHz, the errors become as large as 3.7m and 4.5m when taking mean and median of the candidate estimates, respectively.

With 16KHz sampling rate, the localization algorithm achieves 1m localization error at 50th percentile. As mentioned earlier in Section 3.4, we set node selection range d to 2m such that the average distance from the farthest selected node to the sound source will be 3m. In the following evaluations, we keep the sampling rate at 16KHz.

4.2.2 *Factors Affecting Localization Accuracy*

Based on the experimental results shown in Figure 5, in this section, we discuss two possible factors that may affect the localization performance.

Indoor echoes. After a sound is emitted, one or many echo sounds are bounced back and forth against walls and

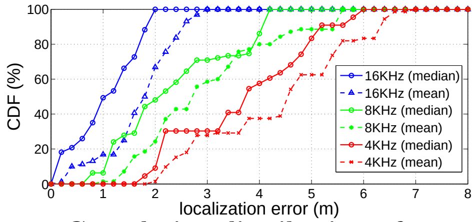


Figure 5: Cumulative distribution of errors in localizing the cough source when changing sampling rates from 16KHz to 4KHz.

other obstacles, overlapping the original sound. The shape and amplitude of the resulting mixed sound differ from the original in both time and frequency domains. This makes it difficult to distinguish time difference of sound arrivals between the two microphones on each node, leading to degradation in the estimation of AoA.

Intrinsic AoA estimation errors. The estimation of AoA values involves quantizing continuous time differences of signal arrivals as discrete audio sample delays. The range of this delays is constrained by the speed of sound, the sampling rate applied, and the inter-microphone distance. Because of these restrictions, the range of possible sample delays is highly limited, resulting in intrinsic AoA errors.

One way to improve AoA estimation accuracy is to employ algorithms that support intra-sample cross-correlation resolution, resulting in continuous sample delays. In addition, increasing the sampling rate or the inter-microphone distance will reduce the errors caused by AoA estimation as well. Such improvements remain the focus of our future work.

4.3 Performance of Cough Detection

By leveraging the information of sound source locations, CoughLoc applies a location-aware acoustic sensing and event detection approach to reduce effects stemming from ambient indoor noises and long distance from the sound sources.

4.3.1 Single-Node Baseline Approach

We first describe performance of a baseline approach that does not consider the location information. In this approach, we use a basic node selection strategy that randomly selects one sensor node in the network to extract audio features after a sound event is detected. The detection performance is expressed as precision and recall values [3], which are defined as

$$precision = \frac{\#correctly\_classified\_coughs}{\#total\_classified\_coughs}, \quad (3)$$

and

$$recall = \frac{\#correctly\_classified\_coughs}{\#total\_real\_coughs}. \quad (4)$$

Further, we use a unified metric F_1 score [3, 7] for each case to simplify the comparison. The F_1 score is a widely used classification performance metric that takes both precision and recall values into account equally. The F_1 score is defined as

$$F_1 = \frac{2 \cdot precision \cdot recall}{precision + recall}. \quad (5)$$

With range from 0 to 1, the higher the F_1 score, the better the classification performance. Figure 6 shows the relation between the F_1 score and the distance between the selected sensor node and the loudspeaker. When the randomly selected node is within 2m from the cough source, the F_1 scores are about 90% on average, indicating a good detection performance. However, as the distance increases from 1m to 7m, the F_1 score decreases significantly. The maximum decrease is 75% under microwave oven case, and for all noise types the average decrease is about 60%. These results indicate the single-node detection approach is greatly affected by the increased distance and thus less applicable for distant cough monitoring. Therefore, in the CoughLoc system, we incorporate the information of sound source location, and use this information to select sensors close to the sound source to improve cough detection performance.

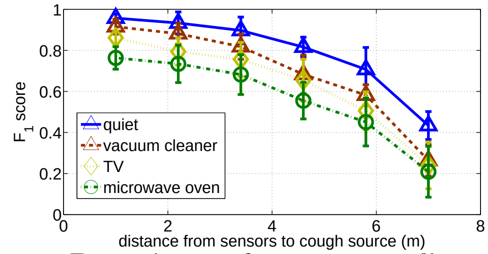


Figure 6: Detection performance vs. distance from the sound source to a randomly selected sensor node. As the distance increases from 1m to 7m, the F_1 score decreases significantly.

4.3.2 Collaborative Detection Performance

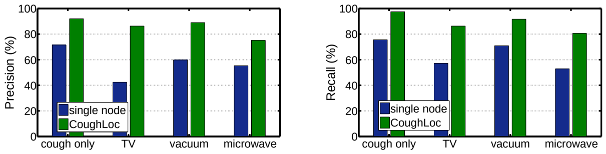
This section presents classification performance of CoughLoc using collaborative multi-node classification. The single-node approach mentioned in the previous section is used as the baseline. In the baseline approach, every audio frame is processed using one randomly chosen sensor node, regardless of its distance to the sound source. In the collaborative approach, the location-aware sensor node selection and multi-node cough detection algorithms described in Section 3 are used. Figure 7 shows the precision and recall values for both the approaches.

As shown in Figure 7(a), the CoughLoc system achieves 92.6%, 86.3%, 89.3%, and 75.1% precisions under the four noise conditions, respectively. In contrast, due to the large difference of the distances between microphones and the cough source, the single-node baseline approach has much lower precision, especially under the TV noise. This is explainable by seeing the similar spectrum energy distribution between cough sounds and other human-voice like sounds.

Figure 7(b) shows that the CoughLoc system achieves 97.5%, 85.7%, 91.7%, and 80.5% recalls under the four noise conditions, respectively. Combining three noisy cases together, CoughLoc achieves 0.95 and 0.84 F_1 scores under quiet and noisy environments, which are 1.3 times and 1.6 times higher than those of the baseline approach. We note that CoughLoc achieves even higher performance gain under noisy environments than under quiet environments. This observation justifies the use of multi-node collaborative detection to counteract indoor noise and echoes, and thus significantly improves the acceptability of ambient sound monitoring.

5. RELATED WORK

The CoughLoc system is designed as a distributed acoustic sensing network for indoor cough detection. As the emerging



(a) Comparison of precisions (b) Comparison of recalls

Figure 7: Detection performance comparing the single-node baseline approach and the multi-node collaborative approach in the CoughLoc system.

development of sensor networks, acoustic sensing has generated considerable research interest. A number of acoustic sensing systems have been developed [12, 13]. However, these systems' goals and constraints differ from ours. For example, some of previous work usually demand high accuracy of sound source localization. As a result, their localization modules either require computationally-intensive algorithms (such as the one used in ENSBox [12]), or are highly optimized for specific applications (such as the one in the CounterSniper system [13]). In contrast, CoughLoc uses a basic AoA-based localization algorithm that requires less computation but achieves sufficient accuracy for the cough detection application.

In the field of pervasive medical monitoring, body-worn sensors are extensively used in previous work. For example, Doukas et al. [14] present a patient monitoring system for activity recognition and emergency treatment using accelerometers and microphones attached on user bodies. Mercury [15] is a system designed for motion analysis of neuromotor disorders that needs patients to wear up to eight sensor nodes for monitoring movement and physiological conditions. A drawback of this type of systems is the patients often forget to wear or take off the sensors due to discomfort. In contrast, a non-intrusive monitoring system like the CoughLoc system can greatly improve the convenience of health monitoring and have greater acceptance among patients.

Prior work with the most similarities to the CoughLoc system has been in the field of ambulatory cough monitoring [3, 4, 10]. Matos et al. [3, 10] present a system that achieves 94.7% precision and 85.7% recall rate under quiet environments, which are very similar to CoughLoc. However, the authors do not mention how their system works in noisy environments. Other automatic cough detection systems, such as [16], adopt acoustic sensing devices that need to be held by caretakers in short range of a patient. The intrinsic inconvenience of such devices makes them difficult to be accepted and used. In contrast, CoughLoc uses distributed acoustic sensing nodes to avoid intrusiveness. By leveraging the information of sound source locations in multi-node collaborative detection, CoughLoc also achieves high detection performance.

6. SUMMARY

In this paper, we present CoughLoc, a non-intrusive cough detection system, which leverages location information to improve classification performance. In addition, we present a location-aware task partitioning strategy that dynamically assigns sensing tasks to different nodes, reducing overall computations and bandwidth requirement. Experimental results show that the system achieves over 90% accuracy under quiet environments. Under noisy environments, the system achieves 1.6 times higher accuracy than a baseline approach without considering location information. Compared to the state-of-the-art ambulatory cough monitors, CoughLoc achieves high cough detection performance for non-intrusive monitoring.

Acknowledgments

This research was supported by CyLab at Carnegie Mellon under grants DAAD19-02-1-0389 and W911NF-09-1-0273 from the Army Research Office.

7. REFERENCES

[1] R.M. Droes, M. Mulvenna, C. Nugent, D. Finlay, M. Donnelly, M. Mikalsen, S. Walderhaug, T. van Kasteren, B. Krose, S. Puglia, and Others. Healthcare Systems and Other Applications. *IEEE Pervasive Computing*, 6(1):59–63, 2007.

[2] KF Chung and JG Widdicombe. *Cough : Causes , Mechanisms and Therapy*. Blackwell Publishing Ltd, 2003.

[3] Sergio Matos, Surinder S Birring, Ian D Pavord, and David H Evans. An automated system for 24-h monitoring of cough frequency: the leicester cough monitor. *IEEE transactions on bio-medical engineering*, 54(8):1472–9, August 2007.

[4] Sung-Hwan Shin, Takeo Hashimoto, and Shigeko Hatano. Automatic detection system for cough sounds as a symptom of abnormal health condition. *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, 13(4):486–93, July 2009.

[5] M.S. Brandstein and H.F. Silverman. A practical methodology for speech source localization with microphone arrays. *Computer Speech & Language*, 11(2):91–126, April 1997.

[6] I.A. McCowan, A.C. Morris, and H. Bourlard. Improving speech recognition performance of small microphone arrays using missing data techniques. In *Seventh International Conference on Spoken Language Processing*, pages 4–7. Citeseer, 2002.

[7] C.M. Bishop. *Pattern recognition and machine learning*, volume 4. Springer New York, 2006.

[8] H. Lu, W. Pan, N.D. Lane, T. Choudhury, and A.T. Campbell. SoundSense: scalable sound sensing for people-centric applications on mobile phones. In *Proceedings of the 7th international conference on Mobile systems, applications, and services*, pages 165–178. ACM New York, NY, USA, 2009.

[9] A.M. Ali, Shadnaz Asgari, T.C. Collier, Michael Allen, Lewis Girod, R.E. Hudson, Kung Yao, C.E. Taylor, and D.T. Blumstein. An empirical study of collaborative acoustic source localization. In *IPSN'07*, volume 57, pages 415–436. ACM, November 2007.

[10] S Matos, S.S. Birring, I.D. Pavord, and H. Evans. Detection of cough signals in continuous audio recordings using hidden Markov models. *Biomedical Engineering, IEEE Transactions on*, 53(6):1078–1083, 2006.

[11] Hidden markov model toolkit. `http://htk.eng.cam.ac.uk/`, 2010.

[12] Andreas M. Ali, Kung Yao, Travis C. Collier, Charles E. Taylor, Daniel T. Blumstein, and Lewis Girod. Acoustic source localization using the acoustic ENSBox. *Proceedings of the 6th international conference on Information processing in sensor networks - IPSN '07*, page 555, 2007.

[13] Á. Lédeczi, A. Nádas, P. Völgyesi, G. Balogh, B. Kusy, J. Sallai, G. Pap, S. Dóra, K. Molnár, M. Maróti, and Others. Countersniper system for urban warfare. *ACM Transactions on Sensor Networks (TOSN)*, 1(2):177, 2005.

[14] Charalampos Doukas and Ilias Maglogiannis. Advanced Patient or Elder Fall Detection based on Movement and Sound Data. *The Proceedings of the Second ICST International Conference on Pervasive Computing Technologies for Healthcare 2008*, 2008.

[15] Konrad Lorincz, Bor-rong Chen, G.W. Challen, A.R. Chowdhury, Shyamal Patel, Paolo Bonato, and Matt Welsh. Mercury: A wearable sensor network platform for high-fidelity motion analysis. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*, pages 183–196. ACM, 2009.

[16] Y. Hata, S. Kobashi, H. Yamaguchi, O. Ishikawa, N. Tsuchiya, and H. Nakajima. Human health monitoring system of systems by non-contacted senors. In *System of Systems Engineering, 2009. SoSE 2009. IEEE International Conference on*, pages 1–5. IEEE, 2009.