



# HearFire: Indoor Fire Detection via Inaudible Acoustic Sensing

ZHENG WANG, Hunan University, China

YANWEN WANG, Hunan University, China

MI TIAN, Hunan University, China

JIAXING SHEN\*, Lingnan University, Hong Kong, China

Indoor conflagration causes a large number of casualties and property losses worldwide every year. Yet existing indoor fire detection systems either suffer from short sensing range (e.g.,  $\leq 0.5m$  using a thermometer), susceptible to interferences (e.g., smoke detector) or high computational and deployment overhead (e.g., cameras, Wi-Fi). This paper proposes HearFire, a cost-effective, easy-to-use and timely room-scale fire detection system via acoustic sensing. HearFire consists of a collocated commodity speaker and microphone pair, which remotely senses fire by emitting inaudible sound waves. Unlike existing works that use signal reflection effect to fulfill acoustic sensing tasks, HearFire leverages sound absorption and sound speed variations to sense the fire due to unique physical properties of flame. Through a deep analysis of sound transmission, HearFire effectively achieves room-scale sensing by correlating the relationship between the transmission signal length and sensing distance. The transmission frame is carefully selected to expand sensing range and balance a series of practical factors that impact the system's performance. We further design a simple yet effective approach to remove the environmental interference caused by signal reflection by conducting a deep investigation into channel differences between sound reflection and sound absorption. Specifically, sound reflection results in a much more stable pattern in terms of signal energy than sound absorption, which can be exploited to differentiate the channel measurements caused by fire from other interferences. Extensive experiments demonstrate that HearFire enables a maximum 7m sensing range and achieves timely fire detection in indoor environments with up to 99.2% accuracy under different experiment configurations.

CCS Concepts: • Human-centered computing → Ambient intelligence; Mobile computing;

Additional Key Words and Phrases: Indoor fire detection, acoustic sensing, channel impulse response, contact-free

## ACM Reference Format:

Zheng Wang, Yanwen Wang, Mi Tian, and Jiaxing Shen. 2022. HearFire: Indoor Fire Detection via Inaudible Acoustic Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 4, Article 185 (December 2022), 25 pages. <https://doi.org/10.1145/3569500>

## 1 INTRODUCTION

Indoor conflagration has long been an enormous threat that causes significant loss of human life and damage to property. Since 2011, an annual average of over 62,000 conflagrations burn millions of acres of land and kill thousands of people each year worldwide [9]. For instance, a recent conflagration of Notre-Dame de Paris

\*This is the corresponding author.

---

Authors' addresses: Zheng Wang, wangzheng@hnu.edu.cn, College of Electrical and Information Engineering, Hunan University, Changsha, Hunan, China; Yanwen Wang, wangyw@hnu.edu.cn, College of Electrical and Information Engineering, Hunan University, Changsha, China; Mi Tian, tianmi@hnu.edu.cn, College of Electrical and Information Engineering, Hunan University, Changsha, China; Jiaxing Shen, jiaxingshen@LN.edu.hk, The Department of Computing and Decision Sciences, Lingnan University, Hong Kong, China.

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2022 Association for Computing Machinery.

2474-9567/2022/12-ART185 \$15.00

<https://doi.org/10.1145/3569500>

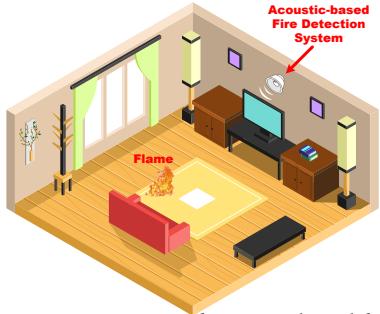


Fig. 1. An usage scenario of acoustic-based fire detection system.

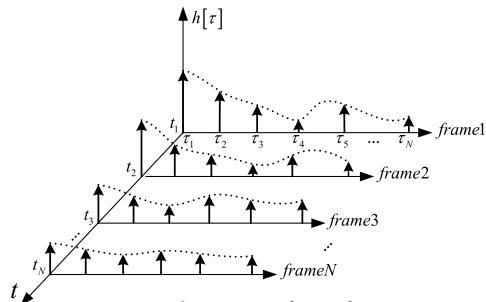


Fig. 2. Three-D surface of CIR.

Cathedral in 2019 almost completely destroyed this world-famous historical relic, resulting in irretrievable damage to human society and contaminating the city with toxic dust and lead [37]. Therefore, achieving timely and effective fire detection becomes increasingly crucial for preventing loss of life and property, especially for early fire alarm and rescue. Recent reports show that the global market for fire alarm and detection systems is projected to reach 98.90 billion dollars by 2030, registering a compound annual growth rate of 5.0 percent [1].

Existing fire detection systems can be roughly divided into two categories: contact-based detection and contact-free detection. Contact-based detection systems have been widely applied in our daily life [20, 21, 36, 44, 51, 53], which employ various types of sensors to sense the nearby environment (i.e., temperature [20, 53], smoke [21, 36], etc.). The intuition is that fire detection can be easily achieved when the nearby environment of the detectors changes and the sensor values reach a pre-defined threshold. However, thermometer-based sensors require close distance or even direct contact with the flame [20, 53]. Existing smoke detectors mainly refer to ionization and photoelectric sensors [33], which enable distant and timely smoke detection. Nevertheless, dust particles, dirt, and other airborne contaminants in the environments may easily cause false detection of fire for photoelectric sensors [14, 34, 44], while ionization smoke detectors will release radioactive material when they are burning, posing an inhalation hazard [10]. Moreover, the non-flexible installation and high maintenance overhead of contact-based detectors further constrain their applicability [36, 51].

Unlike contact-based fire detection, contact-free systems do not require any direct contact with fire while still achieving fire detection. Contact-free fire detection approaches leverage different physical properties of fire, which can be captured by the corresponding detectors. Thermal radiation emitted by flame can be detected by infrared optical detectors, enabling distant fire detection [4, 12, 35]. However, they are easily interfered by the same infrared radiation band caused by other devices (e.g., microwave ovens, remote controllers) and are challenging to meet the timely detection requirement because infrared can be easily blocked by smoke. Image-based fire detection system applies cameras for monitoring fire and deep learning technologies for fire detection [25, 63], which can achieve a high detection accuracy. However, cameras highly rely on Line-Of-Sight (LOS) path and are hard to work with occluded or contaminated lenses. Recent studies exploit Wi-Fi signals to detect fire due to their light-independent and low computational overhead properties [24, 62]. The intuition is that fire inflammation impacts the propagation of Wi-Fi signals, resulting in the variations of amplitude and phase of Channel State Information (CSI). However, in many applicable places, such as military bases or remote rural areas, Wi-Fi connections are not accessible. Moreover, the Wi-Fi measurement is less fine-grained than acoustic signals, as RF signals have a longer wavelength than sound waves [8].

In recent years, acoustic-based sensing technology has attracted extensive attention from both industry and academia due to its many irreplaceable advantages, including pervasive hardware infrastructure, cost-effective deployment as well as high time/space resolution [7, 23, 27, 39, 41, 47, 55, 56]. In this paper, we aim to implement

a cost-effective, easy-to-use and timely room-scale fire detection system using acoustic signals. We envision our proposed system employed indoors, such as warehouses, homes, offices, or any room-scale environment for timely fire detection, as illustrated in Fig. 1. We note that our system is not to replace existing fire detection techniques (e.g., smoke detection and RF sensing), but to provide a complementary technique that could cover most usage scenarios in indoor environments. Our intuition is that fire combustion poses detectable effects on acoustic signals, which can be captured and processed to infer the emergence of flames.

However, implementing an acoustic-based fire detection system is a daunting task. One fundamental challenge facing fire detection is how to measure the influences on acoustic signals caused by fire. Most existing acoustic sensing technologies fulfill sensing tasks leveraging signal reflection effect from target objects [7, 29–31, 38, 39, 41, 47, 56]. The reason behind is that the reflected signals inherently involve objects' temporal and spatial features, which can be utilized to infer diverse sensing tasks. However, unlike previous works, fire is intrinsically incapable of reflecting any acoustic signal. This is because the physical essence of flame is plasma [5], which does not reflect any acoustic signals, making it undesirable for fire detection using signal reflection. More importantly, due to the multipath effect, only a few sound waves are influenced by fire, which are superimposed with other multipath sound waves at the receiver. Therefore, sound signals impacted by flames are difficult to extract from the superimposed signal due to relatively weak impacts posed by fire.

To address this challenge, instead of using the signal reflection effect, we exploit two significant sound propagation characteristics: (1) sound energy can be absorbed by fires when penetrating the flame [2] and (2) the speed of sound propagating in the air varies with different temperatures [3]. Specifically, we transmit a sound wave with high auto-correlation and apply Channel Impulse Response (CIR) to simultaneously measure the energy loss and different Time-of-Flight (ToF) of sound during propagation. On the one hand, high auto-correlation property ensures successful separation of similar sound copies impacted by fire from the superimposed signal by finding the peaks corresponding to their different delays. On the other hand, CIR can characterize all signal multipath in terms of different propagation delay ranges. Signals with similar propagation delays within a specific range are superimposed in a single tap. By focusing on the magnitude of a certain number of CIR taps with signal absorption and sound variation pattern, we can effectively extract useful features from the fire combustion. Our study shows that these two unique characteristics of sound propagation exhibit unique and distinguishable patterns in CIR before and after flame appears, which offers us an opportunity for fire detection using acoustic signals.

Another challenging issue lies in significant signal attenuation, especially in sizable room-scale areas. As the propagation distance increases, the received acoustic signal decays dramatically. More importantly, fire further absorbs the signal energy, resulting in an extremely weak sound signal. Traditional works apply extra devices (e.g., RF amplifier) or antenna with high transmission gain to guarantee sizeable areas, which inevitably increases the hardware complexity as well as system cost. Recent works tackle weak acoustic signals and enlarge sensing range by emitting a longer acoustic sequence, which covers longer propagation paths while maintaining the system cost. However, through our analysis, simply increasing the length of emitted sequence may fail to detect the fire due to many practical constraints, including narrow inaudible frequency band, low system time-resolution, and insufficient feature extraction (will be explained in detail in Section 4.1.2 and Section 4.1.3).

To enable room-scale fire detection, we synthetically analyze the practical constraints of sound covering sizeable areas and correlate the relationship between the length of the transmission frame and the system sensing range. Specifically, we apply Zadoff-Chu (ZC) sequence as the emitted sequence. Different from existing works that primarily use ZC sequence for near-field sensing (e.g., <2m), we extend the usage of ZC sequence to realize far-field sensing (e.g., >6m) by considering comprehensive parameter selection of ZC sequence under fire detection scenario. The parameters of the ZC sequence are carefully configured to achieve inaudible, room-scale, high sensitivity and high accuracy detection.

The third challenge is how to effectively distinguish between fire flame and other environmental interferences. Due to the multipath effect, static objects in the environment (e.g., furniture, desktop, wall, etc.), as well as other moving objects (e.g., people moving nearby) will impose significant interference on CIR measurements caused by fire, degrading the performance of fire detection system.

To distinguish CIR measurements caused by fire from other interfering elements, we deeply investigate the root differences between sounds absorbed by fire and that reflected by other entities. These differences contribute to more stable energy distribution in CIR measurements when reflecting from entities than being absorbed by fire. To this end, we design a simple yet effective method by calculating the sparsity of the CIR measurements caused by fire and other interferences. CIR measurements with high sparsity indicate an unstable energy pattern, which is further input into a classifier for the final detection of fire.

Such a holistic design allows us to implement an acoustic-based fire detection system and achieves high detection accuracy for timely room-scale fire detection, while without posing extra overhead on hardware and computation. In our experiment, HearFire can achieve high detection accuracy in different environments under different experimental configurations. Our contribution can be summarized as follows:

- To the best of our knowledge, HearFire is the first system that exploits inaudible acoustic signals for fire detection. We address fire sensing by leveraging two significant sound propagation characteristics and applying CIR to capture the impacts of sound caused by fire.
- We extend the usage of ZC sequence to realize room-scale sensing by considering the tradeoff among a series of practical constraints that impact the system performance and conducting a comprehensive parameter selection of ZC sequence under a fire detection scenario. The transmission frame is carefully designed with appropriate properties for inaudibility and enlarging the sensing range.
- We remove the interferences caused by other reflections and implement a prototype of HearFire. After conducting extensive evaluation, our experiment results show that HearFire achieves a fire detection accuracy of up to 99.2% in different environments.

## 2 BACKGROUND

### 2.1 Fire and Flame

Fire is a strong and rapid oxidative reaction that occurs during combustion, in which energy is released in terms of light and heat [40]. The visible part of the fire is flame, which is essentially the released photons when the atoms of the air in the excited state jump to the ground state, resulting in intense high-speed motion of air molecules, thereby emitting light and heat. The fire intensity is reflected by the color of the flame, which depends on the substances ignited, and any impurities outside. Combustion releases a large amount of energy due to chemical reaction, which dramatically increases surrounding air temperature (e.g., hundreds of degrees Celsius). Actually, the physical essence of the fire is plasma [5] with unfixed shape and volume and therefore is intrinsically incapable of reflecting any mechanical wave.

### 2.2 Propagation of Acoustic Signals

As a wave of pressure, sound propagates via compressible media and can penetrate, be reflected, and be absorbed by the media. The reflection effect of sound has been widely studied for sensing and detection, as most target objects can reflect the sound signal. However, flame is unable to reflect any sound, which facilitates us to discover alternative propagation characteristics for fire detection.

**2.2.1 Absorption of Sound Energy by Fire.** When a sound wave encounters flame, it will experience two steps: (1) part of the sound energy will be absorbed by the fire, converting to heat [2]. This is because high-speed motion of air molecules nearby the flame generates intense air vibration, which results in viscous friction in the air. This internal frictional force translates a certain amount of sound energy into heat when sound travels nearby, giving

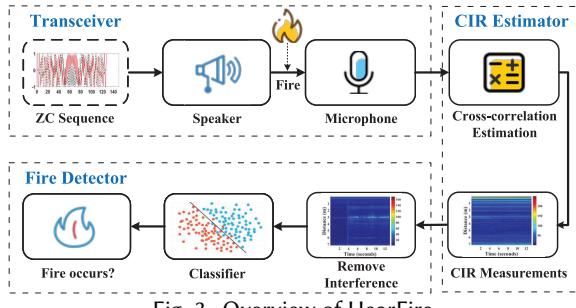


Fig. 3. Overview of HearFire.

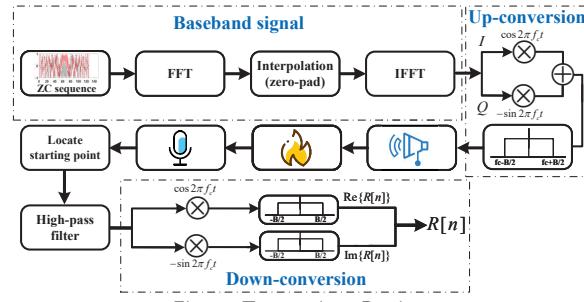


Fig. 4. Transceiver Design.

rise to the energy absorption of the acoustic signal. (2) rest of the sound will penetrate the flame, continuing to propagate in the air. This part of sound can be recorded by acoustic devices if with sufficient residual energy.

**2.2.2 Variation of Sound Speed with Temperature.** The speed of sound is affected by the temperature of the transmission media. A higher temperature of the transmission media results in a faster sound speed. The reason is that molecules at higher temperatures have more energy so that they vibrate faster, allowing sound waves to travel more quickly and, hence, increasing the speed of sound. The relationship between the speed of sound and temperature can be roughly expressed as  $v = v_0 + 0.606T$  [3], where  $v_0 = 331.3$  m/s denotes the sound propagation speed at freezing temperatures,  $T$  is the temperature in degrees Celsius. These two crucial properties of sound propagation potentially offer us an opportunity to sense the fire using acoustic signals. We note that although sound speed is likewise impacted by humidity and air pressure, this impact is several magnitudes less than that caused by fire [13], which can be neglected in our work.

### 2.3 Channel Impulse Response

When a transmitter transmits a signal to the receiver, the signal will propagate along multiple paths due to reflection, scattering, and diffraction. Multipath causes time dispersion, generating multiple copies of signals with different attenuation and delays, which arrive at the receiver and form a superimposed signal. Recent works implement acoustic-based tracking and recognition systems by measuring the CIR of the received signal [26, 39, 47]. CIR is capable of separating multipath signals on the basis of different propagation delays, discretizing multipath signals into delay bins (also known as taps). Each tap represents multipath signals with a specific propagation delay range. Therefore, one may easily focus on those multipath signals merely impacted by the target objects by finding proper bins and can effectively observe signal variations in terms of delay and attenuation.

In practice, CIR can be measured by emitting a pre-defined frame, which is utilized to measure the communication channel and received by the receiver. For complex-valued baseband signals, CIR can be mathematically measured as cyclic cross-correlation between the transmitted and the received baseband signal [39]:

$$h[\tau] = \frac{1}{N} \sum_{n=1}^N s[n] \cdot r^*[n - \tau]. \quad (1)$$

where  $h[\tau]$  is the CIR measurements,  $\tau$  is the multipath delay and  $N$  is the length of the signal.  $s[n]$  denotes the transmitted complex baseband frame and  $r^*[n]$  is the conjugation of the received complex baseband signal  $r[n]$ . The measured CIR is a 3-D matrix, as illustrated in Fig. 2. For each frame in  $r[n]$ , performing cross-correlation with  $s[n]$  to measure the CIR  $h[\tau]$  profiles the signal attenuation corresponding to different delays.

## 3 SYSTEM OVERVIEW

In this work, we propose an acoustic-based room-scale fire detection system in a contact-free manner called HearFire. Fig. 3 illustrates the overview of HearFire. HearFire consists of three main components: Transceiver,

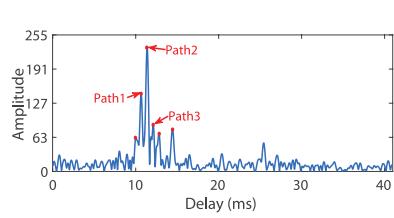


Fig. 5. A single CIR measurement when sound waves penetrate fire.

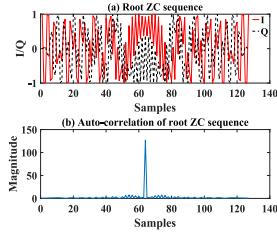


Fig. 6. Root ZC sequence and its auto-correlation.

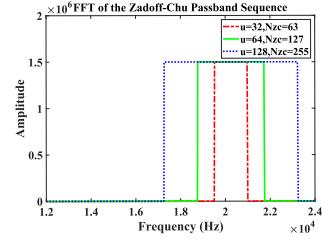


Fig. 7. FFT result of three passband ZC sequences.

CIR Estimator, and Fire Detector. In Transceiver, an inaudible frame is played repeatedly by the speaker and then recorded by the microphone. In our work, we place the speaker and microphone close together and fix the system layout to mitigate the impact of LOS propagation. To increase the sensing range, we dedicatedly design this inaudible frame for channel measurement, which is capable of covering sizable room-scale areas. After receiving the signal frame, a channel estimator is applied to estimate CIR using cross-correlation estimation. The estimated CIR simultaneously captures the residual energy and ToF of the sound wave traveling through the flame. Finally, Fire Detector jointly exploits these two features in extracted CIR measurement to remove the impact of other interfering elements, followed by applying a classifier to detect the fire. HearFire can achieve timely fire detection by immediately issuing an alarm when a flame appears.

## 4 SYSTEM DESIGN

### 4.1 Transceiver Design

The transceiver of HearFire consists of a commodity speaker and a commodity microphone, which play the functions of sending and receiving, respectively, as illustrated in Fig. 4. The transmitter unceasingly emits a modulated pre-defined acoustic frame for measuring the channel, while the receiver records the acoustic signal and measures the real and imaginary parts of the signal for CIR estimation.

**4.1.1 Transmission Frame Selection.** As sound waves have already been applied to measure the attenuation and sound speed, at the beginning of our study, we tried to sense the fire using pure Continuous Wave (CW) signals as the emitted sound wave without any encoding scheme. However, we find that CW signals covering a wide frequency band can only capture the energy absorption and speed variation in an ideal environment (e.g., an empty room with fire placed at the LOS path of HearFire). When we change a room with rich multipath, CW signals are very difficult to capture these characteristics. We explain it as follows:

(1) Pure CW sound waves to capture sound absorption and speed entail a wide range of frequency bands. With a broader frequency band, more channel information can be captured in terms of different frequencies, resulting in more effective fire sensing. However, we aim to implement an indoor fire detection system that can be applied in daily usage scenarios. Considering that HearFire requires unceasing emission of acoustic signals to sense fire during working, the emitted signals should be inaudible and do not pose disturbances to people. We notice that sound waves with frequencies higher than 18KHz are almost inaudible to most people [47]. Besides, current widely used acoustic devices support maximum 24KHz sound waves due to their maximum 48KHz sampling rate. Therefore, the permissible frequency range in our system is 6KHz from 18KHz to 24KHz. Such a narrow frequency band impedes CW sound waves from capturing relevant fire patterns.

(2) In practice, fire combustion affects the acoustic channel in a very complicated way due to incomplete combustion of fuel and signal multipath propagation, which varies over time and across environments, especially in rich multipath environments. The received CW signal at microphone is the superposition of multiple copies

propagating along different paths. As such, signal paths affected by fire are submerged and difficult to extract from the superimposed signal due to relatively weak impacts on sound caused by fire.

Therefore, we aim to separate the sound waves that are only impacted by fire from the superimposed signal using an encoded sound wave. The encoded sound signal should satisfy two key goals. First, the auto-correlation of the encoded sound wave should be sufficiently sharp such that sound copies travel through fire with similar absorption patterns as well as close speeds can be separated by finding the peaks corresponding to their different delays. Said differently, the sharper (narrower peak) the auto-correlation of the transmission frame, the more easily we can separate sound copies with close delays, see Fig. 5. Fig. 5 plots a single CIR measurement when sound waves penetrate fire (we will elaborate how to measure CIR in Section 4.2). Each peak represents a signal path with corresponding delay and power. A narrow peak of auto-correlation guarantees that signal paths with close delays can be successfully separated (e.g., path 1, 2, and 3), which can clearly demonstrate the tiny sound speed variation caused by fire. Second, the bandwidth of the encoded sound wave should be less than 6KHz so that it can fully reside in the narrow inaudible band ranging from 18KHz to 24KHz.

In our work, we apply Zadoff-Chu (ZC) sequence for measuring CIR. The root ZC sequence is composed of a set of complex numbers, which can be mathematically expressed as:

$$ZC[n] = e^{-j \frac{\pi u n(n+c_f+2q)}{N_{ZC}}} \quad (2)$$

where  $ZC[n]$  represents the  $n$ -th element of ZC sequence and  $0 \leq n < N_{ZC}$ .  $N_{ZC}$  denotes the total length of the ZC sequence,  $c_f = \{N_{ZC} \bmod 2\}$  and  $q$  is an integer. The parameter  $u$  is between 0 and  $N_{ZC}$ , and the greatest common divisor of  $u$  and  $N_{ZC}$  is 1.

Compared to other mainstream high auto-correlation sequences (i.e., GSM training sequence, Barker sequence, and M-sequence), ZC sequence is proved to have narrower width of auto-correlation main lobe, higher auto-correlation gain, and lower auto-correlation side lobe level [39], which is suitable in our fire detection scenario to separate multipath sound waves with similar delays. In addition, we can modify the ZC sequence to fully fit it into the narrow permissible frequency band. Fig. 6 shows an example of root ZC sequence and its auto-correlation coefficient with  $u = 64$  and  $N_{ZC} = 127$ . The real line and dashed line in Fig. 6 are the I and the Q components of ZC sequence, respectively. These properties ensure accurate measurement of cross-correlation-based CIR estimation.

**4.1.2 Boosting Sensing Range.** In our work, we expand the sensing range of HearFire by building the quantitative relationship between the length of ZC sequence and sensing distance. As CIR is capable of separating multipath signals corresponding to different delays, given the sound speed  $c = 340m/s$  and sampling rate  $f_s = 48KHz$ , the resolution of each tap in distance can be calculated as  $\Delta d = \frac{c}{f_s} = 0.007m$ , which means each tap covers a path length of 7mm. Said differently, any two propagation paths with a length difference of less than 7mm will be assembled in the same tap with a superimposed signal, while being divided into different taps if the path length difference exceeds 7mm. Given the number of taps  $N$ , the maximum path length difference of two sound signals propagating in the environment can be measured by CIR is  $D = N \times \Delta d$ , which indicates a sensing range of  $\frac{D}{2}$ , considering our compact system layout. As the number of taps can be set by determining the length of the transmission frame [26, 47], we can flexibly expand the sensing range by manually setting different lengths of  $N_{ZC}$  to cater to diverse areas. However, a large  $N_{ZC}$  inevitably decreases the time-resolution of CIR since CIR is measured per frame. On the contrary, a short transmission frame enables high time-resolution of CIR while failing to sense sizeable room-scale area. In our study, we set the length of the extended ZC sequence  $N_{ZC}^{extd}$  to 2048<sup>1</sup> to enable room-scale fire detection, which covers a sensing range of  $\frac{N_{ZC}^{extd} \times \Delta d}{2} = 7.25m$ . Besides, a frame

<sup>1</sup>We use the power of 2 to speed up the FFT computation for timely fire detection.

length of 2048 ensures  $23 \sim 24$  CIR measurements per second, which offers enough time-resolution for sensing the fire.

**4.1.3 Satisfying Inaudible Frequency Band.** A sound wave with a frequency higher than  $18KHz$  is inaudible to most people, while most commodity acoustic devices only support a maximum  $24KHz$  acoustic signal with  $48KHz$  sampling rate. As such, our transmission frame should fall into a frequency band ranging from  $18KHz$  to  $24KHz$ , which means our baseband ZC sequence only bears a maximum  $6KHz$  bandwidth. Simply increasing length of the root ZC sequences by setting the parameter  $N_{ZC} = N_{ZC}^{etd}$  fails to create an inaudible sound since root ZC sequences cover a full range of frequency band.

In our study, we extend  $N_{ZC}$  exploiting the interpolation scheme. The purpose of interpolation is to extend the length of ZC sequence and reduce its bandwidth so that it can reside in the inaudible band. We note that interpolation can be achieved in both time and frequency domains. For time domain interpolation, with a sampling rate of  $48KHz$ , we repeat each sample in root ZC sequence eight times, followed by a low-pass filter to reduce the bandwidth to  $6KHz$ . For frequency domain interpolation, we first apply Fast Fourier Transform (FFT) on the root ZC sequence and perform interpolation by symmetrically padding numbers of zeros after positive and before negative frequency components. Then we apply Inverse Fast Fourier Transform (IFFT) to convert the signal back into the time domain. We ensure that the bandwidth of the frequency domain interpolated ZC sequence reduces to  $6KHz$  as well.

Although an interpolation scheme on both time and frequency domain can reduce the bandwidth to  $6KHz$ , the frequency domain interpolated ZC sequence performs better on auto-correlation than the time domain. Specifically, the main lobe width of auto-correlation for frequency domain interpolated ZC sequences is 6 samples, which increases to 16 samples for ZC sequence interpolated in the time domain. Therefore, in our work, we choose to interpolate the ZC sequence in the frequency domain so that it can separate signal paths with closer delays due to a narrower auto-correlation peak.

On the other side, the interpolation scheme narrows the bandwidth  $B = \frac{N_{ZC} \times f_s}{N_{ZC}^{etd}}$ , where  $B$  is the desired signal bandwidth, which should be narrower than  $6KHz$ . As  $N_{ZC}^{etd}$  is fixed to 2048 to cover sizeable areas, we first focus on three pairs of parameters  $(u, N_{ZC})$  of root ZC sequence,  $(32, 63)$ ,  $(64, 127)$ , and  $(128, 255)^2$ , manifesting bandwidths of  $1.5KHz$ ,  $3KHz$ , and  $6KHz$ , respectively. We interpolate these three root NC sequences all to the same length of 2048 and up-convert them to the passband signal with a carrier wave of  $20.25KHz$ . Fig. 7 illustrates the FFT results of these three passband signals. In general, a sufficient frequency bandwidth of transmission frame is desirable for fire detection such that it is able to capture more compelling features of energy absorption. We note that although a frame length of  $N_{ZC} = 255$  supports  $6KHz$  bandwidth, noises generated by imperfect filters during up- and down-conversion will splash out to nearby frequencies, making the transmission frame audible to people. Therefore, after comprehensive consideration, we select  $u = 64$  and  $N_{ZC} = 127$  as parameters of the root ZC sequence, which is then interpolated to 2048 samples to cover  $3KHz$  frequency band ranging from  $18.75KHz$  to  $21.75KHz$ . To this end, HearFire supports inaudible acoustic sensing and boosts the sensing range to 7.25m for indoor fire detection.

**4.1.4 Up-conversion and Down-conversion.** At the transmitter side, we up-convert the time domain baseband ZC sequence to the inaudible frequency band by multiplying the real and imaginary part of ZC sequence by  $\sqrt{2} \cos 2\pi f_c t$  and  $-\sqrt{2} \sin 2\pi f_c t$ , respectively, where  $f_c$  denotes the carrier wave frequency. Then the up-converted real and imaginary parts are added together to form the transmitted passband signal, followed by a bandpass filter to remove the noise outside the transmission band. Note that our transmitted inaudible frame can be generated offline and saved as a playable audio file (e.g., WAV, Mp3, etc.) that can be unceasingly played by most existing commodity acoustic devices.

<sup>2</sup>ZC sequence is periodic with a period of  $N_{ZC}$  ( $ZC[n + N_{ZC}] = ZC[n]$ ) if  $N_{ZC}$  is an odd value [59].

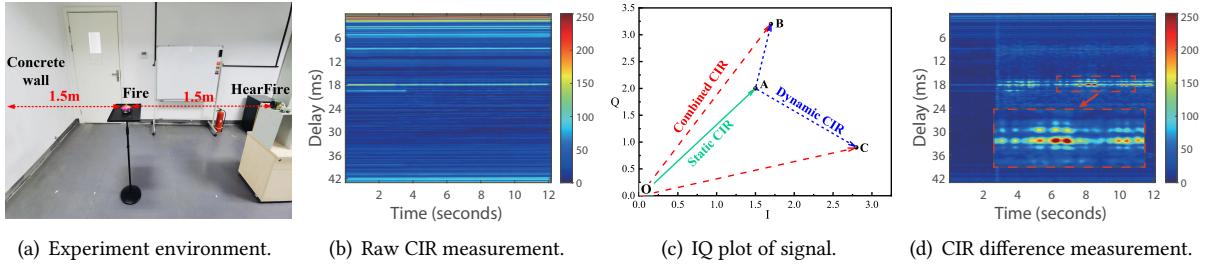


Fig. 8. Fire detection using CIR measurements.

After receiving the acoustic signal, the receiver first locates the starting point of the frame by performing cross-correlation with the passband frame and finding the first peak. As the length of the transmission frame is fixed, the boundary of the rest frames can be easily aligned once the starting point of the first frame is correctly pinpointed. After detecting the first frame, the receiver performs down-conversion using the same carrier frequency. In specific, we multiply the received signal by  $\sqrt{2} \cos 2\pi f_c t$  and  $-\sqrt{2} \sin 2\pi f_c t$  to acquire the real part and the imaginary part of received signal, respectively, which can be further used to measure CIR.

## 4.2 CIR Estimator

**4.2.1 CIR Measurement for Fire Detection.** We apply CIR to measure the channel impacted by flame. We conduct an experiment to show the ability of CIR to capture energy absorption patterns caused by flame. As shown in Fig. 8(a), HearFire is deployed 3 meters away from a concrete wall. The speaker repeatedly emits the inaudible passband frame, which will be received by the microphone. We fix a fire source consisting of ethanol at the center point between HearFire and the wall. During the experiment, we first collect the acoustic data without fire for 3 seconds, after which we ignite the fire source using a long flame lighter and continue to collect sound for 10 seconds. Except for fire combustion, we guarantee that other objects in the environment keep static throughout the experiment. After pinpointing the starting point in the received signal and down-converting the signal to the baseband, CIR can be measured frame by frame by performing cross-correlation.

Fig. 8(b) shows the result of raw CIR measurements. The X-axis is the time duration and Y-axis is the propagation delay. The color represents the signal power. The red lines at the top of CIR heat map represent the sound propagation paths directly from the speaker to the microphone. Since we collocate the speaker and microphone at a close distance, these portions of LOS multipath signals transmit in the shortest delays and experience minimal attenuation with almost constant power. Except for high-energy LOS paths, we observe multiple light-colored lines throughout the propagation delays, indicating multipath signals propagating along diverse paths after being reflected from different objects in the environment. Signal paths with the second highest energy locate around 18ms delays, representing an exact 3 meter's propagation distance between HearFire and the wall. The energy of these paths keeps constant before we ignite the fire, while suddenly decays after 3rd second and varies randomly, demonstrating an energy absorption by fire.

**4.2.2 Remove the Effects of Static Objects.** The measured raw CIR is the superposition of multipath signals, containing both static and dynamic components. Unlike previous works that apply the signal reflection effect, in our work, dynamic components involve signals whose energy is partially absorbed when propagating near the flame, resulting in signal strength variations. However, the static component consists of multipath signals that are not impacted by the fire, which are superimposed on the receiver with constant energy. Thus, we can still cancel the static component by calculating CIR difference  $CIR_{diff}$  between two consecutive frames, preserving the information on CIR change caused by the dynamic component. As illustrated in Fig. 8(c),  $\overrightarrow{OB}$  and  $\overrightarrow{OC}$  represent CIRs of two consecutive frames, which share a mutual static component  $\overrightarrow{OA}$  while with different

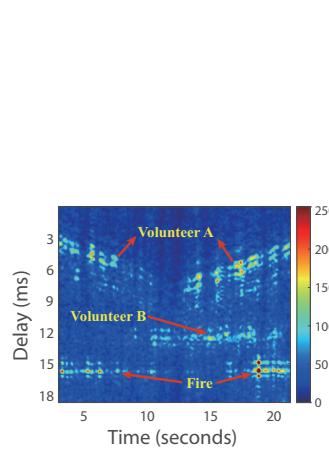


Fig. 9. CIR measurement in a more dynamic environment.

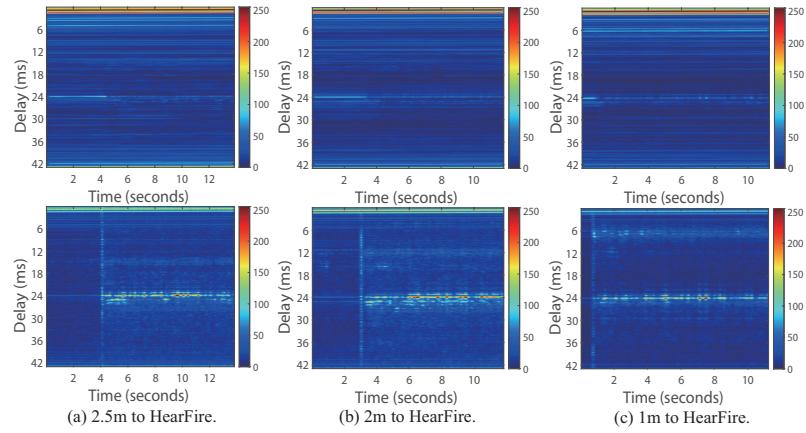


Fig. 10. Fire at different locations share similar effects on CIR measurements.

dynamic components  $\vec{AB}$  and  $\vec{AC}$ . By performing  $\vec{OB} - \vec{OC}$ , the static component  $\vec{OA}$  is removed, remaining  $\vec{AB} - \vec{AC}$ , which describes the energy variation of the signal.

Fig. 8(d) shows the CIR result in Fig. 8(b) after removing the static components. In Fig. 8(d), we have three key observations. First, the strong LOS path and other static reflections have been successfully cancelled out due to their constant signal power. Second, when we zoom-in, we observe that the acoustic signals with a propagation delay of around 18ms (i.e., 6m round trip distance) exhibit similar patterns across different delays. This is because 1) sound travels faster within the area impacted by flame due to higher temperature and 2) air temperature decreases with the increasing distance to the fire. Thus, sound penetrating this high-temperature area will create multiple copies with shorter propagation delays yet similar energy loss of the acoustic signal. Third, all sound signals with different delays exhibit random variations of signal strength after igniting the fire, manifesting a clear signal absorption pattern. This is because incomplete combustion of fire incurs unstable energy absorption of the acoustic signal, resulting in random residual energy of sound after penetrating the flame. Such a series of unique patterns captured by CIR offer the potential for fire detection using acoustic signals.

To further validate that using differences of CIR measurements can effectively remove static interferences and preserve the variation of the channel, we conduct an experiment in a more dynamic environment by asking three volunteers to move simultaneously. In specific, volunteer A is allowed to walk back and forth in a straight line parallel to HearFire, while volunteer B walks between fire source and wall, and volunteer C taps a keyboard 3 meters away from HearFire. During the experiment, we keep the fire ignited.

The corresponding  $CIR_{diff}$  measurement is shown in Fig. 9. Volunteer A moving away and approaching HearFire can be clearly captured by CIR in the propagation delay ranging from 3ms to 9ms. We can also observe the pattern in the delay between 12ms and 18ms, which shows that the signal passing through the flame is first reflected by the wall (before 10s), then by the moving volunteer B (10-18s), and finally again by the wall (after 18s). We notice that volunteer C tapping the keyboard fails to be captured by CIR due to limited impacts on sound waves. Our experiment result validates that using CIR difference to remove the effects of static components of the signal can hold up in a more dynamic environment. The reason relies on two folds: (1) narrow auto-correlation peak of ZC sequence can separate signal paths with close delays. Hence, dynamic and static paths can be accurately isolated in terms of different delays; (2) CIR difference measures the change of the channel. Static components of the signal are eliminated, remaining effective dynamic paths related to channel variations.

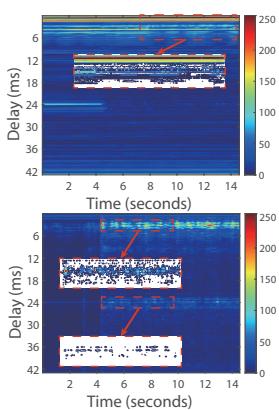


Fig. 11. Delay &lt; 6ms.

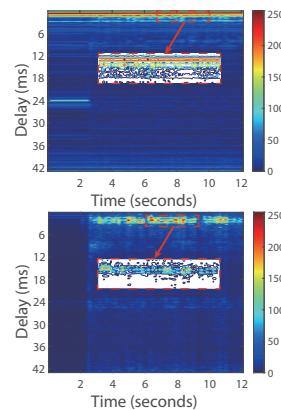


Fig. 12. Delay &lt; 3ms.



(a) Smoke detection setup.



(b) Smoke in the early stage of fire.

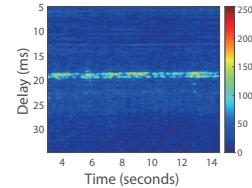
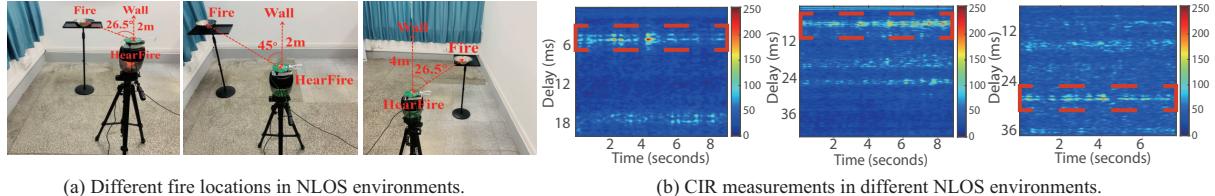


Fig. 13. Smoke detection in the early stage of fire.

**4.2.3 Validate Energy Absorption in CIR Measurement.** To validate sound energy absorption caused by fire, we first move the fire source to random positions along the LOS path between HearFire and wall and measure the CIR after igniting the fire at a random time point. Interestingly, all CIR measurements of each experiment exhibit similar patterns. When the fire resource locates at different distances to the HearFire, the acoustic signals keep arising after igniting the fire while appearing at the same position with a propagation delay of 24ms, as plotted in Fig. 10. Similarly, the energy absorption patterns remain at the same propagation range in CIR after removing static components. Such a pattern demonstrates that the position of flame imposes no impact on the signal propagation delay, proving that fire is incapable of reflecting any sound. As a result, CIR measurement at 24ms delay only captures the signals propagating along the LOS path between HearFire and the wall with a fixed delay of 24ms, even though involving the penetration of fire.

However, when we move the fire close to the HearFire ( $\leq 0.5\text{m}$  in our experiment), we observe different CIR patterns, as illustrated in Fig. 11. Except for CIR capturing the LOS path between HearFire and the wall at a 24ms delay, signal energy exhibits a clear unstable pattern covering multiple paths with shorter propagation delays (i.e., around 5ms). When we zoom in the CIR, we can observe a continuous pattern in raw CIR due to stable reflections by the container. However,  $CIR_{diff}$  measurements exhibit a clear unstable pattern compared to raw CIR. This is because the fire directly affects the surrounding area of HearFire at this delay, absorbing the energy of sound reflecting from the fire's container. This effect becomes more distinct if we continue to move the fire towards HearFire (e.g.,  $\leq 0.2\text{m}$ ). As shown in Fig. 12, even the LOS path between speaker and microphone is impacted by the flame, which can be directly observed in the raw CIR measurement. Moreover,  $CIR_{diff}$  exhibits various delays in both Fig. 11 and Fig. 12 while with similar energy absorption patterns due to different temperatures. We notice that CIR measurements at 24ms delay almost disappear in both raw CIR and  $CIR_{diff}$  since the environment nearby HearFire dominates the CIR change, resulting in more energy absorption and more drastic variations in signal strength. Our experiment confirms that sound energy is indeed absorbed by fire, and sound speed varies with temperature, which can be captured by CIR.

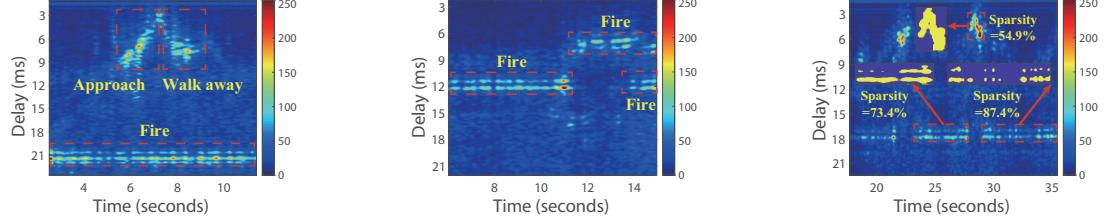
Being able to detect fires at an early stage is critical for fire rescue. In the early stage of fire, the fuel generates smoke due to incomplete combustion. Therefore, we conduct an experiment to test whether HearFire can detect fire in the early stage (e.g., only smoke without flame). As shown in Fig. 13(b), we use a smoldering cotton wick as the fuel and place leaves and flakeboard inside a container to generate smoke. In our experiment, HearFire with an omnidirectional speaker is deployed close to the wall, and the smoke is randomly placed in the room, as illustrated in Fig. 13(a). The speaker repeatedly emits the inaudible passband frame, which will be received by the microphone.



(a) Different fire locations in NLOS environments.

(b) CIR measurements in different NLOS environments.

Fig. 14. NLOS environment settings and corresponding CIR measurements.



(a) Walking parallel to the LOS path between HearFire and wall.

(b) Traversing the LOS path between fire and wall.

Fig. 15. Extract effective CIR measurements.

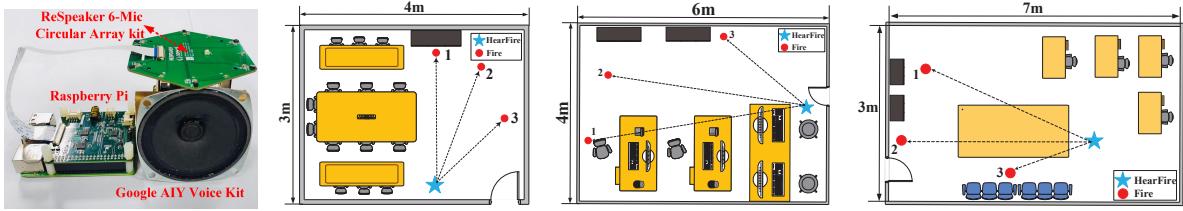
The experiment result is depicted in Fig. 13(c). We observe that the acoustic signals with a propagation delay of around 20 ms penetrate the fire and exhibit obvious signal absorption and speed variation patterns, which shows that HearFire is capable of detecting smoke as well in the early stage of fire. This is because the fundamentality of our HearFire to detect fire is to sense the emitted heat nearby fire sources. As smoke in the early stage of fire also generates heat and significantly increases the temperature of nearby air, the energy absorption and varying speed of sound can be captured by CIR measurements.

Currently, we apply a directional speaker and place the fire source at the LOS path of HearFire to ensure that sound waves can travel through fire. However, the fire location is not guaranteed in LOS in practice. Therefore, we further explore the capability of HearFire when the location of the fire source is not in LOS. In this study, we replace the directional speaker with an omnidirectional speaker and test HearFire when fire sources are located in NLOS scenarios. In experiments, we fix the location of HearFire and place the fire source at three different locations in a room with a size of 3m\*7m\*3.5m, as shown in Fig. 14(a). The measured CIR for each scenario is shown in Fig. 14(b). From Fig. 14(b), we observe that the fire at different locations is successfully captured by CIR, which all exhibit clear patterns of sound absorption and varying propagation speeds, i.e., CIR patterns in the red rectangle. This is because the omnidirectional speaker transmits sound in all directions, which increases the probability of sound waves penetrating the flame and being received by the microphone. Note that we use different delay ranges (Y-axis) to better exhibit the CIR patterns caused by fire.

### 4.3 Fire Detector

Our fire detector design aims to accurately detect fire from extracted CIR measurements. We take CIR after removing static components as the input of our fire detector. In specific, we measure the raw CIR of two consecutive frames  $CIR_{t-1}$  and  $CIR_t$ , and calculate their differences  $CIR_{diff} = \|CIR_t - CIR_{t-1}\|$ , which results in a 2048-point real vector describing the energy of CIR change, where  $\|\cdot\|$  represents the 2-norm operator.

**4.3.1 Extract Effective CIR Measurements.** Unfortunately, other moving objects in the environment will cause multipath reflections of acoustic signals, which greatly affects the extracted CIR measurement. In this experiment, we vary the distance between HearFire and wall to better display the effects caused by moving objects. After igniting the fire, we first ask a volunteer to walk backward and forward following a straight line parallel to the



(a) Speaker and microphone.

(b) Room 1.

(c) Room 2.

(d) Room 3.

Fig. 16. Commercial speaker and microphone applied in HearFire and three different rooms.

LOS path between HearFire and wall. Fig. 15(a) shows the  $CIR_{diff}$  result. In addition to the signal absorption by fire, we can observe a clear reflection pattern caused by human body, indicating a walking toward followed by leaving away from HearFire. Surprisingly, the CIRs caused by fire absorption and human walking can be divided on the basis of propagation delay. In addition, we notice that signals reflected from moving objects generally cover a wider range of propagation delays than that of fire and exhibit a more continuous pattern. This is because human body, as a reflector, results in more stable reflections, whose propagation delay changes successively.

We further investigate the impact of moving objects by allowing a volunteer to traverse the LOS path between HearFire and the wall. After igniting the fire, the volunteer first approaches flame, halting for several seconds in the middle of the wall and fire, and continues to walk away from fire. Interestingly, CIR measurements at 2 meters suddenly disappear when the volunteer is halting, which generates unstable yet similar CIR measurements with shorter propagation delays, as shown in Fig. 15(b). This is because sounds are reflected by human body rather than the wall, therefore, shortening the propagation delays of signals penetrating the fire. We finally ask a volunteer to traverse between HearFire and fire two times. At this time, flame is occluded by the body, incurring the vanishing of CIR measurements caused by fire while at the same time creating stronger CIR measurements close to HearFire, as shown in Fig. 15(c). However, different from the CIR pattern in Fig. 15(b), due to sound reflections from human body, CIR measurements exhibit a stable pattern across successive delays in Fig. 15(c).

#### 4.4 Fire Detection

Based on those observations, we design a two-step scheme for fire detection. First, instead of applying a machine learning framework, we apply a simple yet effective method to remove reflection interferences by measuring the sparsity. In specific, we pick a fixed number of consecutive  $CIR_{diff}$  traces and locate the area of the CIR measurements whose values exceed a pre-defined threshold. Then we measure the sparsity for each area as the number of zero-valued taps divided by the total number of taps in the area. Our study empirically set the sparsity to 60% to distinguish between fire and other moving objects. The areas with sparsity lower than 60% are determined as interferences (e.g., human motion), and we manually set the value of all grids in the areas to a very small number to remove the interferences. We note that this threshold of sparsity can be flexibly configured within a certain range due to a clear difference between energy absorption and signal reflection patterns in CIR.

HearFire then inputs these interference-removed CIR traces one by one into a model pre-trained by Support Vector Machine (SVM). The SVM model will classify each 2048-point CIR trace into two categories: fire and non-fire. Finally, HearFire takes multiple inferences and applies the following engineering strategy: HearFire issues a fire alarm if and only if when at least  $k$  of  $n$  successive measurements report “fire”, where  $k \in [1, 2, 3, \dots, n]$ . Otherwise, HearFire will determine that there is no fire in the monitored area and keep silent.

## 5 EVALUATION

### 5.1 Experiment setup

**5.1.1 Transmission Frame.** To capture the impact of fire on acoustic signal and measure CIR, a transmission frame in Section 4.1.2 is applied to satisfy the inaudible frequency band and support room-scale fire detection.

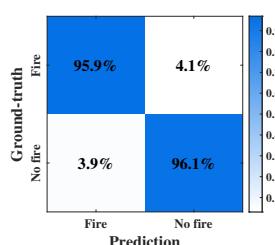


Fig. 17. Overall performance for each measurement.

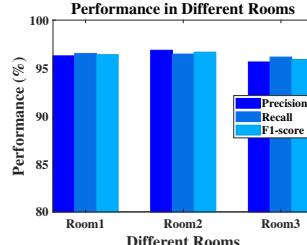


Fig. 18. System performance for each room.

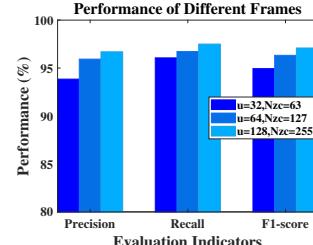


Fig. 19. Performance on different frames.

Under these parameter settings, a frame takes 42.6ms in transmission with a sampling rate of 48KHz, causing around 23~24 frames received per second, which provides sufficient time-resolution for CIR measurements.

**5.1.2 Hardware.** In our evaluation, we apply a commodity speaker and microphone pair as our transmitter and receiver, as shown in Fig. 16(a). For the speaker, we use google AIY Voice Kit version 2.0, which contains a fully programmable speaker supporting a 48KHz sampling rate. The speaker is connected to a built-in Voice Bonnet with a special audio codec processor powered and controlled by a Raspberry Pi Zero board with 2.1 Amps of power. All of these connected components cost \$30 with small size, which is convenient and easy to deploy by users. We program the speaker to unceasingly transmit our pre-defined inaudible transmission frame. We tune the voice to 80% of its maximum volume to avoid being audible to people in quiet environments when they really pay attention. Users can flexibly tune the volume to guarantee that the frame is inaudible to them while not much affecting the system performance. For the microphone, we choose a low-cost ReSpeaker 6-Mic Circular Array kit to receive the frame. The microphone embedded in ReSpeaker kit supports a 48KHz sampling rate, enabling signal reception at an inaudible frequency band (18KHz ~ 24KHz). This microphone kit is powered and can be controlled by a Raspberry Pi as well, which is fully programmable and designed for AI and voice applications. Note that we only open one microphone as our fire detection system does not require information from multiple microphones. Throughout the experiment, we place the speaker and microphone close enough to mitigate the impacts from LOS propagation.

**5.1.3 Software.** We use MATLAB to generate the transmission frame, which is saved as an audio file in a Raspberry Pi Zero connected to the speaker. The Simpleaudio and PyAudio packages are imported in Python to control the speaker's volume and play this audio file repeatedly. The signal is then received and stored in Raspberry Pi, which is further processed using Python to measure the CIR.

**5.1.4 Data Collection and Model Training.** We collect the data in three different rooms with sizes of room1:  $3m \times 4m$ , room2:  $4m \times 6m$ , and room3:  $7m \times 7m$ , respectively, as illustrated in Fig. 16(b)-16(d). We consider three types of fuel: ethanol, paper and charcoal to generate fire, representing different combustion levels. We randomly choose different fuels and place the fire source at different locations to collect CIR. In each room, we equally collect 3000  $CIR_{diff}$  traces for fire scenario and non-fire scenario to avoid unbalanced training. For each scenario, 1500 traces are used for training and 1500 traces for testing. Therefore, we have 9000 training traces and 9000 testing traces in total. We choose offline training by using a Laptop with an Intel(R) Core(TM) i5-6300U CPU, 20GB ram and a integrated graphics. The trained model is then saved in Raspberry Pi for online testing. In our evaluation, we take 10-fold cross-validation.

## 5.2 Overall Performance

We evaluate the overall system performance of HearFire in two stages. In the first stage, we evaluate the performance of HearFire to deal with a single measurement (i.e., one CIR per ZC frame), as depicted in Fig. 17. We mix the traces collected in all three rooms and take 50% and 50% for training and testing, respectively. We

Table 1. True Positive and True Negative Rate under Real Scenarios.

	n=1	n=2	n=3	n=4	n=5	n=6
TP Rate	0.959	0.964	0.972	0.988	0.992	0.996
TN Rate	0.961	0.968	0.970	0.984	0.993	0.998

Table 2. Performance on Different Classifiers.

Classifier	SVM	BP	KNN	RF	NB
Average Performance	96.1%	96.00%	95.5%	83.9%	61.3%
Best Performance	97.0%	97.1%	96.5%	93.6%	71.0%
Worst Performance	94.8%	94.4%	93.9%	71.6%	56.0%

apply SVM as the classifier since its performance exceeds other classifiers (will be discussed in Section 5.3). In Fig. 17, the ground truth is presented in each column, while each row shows the results of each measurement detected by HearFire. In this evaluation, we calculate the Accuracy:  $\frac{TP+TN}{TP+FP+TN+FN}$ , Precision:  $\frac{TP}{TP+FP}$ , and F1\_score:  $\frac{2 \times Precision \times Recall}{Recall + Precision}$  of the detection result for each measurement in the fusion matrix, where TP,TN, FP and FN represent the true positive, true negative, false positive and false negative, respectively.

For each measurement, HearFire achieves an average accuracy of 96% for all three rooms, with both TP and TN exceeding 95%. When we separate the measurement performance based on each room (i.e., model that is trained and tested using data collected in the same room), the detection results for each measurement, including Precision, Recall, and F1\_score, reach over 95% for all three rooms, as shown in Fig. 18, indicating that CIR measurements of acoustic signal indeed have the ability to capture fire and are robust for different environments. This is due to unique sound propagation characteristics that we exploit, which can be captured by CIR measurements. Moreover, our interference cancellation scheme filters out the effects caused by other fire-irrelevant objects.

In the second stage, we evaluate the overall performance of our HearFire system. Recall that we apply multiple measurements to increase the detection accuracy for HearFire. We apply the detection strategy in Section 4.4 to test HearFire in real scenarios in different rooms, fire positions, and under interferences. We notice that the configuration of n and k will significantly affect the system performance. On the one hand, a large n improves the system accuracy while degrading the system's real-time performance. On the other hand, an appropriate k results in an acceptable FN rate and the FP rate. In our study, we take  $n \in [1,2,3,4,5,6]$  and different k, and measure the TP and TN rate in the experiment as shown in Table. 1 (we vary k under each n and present the best performance). Both TP and TN rates improve as the number of combined inferences increases. Finally, we empirically set  $n=5$  and  $k=3$ , which balances the system's real-time performance and achieves a TP and TN rate both exceeding 99%, resulting in an overall fire detection accuracy of 99.2% for the HearFire system.

We note that the system accuracy highly relies on the accuracy of each measurement. Therefore, we use the measurement accuracy to show the performance of HearFire in the following evaluation.

### 5.3 Performance on Different Classifiers

In this evaluation, we select 6 mainstream classifiers to test the system performance, including SVM, Back Propagation (BP) Neural Network, K-Nearest Neighbors (KNN), Random Forest (RF) and Naive Bayes (NB). All the classifiers are implemented using MATLAB with their default parameter settings. For example, the default C value in SVM is 1, and the default maximal number of decision splits and minimum number of leaf node observations in DT are set to 1 and 2, respectively. For BP network, we achieve the best performance by configuring a 2-layer structure with 20 and 10 neurons in each layer, followed by a Tanh activation function with a learning rate of  $1e^{-4}$ . In KNN, we set  $k = 7$  to achieve the best performance. We randomly choose fuels and place the fire source in room3 to collect 3000 CIR traces with and without fire, respectively, in which 1500 traces are used for training and 1500 for evaluation<sup>3</sup>. We also calculate the Accuracy, Precision, Recall and F1\_Score for each detection.

Table. 2 illustrates the system performance when using different classifiers. Except for NB, all other four mainstream classifiers achieve a performance of over 80% detection accuracy. SVM achieves the best performance

<sup>3</sup>Unless otherwise mentioned, the rest of evaluations are conducted in room3 and all apply the same data collection and split strategy.

Table 3. Detection Time of HearFire.

CIR Calculation (ms)			Fire Detection (ms)				
Frame Detection	Down Conversion	CIR	SVM	BP	KNN	RF	NB
310	30	44	1.35	2.62	2.60	1.72	1.81

Table 4. A Comparison between Wi-Fire and HearFire.

Fire detection	Detection Accuracy		Overall Detection under 26.5° Accuracy	Detection Speed
	Wi-Fire	HearFire		
Wi-Fire	95.5%	98.8%	87.9%	<1.0s
HearFire			96.4%	<0.7s
			99.2%	

in these classifiers with an average accuracy of 96.1%, which is 12.2% and 34.8% higher than that obtained by RF and NB, respectively. This is because SVM is able to create the best hyperplane, representing the largest separation between two classes, which is suitable for binary classification. We notice that NB has the worst performance because NB incurs zero probability problem, requiring the testing data to be very similar to the training data. Although BP and KNN achieve similar performance to SVM, they take a longer time to process a single CIR trace than that of SVM (will be shown in Section 5.4). Therefore, we finally select SVM as our classifier for fire detection.

#### 5.4 Detection Time

The result of detection time is summarized in Table. 3. To measure the detection time, we track the processing delay for each step after receiving the frame and calculate the total time for issuing a detection result. For this evaluation, we first measure the processing time for locating the starting point in the first frame of the received signal. In this stage, cross-correlation coefficients are calculated with a sliding window with a size of 2048 and step length 1. We measure this step 100 times, and HearFire takes an average of 310ms to detect the first frame. This frame detection step can be further shrunk with strict time synchronization between speaker and microphone. Note that this frame synchronization is a one-time step at the beginning of the detection, and the rest frame can be aligned with a fixed frame length.

Then we take all 3000 testing CIR traces and measure the processing delay of down-conversion and CIR extraction step. In HearFire, each received frame is first down-converted by multiplying a carrier wave, followed by CIR extraction using cross-correlation. In our experiment, HearFire takes an average of 30ms and 44ms to perform down-conversion and extract CIR, respectively. This is because the down-conversion step involves two 2048-point dot products (i.e.,  $\cos$  and  $-\sin$ ), and CIR measurements extraction includes convolution calculation (i.e., cross-correlation), resulting in a longer processing time. Note that as we employ a simple yet effective matrix sparsity method to mitigate the interferences, the processing delay of this step is magnitudes level shorter than other steps, which can be neglected in our evaluation.

Finally, we evaluate the detection time for all five mainstream classifiers to test a single 2048-point CIR trace. BP has the longest processing time because of its 2-layer model, which bears more parameters and incurs the highest computational overhead. SVM achieves the fast response and only requires 1.35ms to test a CIR trace. Since HearFire extracts CIR traces for 5 consecutive frames and issues final detection results, which takes around 0.25 seconds to transmit by a transmitter. As a result, HearFire takes an average of less than 0.7 seconds in total for detecting fire when a flame appears, which is sufficient for an indoor fire detection system.

#### 5.5 Comparison with the State-of-the-art Fire Detection Approach

To compare with the state-of-the-art fire detection methods, we take Wi-Fire [62] as the benchmark that utilizes Wi-Fi signals to detect fire, which is a contact-free fire detection system using wireless signals. Wi-Fire leverages the fact that fire affects the RF signal propagation, which can be captured by Channel State Information (CSI) to achieve room-scale fire detection. In the experiment, we test the system accuracy of HearFire using an omnidirectional speaker under the same experiment settings as Wi-Fire. Specifically, we place the fire in two particular directions of  $26.5^\circ$  and  $45^\circ$  to the LOS between HearFire and wall. For each direction, we place the fire source 2m and 4m to the HearFire device, respectively. We list the result in Table. 4.

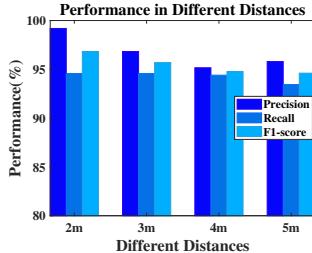


Fig. 20. Different distances between HearFire and fire.

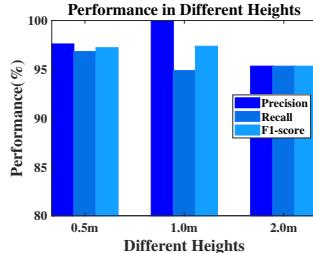


Fig. 21. Different heights of fire.

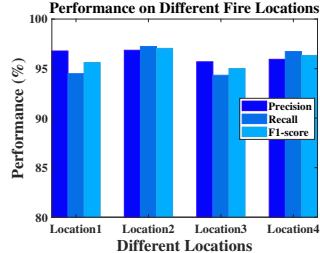


Fig. 22. Different fire locations.

The overall detection accuracy of HearFire system achieves 99.2%, which is 2.6% higher than Wi-Fire. Moreover, HearFire performs better when fire source is located in both directions, which can achieve an accuracy of 98.8% and 96.4% for fire detection at 26.5° and 45°, respectively. We notice that the accuracy at 45° is lower than that of 26.5° because of more deviation to the LOS, resulting in more weak signals involving fire impacts. One may use the beamforming technique to enhance the fire detection performance. Our system can detect fire within 0.7 seconds, which is 30% faster than Wi-Fire, offering more urgent alarm for fire rescue, especially considering the high timeliness of fire detection.

## 5.6 Performance on Different Transmission Frames

As we discussed in Section 4.1.2, the configuration of parameters for root ZC sequence will pose impacts on the system performance. We evaluate three transmission frames with different parameter pairs of  $(u, N_{ZC})$ : (32, 63), (64, 127), and (128, 255), respectively. We interpolate all three sequences to the same length of 2048. During experiments, we keep all other settings the same while only changing the parameters of the root ZC sequence. Fig. 19 shows the system performance for different transmission frames. As the length of root ZC sequences increases, HearFire achieves better performance since the transmission frame covers a wider frequency band, involving more information on energy absorption patterns. However, although achieving the best performance, transmission frame with  $u = 128$  and  $N_{ZC} = 255$  is indeed audible to people, disturbing them with squeals throughout our experiment. This is because frequencies leak into the audible frequency band below 18KHz due to imperfect frequency responses of filters. Therefore, we eventually select the root ZC sequence with  $u = 64$  and  $N_{ZC} = 127$  to implement HearFire, which covers a frequency range of 3KHz and is fully located inside the inaudible frequency band with sufficient interval to the boundary.

## 5.7 Performance on Different Fire Locations

In this evaluation, we first fix the location of HearFire and place the fire source 2, 3, 4, and 5 meters away from HearFire while at the same height, respectively. Then we fix the distance between fire and HearFire to 4m while with different heights of the fire. Fig. 20 illustrates the evaluation results for different fire locations. As fire is unable to reflect any sound, varying locations between fire and HearFire does not impact the LOS path between HearFire and the wall, imposing almost no effect on CIR measurements. Furthermore, our interference cancellation scheme effectively removes the impact of other objects, resulting in a clear CIR pattern merely caused by fire. Thus, the system performance remains at the same level, as shown in Fig. 20. Similarly, different heights of fire source give rise to limited impacts on system performance as well since fire and its affected surroundings are within the sensing range of HearFire, which can be detected using acoustic signals, as shown in Fig. 21.

To further test HearFire's effectiveness across the room scale, we randomly place the fire source in three different locations at different distances and angles in room3, as shown in Fig. 16(d). For each fire location, we collect 1000 CIR traces for fire and non-fire. We use half of them for training and testing, respectively. The performance is shown in Fig. 22. The precision, recall and F1\_score for all four locations all exceed 94% for each

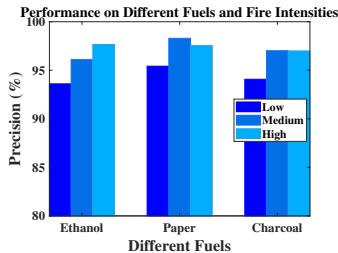


Fig. 23. Performance on different fuels and fire intensities.

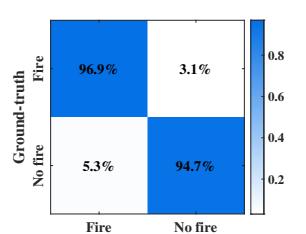


Fig. 24. Performance under interferences.

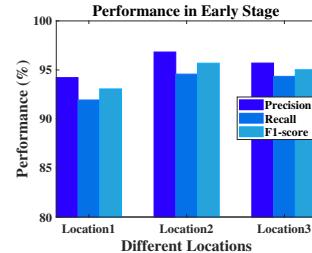


Fig. 25. Performance in early stage.

CIR measurement, demonstrating that our HearFire enables room-scale fire detection when fire is at different locations. This is because HearFire can capture sound absorption and speed variation while the omnidirectional speaker increases the probability of sound traveling through fire.

### 5.8 Performance on Different Fuels and Fire Intensities

We evaluate whether different fuels influence HearFire’s performance. Different types of fuel cause diverse combustion levels due to different chemical reactions. In this evaluation, we fix the location of the fire source while using three types of fuel: ethanol, butane, and paper. To generate different fire intensities, we tune the opening size of the container (small, medium, and large sizes) to control the contact area between the flame and the air. For each fuel in the container with each opening size, we collect 1000 CIR traces with and without fire, and use 50% and 50% for training and testing. Fig. 23 plots the system performance results. The system performance rises with increasing container’s opening size due to the higher intensity of fire, affecting a more extensive area near the fire. As such, CIR’s absorption of sound energy can be more easily captured, thus improving the system performance. However, different fuels pose limited effects on system performance since combustion for each fuel remarkably elevates air temperature and causes significant energy absorption of sound.

### 5.9 Performance under Interferences

Multipath caused by other moving objects in the environment can be captured by CIR and affects the system performance. In this evaluation, the distance between fire and HearFire is fixed at 5m. A volunteer is allowed to randomly walk with large torso movement in room3 at the same time when igniting the fire. For small movements, we ask a volunteer sitting 1 meter away from fire and perform gestures by raising and dropping his arm. 1000 CIR traces are collected for both large and small movements. We then perform the sparsity scheme to identify the interference caused by body movement and CIR caused by fire. Fig. 24 plots the confusion matrix of detection results. Our sparsity scheme can effectively differentiate CIR patterns of signal absorption from signal reflection with an accuracy of 95.8%. We note that 3.1% of CIR caused by fire is classified into human movement, and 5.3% of CIR caused by human movement is classified into fire. This is because the volunteer is allowed to traverse the LOS between fire and HearFire during walking, resulting in errors when human torso blocks the fire.

### 5.10 Performance on Detecting Fire in the Early Stage

In this evaluation, we place smoke at different locations in room3 (i.e., red dots) and collect 1000 CIR traces with and without smoke, and use 50% and 50% for training and testing, respectively. The performance on detecting fire in the early stage is shown in Fig. 25. The SVM classifier result shows that the measurement accuracy when smoke is placed at different locations reaches an average of 94.4%, with precision, recall and F1\_score all exceeding 92%, which validates that HearFire can detect fire even in the early stage. This is because the fundamentality of HearFire is to capture the sound pattern impacted by the emitted heat nearby fire sources.

### 5.11 System Power Consumption

Our HearFire system consists of two Raspberry Pi systems, a high gain speaker and a 6-mic array. To obtain the power consumption of such an integration of multiple modules, we first refer to the official power consumption specifications for each component of HearFire. The theoretical total power of HearFire is 28.35W, which is less than 0.03KW·h for an hour of use.

To test the real power consumption of HearFire in practice, we apply a Watt Meter. Specifically, we open the HearFire to continuously monitor a 7m\*3m\*3.5m room for two hours and randomly ignite a fire source several times. The power consumption of HearFire for 2 hours of continuous operation is 0.018kw·h, which is less than the theoretical one as the official specification is the rated power. The Raspberry Pi 4B can be replaced with a more lightweight Raspberry Pi Zero WH, further reducing power consumption while maintaining the system's function. We note that although the energy consumption is higher than some traditional methods, HearFire's power consumption is still acceptable in most indoor usage scenarios such as home, office, and warehouse, especially when being connected to a constant power supply.

## 6 LIMITATIONS

**Limit of measurement.** In our study, we use three types of fuel, which all generate flame during combustion and impact the air nearby. To test the limit of our system, we replace our fire source with a lighted cigarette. The top of the cigarette has only sparks but no flames. Our experiment result shows that HearFire fails to detect the lighted cigarette no matter where we place it, even in the smallest room<sup>1</sup>. At the current stage, HearFire is unable to sense sparks since their impacts on air are too minute to capture by CIR, resulting in almost constant CIR in different propagation delays. This is because of the time/space resolution limitations of our current commodity acoustic devices. Most commodity acoustic devices support a maximum sampling rate of  $f_s=48KHz$ , which introduces a time-resolution and space-resolution of  $\Delta t = \frac{1}{f_s} = 2.083 \times 10^{-5}s$  and  $\Delta d = \frac{c}{f_s} = 0.007 m$ , respectively. As such, the heat emitted by a single cigarette is too low to be effectively detected by HearFire. One possible solution to expand the range and time/space resolution is to improve the resolution by increasing the sampling rate of the device, which, however, will increase the system cost.

**Limit of sensing range.** In this work, we apply acoustic signals to sense the fire. Considering the inherent feature of the acoustic wave, it is limited in long-distance sensing. Another issue that limits the sensing distance is that we apply a 2048-point ZC sequence, which corresponds to a 7.25m sensing distance. Intuitively we can enlarge the length of ZC sequence to achieve a longer sensing range. For example, a 4096-point and 8192-point frame support 15.5 m and 31m sensing distance, respectively. However, a more extended ZC sequence inevitably reduces the time-resolution of CIR since CIR is measured per frame. One possible solution to realize a larger indoor or outdoor fire detection is to apply a speaker with a higher transmission power, which, however, increases the system costs. Another method to enlarge the sensing area is to increase the sampling rate using specialized high-end acoustic devices. However, those specialized acoustic devices may not be supported by most smart devices, which constrains their wide-scale use.

**Limit of background audio in similar frequency bands.** Our current version of HearFire can work when there is no background audio in similar frequency bands. This is because the background noise with a similar frequency to the working frequency of HearFire might be difficult to separate and filter out, resulting in invalid fire detection. In our current work, HearFire can accurately detect fire in daily indoor scenarios, where most of the background noises only reside in a relatively lower frequency band (e.g., below 8KHz) [19]. We notice that signal denoising algorithms might be applied to remove the background audio in a similar frequency band. Also, the high auto-correlation property of emitted ZC sequence may be leveraged to mitigate the background noises. We will study it in our future work to expand the application scenarios of HearFire.

**Detection of multiple fire sources.** Our experiment only detects a single fire source. We notice that multiple flames appearing at different locations influence sound signals simultaneously, which can be divided by CIR

if their propagation delays differ significantly. However, signals affected by fire but with similar propagation delays are difficult to separate as they superimpose together in CIR measurements, causing entirely different CIR patterns. We plan to apply beamforming technique such that signals with similar propagation delays can be differentiated using different angle-of-arrivals (AoA). Moreover, having AoA may potentially implement fire source localization, adding a new function to our HearFire system.

**Fire detection behind wall.** Due to the inherent characteristics of sound, acoustic devices have difficulty in sensing fire behind walls since wall completely blocks the LOS of HearFire. We note that our HearFire system is not to replace existing fire detection techniques (e.g., smoke detection and RF sensing) but to provide a complementary technique that could cover most usage scenarios in indoor environments. For behind wall fire detection, although our HearFire and other smoker detectors may fail to detect fire, we may apply Wi-Fi sensing technology as Wi-Fi signal can penetrate wall. The development of HearFire facilitates that one may flexibly choose appropriate sensing technologies to cater to most daily usage scenarios.

**Detection with occluded LOS path.** In Section 5.9, we have evaluated our system performance when a user moves nearby the fire. The detection errors primarily derive from the complete blockage of LOS path between HearFire and the flame by the human body. In this scenario, HearFire fails to detect fire. We further conduct an experiment using a plastic plate to completely block the LOS of HearFire and obtain similar results, even using an omnidirectional speaker. This is because the sound waves traveling through fire are too weak to capture by CIR measurements. In practice, we notice that fire may not be easily observed by people even if they are in the same area. The current version of HearFire will detect the blockage scenario as other dynamic interferences and simply drops this detection result since the occlusion of LOS path merely lasts a very short period (< 1 second). As such, HearFire can continue to sense the fire if LOS path between HearFire and fire reappears. We plan to address this practical challenge in the future. One possible solution is to employ multiple HearFire devices at different positions in the same room to fully cover 360-degree sensing and reduce the probability of being occluded by objects. Another approach is to apply the beamforming technique using received mic-array to enhance the signal in a particular angle involving fire patterns. We plan to address this practical challenge in the future.

**System sustainability.** We envision our HearFire can be pervasively applied in many indoor scenarios, especially where constant power supplies can be connected, such as the living rooms, offices, corridors, warehouses, etc. However, we notice that when the constant power supplies are not guaranteed, the energy consumption of HearFire is relatively higher than traditional approaches. To further reduce power consumption, the Raspberry Pi 4B can be replaced with a more lightweight Raspberry Pi Zero WH or other lower power control devices.

On the other hand, the system energy may not be effectively used considering the rare occurrence of fire while HearFire continuously emits sound all the time. To improve the effective usage of system energy, we may integrate many useful sensing tasks simultaneously, such as human activities sensing (e.g., falling, walking), human gestures sensing, localization, etc., in HearFire. Such potential functions can provide users with a full-featured system, which significantly expands the usage scenarios and increases the installation value of our proposed system.

## 7 RELATED WORK

### 7.1 Fire Detection Systems

**7.1.1 Contact-based Fire Detection.** Contact-based fire detection systems utilize physical phenomena and chemical reactions between detectors and fire to sense the flames. Thermometer-based fire detection typically utilizes thermal elements to detect fires by sensing the heat released during combustion [20, 21, 53]. However, thermometer-based fire detection is either only suitable for specific materials [53] or inevitably incurs high economic costs and inconvenient system deployment [20]. Smoke detection systems primarily refer to ionization and photoelectric sensors [33], which have been pervasively applied in indoor environments [10, 14, 21, 34, 36, 44] due to their accurate and timely smoke detection. Photoelectric smoke detectors trigger alarms based on the

intensity of scattered light, which varies with the concentration of smoke. However, non-fire particles in the environment (i.e., haze and dust) may also scatter light, which will cause false alarms when their concentrations exceed a set threshold[14, 34, 44]. The ionization sensor leverages the fact that fire smoke can affect the radioactive material, which in turn changes the electrical current traveling through it to issue the alarm. Although it is resilient to the environment, radioactive americium will be released when ionization smoke detectors are burning, incurring an inhalation risk for humans [10].

Although achieving high accuracy, those detection systems require direct contact with or close to the fire, which may pose a danger to humans. Moreover, these systems neither support timely detection [36, 51] nor can be generalized to wide use [20, 53].

**7.1.2 Contact-free Fire Detection.** Unlike contact-based detection, contact-free fire detection systems leverage the combustion phenomenon that can be remotely captured by detectors. Optical fire detectors apply low-cost optics with a simple structure, working in a specific spectrum, and can capture the thermal radiation emitted during combustion in the visible light with corresponding wavelength [4, 12, 15, 35, 42]. The intuition of optical fire detection is to trigger an alarm when the measured thermal intensity exceeds a threshold [4, 12]. However, visible light may be difficult to scatter out from derivatives (i.e., black smoke) generated by combustion, resulting in significant delays or even failure to capture the thermal radiation. Image-based fire detection is often implemented based on video sensors, recording the data as images or videos and performing fire detection by measuring the variations in pixel intensity [11, 17, 25, 63]. However, cameras require operators to accurately investigate a monitored area [11]. Although deep learning frameworks have been applied for automatic fire detection [25, 63] in recent years, high dimensional image data as well as high requirement for network traffic hinder the real-time performance of the detection systems [48–50].

Recently, researchers use Wi-Fi signals to implement fire detection since fire impacts the RF signal propagation, giving rise to the variations of CSI in terms of amplitude [24, 62]. Wi-Fire designs a series of signal processing methods to extract features from CSI and trains a supervised learning model using Random Forest classifier [62]. The performance of Wi-Fire can be further improved by selecting a set of useful subcarriers in CSI in both 2.4GHz and 5GHz channels [24]. However, CSI is only accessible by using the CSI tool for Ath9k series Wi-Fi chips and Intel 5300 Wi-Fi chips [18, 52, 58] and does not support 4G/5G band. CSI extraction for most mainstream commodity Wi-Fi chips in the market is not available for public access, making it undesirable for fire detection.

## 7.2 Acoustic Sensing Systems

In recent years, a vast body of works on acoustic sensing have been studied due to pervasive deployment of speakers and microphones in smart devices [7, 8, 16, 22, 23, 26–30, 30–32, 38, 39, 41, 43, 45–47, 54–57, 61].

**7.2.1 Target Tracking.** LLAP tracks hand motion by transmitting a single-frequency continuous wave (CW) signal and extracting phase values from the signal reflected from hand [46]. EchoTrack measures the time-of-flight (ToF) information of the sound to track finger motion since sound reflected from fingers causes different propagation delays [7]. Covertband achieves centimeter-level activity tracking for multiple persons by differentiating patterns of different spectrums derived from reflected sounds [32]. Strata improves the finger tracking accuracy to millimeter level by estimating the phase change of CIR to detect finger motions [57]. Different from existing works that primarily use the signal reflection effect to track the objects, we exploit sound energy absorption to detect fire because of the unavailability of any signal reflection from fire.

**7.2.2 Gesture Recognition.** Acoustic sensing has been applied to fulfill gesture recognition tasks since slow speed and short wavelength of sounds can capture more minute finger movement. Doppler frequency shift is measured to sense the gesture since hand movement causes changes in sound signal frequency, which can be characterized to infer gestures [16, 38]. UltraGesture measures the amplitude of the CIR and is able to recognize hand gestures with high accuracy [26]. RobuCIR applies both amplitude and phase of CIR from signal reflection and conducts

a series of interference cancellation schemes and data augmentation techniques to increase the robustness of gesture recognition [47]. VSkin enables the identification of even smaller gestures on the back of mobile devices by capturing the subtle multipath signals reflected from different fingers using CIR [39]. However, all these works can only sense the motions performed nearby smart devices (e.g.,  $\leq 0.5\text{m}$ ), which is inadequate for fire detection. In our work, we boost the sensing range by deeply analyzing the correlation between CIR measurements and system's sensing range, and selecting an appropriate acoustic frame by balancing the constraints of many practical factors.

**7.2.3 Vital Sign Detection.** Recent researches show that human vital sign can be detected via acoustic sensing by emitting frequency-modulated sound signals. ApneaApp enables apnea alarming for multiple persons by transmitting a frequency-modulated continuous wave to measure the ToF of sound reflected by human chest and abdomen during breathing [30]. C-FMCW measures the signal delay caused by human respiration by calculating the cross-correlation between FMCW sound waves and achieves high accuracy of respiration rate estimation [45]. Unlike these works that measure ToF caused by signal reflection, the ToF variation in our work is mainly due to the increase in sound speed propagated in high-temperature air.

**7.2.4 Temperature Sensing.** Acoustic-based temperature sensing has been researched in recent years. Acoustic Thermometer (AcuTe+) uses on-board dual microphones of smartphones to measure ambient temperature by estimating airborne sound propagation speed [6]. However, AcuTe+ only measures ambient temperature ranging from 5 to  $38^{\circ}\text{C}$ , which is much lower than that generated by fire. AcousticThermo applies a commodity acoustic sensor to measure temperature by calculating the sound speed within a certain distance [60]. However, AcousticThermo requires strict initialization of the system by measuring the distance of the round trip in advance (i.e., placing an obstruction in front of the device), which limits its usage scenarios. Different from these works, HearFire jointly combines acoustic absorption and sound speed characteristics and can achieve fire detection at much higher temperatures. In addition, HearFire can detect fire without any pre-setting before detection due to the ability of CIR to measure the sound propagation speed as well as the acoustic absorption pattern.

## 8 CONCLUSION

This paper proposes a cost-effective, easy-to-use, and timely indoor fire detection system via acoustic sensing using commodity devices. Our main contribution in this paper is to offer the potential for sensing fire using acoustic signals by leveraging the fact that (1) sound energy can be absorbed by fire and (2) sound speed varies with different temperatures. We use CIR to measure the impact of these two sound propagation characteristics on acoustic signals. To address the problem of short sensing range in existing fire detection and acoustic sensing systems, we mathematically map the sensing distances to transmission signal lengths. A series of practical constraints in inaudible acoustic sensing for fire detection are considered and balanced to select the appropriate transmission sound waves, achieving room-scale fire detection. We cancel the environmental interference generated by sound reflection by designing an effective sparsity-based approach based on signal reflection and absorption differences. We conduct extensive experiments to evaluate our system under different experiment settings. The results show that our proposed system can detect fire in real-time in sizeable indoor areas with an accuracy of up to 99.2%.

## ACKNOWLEDGMENTS

This work was supported by the National Nature Science Foundation of China under Grant 62102139, the Nature Science Foundation of Hunan Province of China under Grant 2022JJ30168 and the Fundamental Research Funds for the Central Universities under Grant 531118010612. This research has benefited from financial support of Lingnan University, Hong Kong Special Administrative Region, China (Jiaxing Shen is the corresponding author). We are grateful to reviewers for their insightful comments and help.

## REFERENCES

- [1] Allied Market Research. Fire Alarm and Detection System Market. <https://www.alliedmarketresearch.com/fire-alarm-and-detection-system-market-A12493>, 2021.
- [2] Jorge P Arenas and Malcolm J Crocker. Recent trends in porous sound-absorbing materials. *Sound & vibration*, 44(7):12–18, 2010.
- [3] David A Bies, Colin H Hansen, and Carl Q Howard. *Engineering noise control*. CRC press, 2017.
- [4] Rob Bogue. Sensors for fire detection. *Sensor Review*, 2013.
- [5] M. I. Boulos. Flow and temperature fields in the fire-ball of an inductively coupled plasma. *IEEE Transactions on Plasma Science*, 4(1):28–39, 1976.
- [6] Chao Cai, Henglin Pu, Liyuan Ye, Hongbo Jiang, and Jun Luo. Active acoustic sensing for hearing temperature under acoustic interference. *IEEE Transactions on Mobile Computing*, 2021.
- [7] Huijie Chen, Fan Li, and Yu Wang. Echotrack: Acoustic device-free hand tracking on smart phones. In *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, pages 1–9, 2017.
- [8] Haiming Cheng and Wei Lou. Push the limit of device-free acoustic sensing on commercial mobile devices. In *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, pages 1–10, 2021.
- [9] Congressional Research Service. Fire Statistics. <https://sgp.fas.org/crs/misc/IF10244.pdf>, 2022.
- [10] R. Daudel and M. Jean. The role of the electronic configuration in the phenomena of radioactivity. *compt rend*, 1949.
- [11] Nguyen Manh Dung and Soonghwan Ro. Algorithm for fire detection using a camera surveillance system. In *Proceedings of the 2018 International Conference on Image and Graphics Processing*, ICIGP 2018, page 38–42, New York, NY, USA, 2018. Association for Computing Machinery.
- [12] W Dunkhorst, P Lipowicz, and W Koch. Characterization of highly concentrated organic aerosols by optical extinction in the mid infrared regime: Application to e-cigarettes. *Journal of Aerosol Science*, 94:33–42, 2016.
- [13] Tony F. W. Embleton. Tutorial on sound propagation outdoors. *The Journal of the Acoustical Society of America*, 100(1):31–48, 1996.
- [14] First Alert Store. Smoke Alarms and Smoke Detectors. <https://www.firstalertstore.com>, 2022.
- [15] Sheldon Kay Friedlander. Smoke, dust and haze: Fundamentals of aerosol behavior. *New York*, 1977.
- [16] Sidhant Gupta, Daniel Morris, Shwetak Patel, and Desney Tan. Soundwave: Using the doppler effect to sense gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’12, page 1911–1914, New York, NY, USA, 2012. Association for Computing Machinery.
- [17] Winfried Halle, Christian Fischer, Thomas Terzibaschian, Adina Zell, and Ralf Reulke. Infrared-image processing for the dlr firebird mission. In Michael Cree, Fay Huang, Junsong Yuan, and Wei Qi Yan, editors, *Pattern Recognition*, pages 235–252, Singapore, 2020. Springer Singapore.
- [18] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. Tool release: Gathering 802.11n traces with channel state information. *SIGCOMM Comput. Commun. Rev.*, 41(1):53, jan 2011.
- [19] Young Hoon Jung, Trung Xuan Pham, Dias Issa, Hee Seung Wang, Jae Hee Lee, Mingi Chung, Bo-Yeon Lee, Gwangsu Kim, Chang D Yoo, and Keon Jae Lee. Deep learning-based noise robust flexible piezoelectric acoustic sensors for speech processing. *Nano Energy*, page 107610, 2022.
- [20] T. Kaiser. Fire detection with temperature sensor arrays. In *Proceedings IEEE 34th Annual 2000 International Carnahan Conference on Security Technology (Cat. No.00CH37083)*, pages 262–268, 2000.
- [21] Do-Hun Kang, Min-Sung Park, Hyoung-Sub Kim, Da-young Kim, Sang-Hui Kim, Hyeyon-Ju Son, and Sang-Gon Lee. Room temperature control and fire alarm/suppression iot service using mqtt on aws. In *2017 International Conference on Platform Technology and Service (PlatCon)*, pages 1–5, 2017.
- [22] Bryce Kellogg, Vamsi Talla, and Shyamnath Gollakota. Bringing gesture recognition to all devices. In *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*, pages 303–316, Seattle, WA, April 2014. USENIX Association.
- [23] Dong Li, Jialin Liu, Sungsoon Ivan Lee, and Jie Xiong. Lasense: Pushing the limits of fine-grained activity sensing using acoustic signals. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 6(1), mar 2022.
- [24] Junye Li, Aryan Sharma, Deepak Mishra, and Aruna Seneviratne. Fire detection using commodity wifi devices. In *2021 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, 2021.
- [25] Pu Li and Wangda Zhao. Image fire detection algorithms based on convolutional neural networks. *Case Studies in Thermal Engineering*, 19:100625, 2020.
- [26] Kang Ling, Haipeng Dai, Yuntang Liu, Alex X. Liu, Wei Wang, and Qing Gu. Ultragesture: Fine-grained gesture sensing and recognition. *IEEE Transactions on Mobile Computing*, pages 1–1, 2020.
- [27] Chao Liu, Penghao Wang, Ruobing Jiang, and Yanmin Zhu. Amt: Acoustic multi-target tracking with smartphone mimo system. In *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, pages 1–10, 2021.
- [28] Wenguang Mao, Jian He, and Lili Qiu. Cat: High-precision acoustic motion tracking. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, MobiCom ’16, page 69–81, New York, NY, USA, 2016. Association for Computing

- Machinery.
- [29] Wenguang Mao, Mei Wang, and Lili Qiu. Aim: Acoustic imaging on a mobile. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '18, page 468–481, New York, NY, USA, 2018. Association for Computing Machinery.
  - [30] Rajalakshmi Nandakumar, Shyamnath Gollakota, and Nathaniel Watson. Contactless sleep apnea detection on smartphones. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '15, page 45–57, New York, NY, USA, 2015. Association for Computing Machinery.
  - [31] Rajalakshmi Nandakumar, Vikram Iyer, Desney Tan, and Shyamnath Gollakota. Fingerio: Using active sonar for fine-grained finger tracking. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 1515–1525, New York, NY, USA, 2016. Association for Computing Machinery.
  - [32] Rajalakshmi Nandakumar, Alex Takakuwa, Tadayoshi Kohno, and Shyamnath Gollakota. Covertband: Activity information leakage using music. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(3), sep 2017.
  - [33] National Fire Protection Association. Ionization vs photoelectric. <https://www.nfpa.org/Public-Education/Staying-safe/Safety-equipment/Smoke-alarms/Ionization-vs-photoelectric>, 2022.
  - [34] National Institute of Standards and Technology. How Do Smoke Detectors Work. <https://www.nist.gov/>, 2022.
  - [35] V F Panin and A G Dashkovskii. Improvement of optical systems for detection of smokes. *IOP Conference Series: Materials Science and Engineering*, 81:012115, apr 2015.
  - [36] Sancha Panpaeng, Phattharamon Phanpeang, and Ekkasit Metharak. Cigarette smoke detectors for non-smoking areas in the building. In *2018 22nd International Computer Science and Engineering Conference (ICSEC)*, pages 1–4, 2018.
  - [37] Reuters. Paris Notre Dame Cathedral on fire. <https://www.cnbc.com/2019/04/15/paris-notre-dame-cathedral-on-fire-reuters.html>, 2019.
  - [38] Wenjie Ruan, Quan Z. Sheng, Lei Yang, Tao Gu, Peipei Xu, and Longfei Shangguan. Audiogest: Enabling fine-grained hand gesture detection by decoding echo signal. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '16, page 474–485, New York, NY, USA, 2016. Association for Computing Machinery.
  - [39] Ke Sun, Ting Zhao, Wei Wang, and Lei Xie. Vskin: Sensing touch gestures on surfaces of mobile devices using acoustic signals. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, MobiCom '18, page 591–605, New York, NY, USA, 2018. Association for Computing Machinery.
  - [40] National Wildfire Coordinating Group. Incident Operations Standards Working Team. *Glossary of wildland fire terminology*. National Wildfire Coordinating Group, 2012.
  - [41] Yu-Chih Tung, Duc Bui, and Kang G. Shin. Cross-platform support for rapid development of mobile acoustic sensing applications. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '18, page 455–467, New York, NY, USA, 2018. Association for Computing Machinery.
  - [42] Ruff Urban, Greenberg Meyer, Cleary Fisher, et al. The smoke aerosol modelling experiments (same) conducted on the international space station. In *15th Int. Conf. Autom. Fire Detect. AUBE '14*, 2014.
  - [43] Anran Wang, Jacob E. Sunshine, and Shyamnath Gollakota. Contactless infant monitoring using white noise. In *The 25th Annual International Conference on Mobile Computing and Networking*, MobiCom '19, New York, NY, USA, 2019. Association for Computing Machinery.
  - [44] Shu Wang, Xiao Xiao, Tian Deng, Ang Chen, and Ming Zhu. A sauter mean diameter sensor for fire smoke detection. *Sensors and Actuators B: Chemical*, 281:920–932, 2019.
  - [45] Tianben Wang, Daqing Zhang, Yuanqing Zheng, Tao Gu, Xingshe Zhou, and Bernadette Dorizzi. C-fmcw based contactless respiration detection using acoustic signal. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(4), jan 2018.
  - [46] Wei Wang, Alex X. Liu, and Ke Sun. Device-free gesture tracking using acoustic signals. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, MobiCom '16, page 82–94, New York, NY, USA, 2016. Association for Computing Machinery.
  - [47] Yanwen Wang, Jiaxing Shen, and Yuanqing Zheng. Push the limit of acoustic gesture recognition. *IEEE Transactions on Mobile Computing*, 21(5):1798–1811, 2022.
  - [48] Kun Xie, Xiaocan Li, Xin Wang, Jianmeng Cao, Gaogang Xie, Jigang Wen, Dafang Zhang, and Zheng Qin. On-line anomaly detection with high accuracy. *IEEE/ACM Transactions on Networking*, 26(3):1222–1235, 2018.
  - [49] Kun Xie, Xiaocan Li, Xin Wang, Gaogang Xie, Jigang Wen, Jianmeng Cao, and Dafang Zhang. Fast tensor factorization for accurate internet anomaly detection. *IEEE/ACM Transactions on Networking*, 25(6):3794–3807, 2017.
  - [50] Kun Xie, Can Peng, Xin Wang, Gaogang Xie, Jigang Wen, Jianmeng Cao, Dafang Zhang, and Zheng Qin. Accurate recovery of internet traffic data under variable rate measurements. *IEEE/ACM Transactions on Networking*, 26(3):1137–1150, 2018.
  - [51] Xin-hui Xie, Yue-fei Lv, Zheng-yang Mu, Peng-xin Bian, Jie Ji, Lian-ming Xu, and Neng Wan. An intelligent smoke detector system based on lora and indoor positioning. *DEStech Transactions on Computer Science and Engineering*, pages 268–273, 2018.
  - [52] Yaxiong Xie, Zhenjiang Li, and Mo Li. Precise power delay profiling with commodity wifi. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, MobiCom '15, page 53–64, New York, NY, USA, 2015. Association for Computing Machinery.

- [53] Hui Xu, Yang Li, Neng-Jian Huang, Zhi-Ran Yu, Peng-Huan Wang, Zhao-Hui Zhang, Qiao-Qi Xia, Li-Xiu Gong, Shi-Neng Li, Li Zhao, Guo-Dong Zhang, and Long-Cheng Tang. Temperature-triggered sensitive resistance transition of graphene oxide wide-ribbons wrapped sponge for fire ultrafast detecting and early warning. *Journal of Hazardous Materials*, 363:286–294, 2019.
- [54] Xiangyu Xu, Jiadi Yu, Yingying Chen, Yanmin Zhu, Linghe Kong, and Minglu Li. Breathlistener: Fine-grained breathing monitoring in driving environments utilizing acoustic signals. In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys ’19, page 54–66, New York, NY, USA, 2019. Association for Computing Machinery.
- [55] Qiang Yang and Yuanqing Zheng. Model-based head orientation estimation for smart devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 5(3), sep 2021.
- [56] Yanni Yang, Yanwen Wang, Jiannong Cao, and Jinlin Chen. Hearliquid: Non-intrusive liquid fraud detection using commodity acoustic devices. *IEEE Internet of Things Journal*, pages 1–1, 2022.
- [57] Sangki Yun, Yi-Chao Chen, Huihuang Zheng, Lili Qiu, and Wenguang Mao. Strata: Fine-grained acoustic-based device-free tracking. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys ’17, page 15–28, New York, NY, USA, 2017. Association for Computing Machinery.
- [58] Youwei Zeng, Dan Wu, Jie Xiong, Enze Yi, Ruiyang Gao, and Daqing Zhang. Farsense: Pushing the range limit of wifi-based respiration sensing with csi ratio of two antennas. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 3(3), sep 2019.
- [59] Hans-Jurgen Zepernick and Adolf Finger. *Pseudo random signal processing: theory and application*. John Wiley & Sons, 2013.
- [60] Fusang Zhang, Kai Niu, Xiaolai Fu, and Beihong Jin. Acousticthermo: Temperature monitoring using acoustic pulse signal. In *2020 16th International Conference on Mobility, Sensing and Networking (MSN)*, pages 683–687. IEEE, 2020.
- [61] Huanle Zhang, Wan Du, Pengfei Zhou, Mo Li, and Prasant Mohapatra. Dopenc: Acoustic-based encounter profiling using smartphones. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, MobiCom ’16, page 294–307, New York, NY, USA, 2016. Association for Computing Machinery.
- [62] Shuxin Zhong, Yongzhi Huang, Rukhsana Ruby, Lu Wang, Yu-Xuan Qiu, and Kaishun Wu. Wi-fire: Device-free fire detection using wifi networks. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–6, 2017.
- [63] Kun Zhou and Xi Zhang. Design of outdoor fire intelligent alarm system based on image recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, 34(07):2050018, 2020.