AudioSense: Enhancing Smart Speakers with Single-point Appliance Detection using Sound from Power Supplies

ABSTRACT

Over the past years, smart speakers proliferated and take a critical role in smart ecology, for the reason that they provide voice interactive capabilities using a microphone. However, smart speakers are limited to only interact with appliances that contain smart modules, which require additional hardware for traditional appliances. Existing electrical appliance identification approaches such as distributed sensing, energy disaggregation, infrastructure-mediated sensing suffer from the drawbacks of limited range, complex installation, lack of timeliness and requirement of additional hardware, respectively.

In order to overcome above shortcomings, we propose AudioSense, which breaks through the limitation of smart speakers for non-intrusive, remote detection of electrical appliances at a single point without additional hardware. AudioSense leverages the audio signal from its power supply, whose electric components are activated by the current to emit sound. The audio signal contains the frequency components of power factor correction module of a remote electrical appliance in a house. In realizing AudioSense, we proposed an optimized Variation Mode Decomposition scheme to extract the frequency component, and propose a frequency extraction scheme to build a multi-label classification model from only single label training data. Extensive experiments demonstrate that AudioSense can achieve 0.94 F1-score for classification.

1 INTRODUCTION

The smart home ecosystem has long fascinated technology visionaries who want to create a monitoring and communication mechanism between appliances. The use of smart speakers is growing rapidly in the smart home due to their ability to bring voice interaction into the home and users are quickly embracing smart speakers as a convenient hub for monitoring and controlling smart appliances. Recent reports say that 34.7% of U.S. adults own a smart speaker [57]. However, commercial off-the-shelf (COTS) services specifically require communication modules (Bluetooth, WiFi, ZigBee, etc.) to be pre-installed on appliances, which requires users to replace old appliances or install additional smart modules, thus limiting the widespread development of the smart home ecosystem. Therefore, enhancing a COTS smart speaker so it can **remotely** detect the states of appliances at a **single point** (i.e., use a single smart speaker at a fixed location to detect all appliance in the house) without the need for **additional hardware** can contribute to the development of the smart home ecosystem.

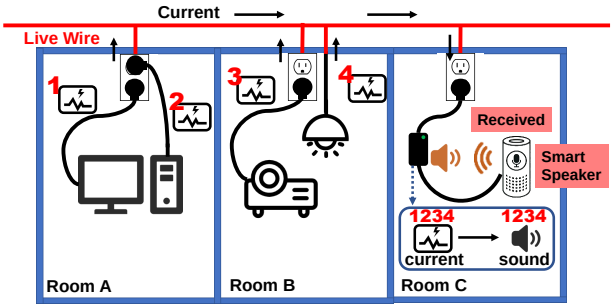
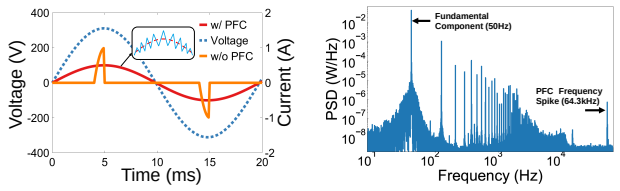


Figure 1: An illustration of AudioSense that using a smart speaker to identify electrical appliances at a single point in a house. i) The appliances generate specific current (PFC signal) while working. ii) The PFC signal thereby interferes with the branch current of the smart speaker. iii) The current cause the power supply of smart speaker to emit sound. iv) The smart speaker capture the sound with microphone, and then identify appliances from the audio signal.

In order to detect the operating status of an electrical appliance, the existing works can be divided into three categories. The first is the installation of *distributed sensors* on each appliance, thus imposing additional deployment costs on the user. It includes but not limited to RFID tags [5, 48], magnetometers [23], microphones [42], or smart plugs [31]. The second way is called *non-intrusive load monitoring* (NILM) [22] which identifies the states of appliances by extracting the consumption characteristics from the total energy usage monitored by the electric meter. Unfortunately, this approach fails in identifying complex time-varying appliances (e.g., computers) and distinguishing appliances of the same model [12, 32, 62]. The third way is called *infrastructure-mediated sensing* (IMS) which utilizes side-channels of the appliances that affect the house infrastructure (including gas [11], water [16, 17, 53], electricity [9, 15, 19–21, 41], etc) to realize a single-point appliance detection. The three types of works, however, either require additional hardware or have limited detection for whole home detection.

Different from the approaches above, we sought to propose AudioSense, a novel software-based solution which



(a) Current waveform (b) Power spectrum density

Figure 2: An illustration of power factor correction. (a) the current waveform generated by PFC module. (b) the power spectrum density of a dell desktop current. There exists a frequency spike generated by PFC.

enables a smart speaker to detect the working status of **remote** appliances (e.g., at different rooms) without installing additional hardware. As shown in Fig. 1, the working principle of AudioSense is that, when an appliance is working, it's power factor correction (PFC) module[1] in the power supply generates high-frequency **current** (i.e., PFC signal) (Sec. 2.1), which thereby interferes the power network and propagates to the branch **current** of a smart speaker's power supply (Sec. 2.2). The current through the power supply activates internal electric components to emit high-frequency **sound** (Sec. 2.3) due to Magnetrostriction and Piezoelectric effect. The sounds containing particular components of PFC signals from the appliance can be further received by the microphone on the smart speaker. By analyzing the received audio signals, AudioSense can identify the working state of the appliance.

In order to support the identification of various types of appliances and detect them at a long distance (e.g., $> 50m$), AudioSense needs to address several challenges. First, the frequency bands of PFC signals from different appliances may overlap each other, resulting in a complex mix of signals. In the case that there are multiple appliances working at the same time, AudioSense needs to distinguish these appliances in the overlapping spectrum. Second, the PFC signal from appliances farther away from the smart speaker attenuates with distance, resulting in weaker sounds. These sounds are also interfered by noise from the smart speaker itself, the environment, and other appliances. We need a denoising scheme to effectively extract PFC features from the noisy data. In addition, there are various types of PFC modules that generate various signal patterns. The denoising scheme needs to dynamically optimize its configuration to fit the current set of working appliances.

To solve the above problems, we first apply a spectral subtraction scheme followed by a high-pass filter to mitigate the noise from the smart speaker and the environment (Sec. 6.1.1). Then, optimal variational mode decomposition (OVMD) is proposed to adaptively optimize the bandwidth

[1]Not all appliances have a PFC module but it's becoming more and more common. We will discuss it in Sec. 2.1

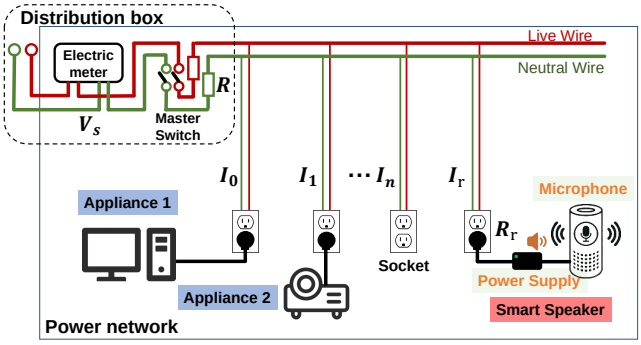


Figure 3: An illustration of the electrical network in a house.

of each intrinsic mode function to decompose the PFC signal (Sec. 6.1.2). To address the problem of overlap, we utilize Principal Component Analysis (PCA) to extract linearly additive features (Sec. 6.2). Finally, we build a multi-label classification model using only single-label data as the training set (Sec. 6.3)

We implement AudioSense on 26 different electrical appliances (13 of appliances are of the same model.) The proposed system was also tested in 5 real-world families. The contributions of the work are summarized as follows:

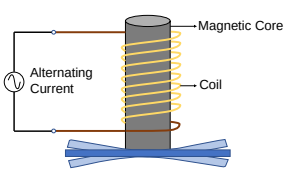
- To the best of our knowledge, we are the first to take advantage of the sound in power supply at a single point to identify remote electrical appliances using PFC signal by investigating a current-acoustic model.
- We proposed an Optimized Variation Mode Decomposition method to adaptively optimize the bandwidth of each mode to decompose the PFC signal.
- We proposed a feature extraction scheme to extract stable features to confront overlapping and aliasing problem. By utilizing the features, we built a multi-label classification model that only uses single-label training data, which improved the usability.
- We implemented AudioSense in real-world and conducted exhaustive experiments. The results show that AudioSense performs over 0.94 F1-score for appliance identification.

2 BACKGROUND

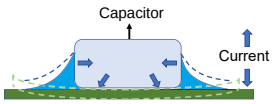
In this section, we will describe how appliances affect the smart speaker, resulting in the sound emission from the power supply.

2.1 Power Factor Correction Circuit

Power factor is an important indicator in power systems, which represents the effectiveness of energy utilization. Power factor is defined as the ratio of actual instantaneous power to apparent power [28]. Specifically, real power is the power absorbed by the load and apparent power is the power provided

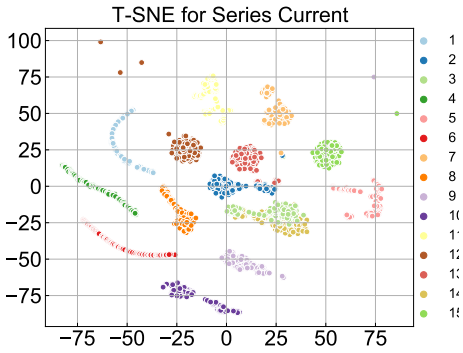


(a) Magnetrostriction

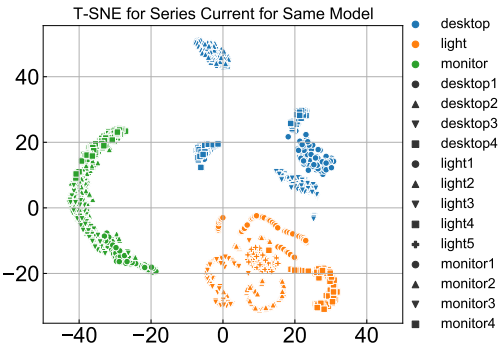


(b) Piezoelectric effect

Figure 4: An illustration of Magnetrostriction and Piezoelectric effect.



(a) Different Model



(b) Same Model

Figure 5: The t-SNE for series current of different model and same model, respectively. (a) 15 different models of electrical appliances. (b) the same model including 5 lights, 4 desktops and 4 monitors.

by the distribution system. A lower power factor requires a larger apparent power to provide the same amount of real power, which leads to energy waste. Therefore, the power factor should be increased.

The *power factor correction* (PFC) circuits are designed to improve energy utilization efficiency by increasing the power factor. An appliance with low power factor produces a non-sinusoidal current waveform, which generates multiple harmonics as shown in Fig. 2(a). These harmonics do not perform real power. In this case, PFC works by generating a high frequency current ripple using a pulse width modulator (PWM) to produce an approximate sinusoidal current waveform to reduce harmonics. PWM switches between on and off to control the input current to rise or fall periodically to fit the sine wave. The switching frequency (ω_{pfc}) is the reciprocal of the period T_{pfc} (the sum of the rise and fall times) , denoted by $\omega_{pfc} = 1/T_{pfc}$. Fig. 2(b) shows an example of frequency spike generated by a dell desktop. Particularly, PFC switching frequency typically falls into the range of $40kHz$ to $150kHz$ [46], which will not be filtered by an electromagnetic interference (EMI) filter [2]. According to IEC61000-3-2 standard [24], PFC modules must be installed in power supplies for lighting equipment with power greater than $5W$ and Class D equipment (e.g., monitors, TVs, computers, etc.) with power greater than or equal to $75W$ [49]. In addition, it is widely used in chargers to save energy [27]. PFC modules are common, as appliances containing PFC modules are reported to account for 35% of electricity consumption in the US [43].

[2]The cut-off frequency of a typical EMI filter is $200kHz$. [36]

2.2 Powerline Interference

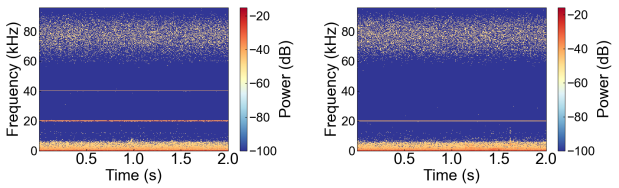
We then describe how PFC circuits in appliances interfere with power networks. A typical indoor electrical supply network is composed of a distribution box and multiple distributed outlets. An electricity meter and a master switch are usually installed in the distribution box to control the distribution and record the consumption. The distribution box delivers single-phase electricity in parallel to outlets in different rooms of the house. Fig. 3 highlights the parallel connection of power outlets and appliances. The smart speaker and other appliances are plugged into any outlet in the house. Since the power to each outlet comes from the distribution system, the relationship between the current in each branch I_i and the source voltage V_s supplied by the distribution system satisfies the following equation:

$$I_r R_r + (I_0 + I_1 + \cdots + I_n + I_r)R = V_s$$

where I_i denotes the current of the i^{th} appliance, R denotes the common resistance of the distribution box, and R_r and I_r denote the self-resistance and current of the smart speaker. Therefore, the current of the smart speaker (I_r) can be deduced as follows:

$$I_r = \frac{V_s}{R + R_r} - \frac{R}{R + R_r} \sum_{i=1}^{n} I_i \tag{1}$$

From Eq. 1, the current I_r of the smart speaker contains the current components of other appliances. As mentioned in Sec. 2.1, the PFC circuit generates output current ripples with special frequency spikes, which makes it possible to detect appliances from the smart speaker.



(a) inductor (b) capacitor

Figure 6: Spectrogram of the acoustic emission from an inductor and a capacitor when the frequency of passing current is $20kHz$. The inductor emits sound at $20kHz$ and $40kHz$ and the capacitor emits sound at $20kHz$, respectively.

2.3 Acoustic Emission

Then we show how the current I_r causes the power supply of the smart speaker to emit sound. A typical power supply structure contains an EMI filter, a rectifier, and a transformer. These internal components consist of inductors and capacitors that are affected by the AC current to generate sound. Specifically, there are two main causes of acoustic noise from power supplies: Magnetostriction and Piezoelectric effects.

Magnetrostriction, illustrated in Fig. 4(a), refers to the phenomenon that a magnetic material elongates or contracts in the direction of magneticzation when it is magnetized in a magnetic field. Specifically, the alternating current $I_m sin(\omega t)$ generates an alternating magnetic flux $\phi_m sin(\omega t)$, which deforms the magnetic core and thus causes vibrations. According to Maxwell's equations of electromagnetic attraction [38], the electromagnetic attraction force F_e can be expressed as:

$$F_e = \frac{\phi^2 S}{2\mu_0} = \frac{\phi_m^2 S}{2\mu_0}\sin^2(\omega t) = \frac{\phi_m^2 S}{4\mu_0} - \frac{\phi_m^2 S}{4\mu_0}\cos(2\omega t) \quad (2)$$

where μ_0 represents permeability of the medium and S represents cross-sectional area of the magnetic core, which are constant. From Eq. 2, the core vibrates at double frequency 2ω, resulting in sound of 2ω frequency (as observed in Fig. 6(a)). This sound will be amplified as the core comes into contact with other components, accordingly the sound will be heard.

Besides, the coil wound on magnetic core is meanwhile subjected to electrical damping force [13]. The electrical damping force can be derived as

$$F_{elec} = K_e I_m \sin(\omega t)$$

where K_e is an electromechanical constant. It illustrates that the coil emits the sound of the same frequency to the current. In general, these two acoustic emission mainly occurs in components that include coils, bobbins, and magnetic cores, such as inductors and transformers.

Piezoelectric effect is a phenomenon in which a high frequency electrical signal is applied to a capacitor, producing high frequency mechanical vibrations 4(b). The capacitor vibrates on the printed circuit board (PCB), which generates acoustic noise. According to Piezoelectric equation [2], the displacement of the capacitor surface D can be expressed as:

$$D = eS_d + \varepsilon E \quad (3)$$

$$\nabla \times E = -\frac{\partial \phi_m sin(\omega t)}{\partial t S} \quad (4)$$

where e, S_d, ε represent Piezoelectric stress constant, Strain constant, and Dielectric constant, respectively. E is the electric field which conforms to Maxwell's equations. Therefore, the frequency of the sound depends on the frequency of the current through the capacitor (as we can observe in Fig. 6(b)). This noise occurs mainly in capacitors and crystal oscillators.

Remarks. In summary, both inductors and capacitors can emit sound due to alternating currents passing through the circuit, and the frequency of the sound corresponds to the frequency of the current passing through. Fig. 6 shows the spectrum of the sound emitted by an inductor and a capacitor at a current of $20kHz$.

3 FEASIBILITY STUDY

We conducted experiments to preliminarily study the feasibility of using sound emitted by a single-point power supply to identify electrical appliances. As mentioned in Sec. 2.3, the frequency of sound is closely related to that of PFC signal. Since the PFC signal will be extremely weakened due to parallel transmission, we separately investigate the series PFC signal and the audio signal received by a microphone. There are two questions we mainly focus on: i) Can PFC signal be used as a fingerprint to identify electrical appliances? ii) Does the sound contain the corresponding PFC current component, and whether it can be transmitted remotely?

3.1 PFC Signal as Fingerprint

If the PFC signal is to be used as a feature by which to identify electrical appliances, we must first verify its uniqueness and stability. We used a current sensor ACS712 [39] sampling at $160kHz$ to measure the series current of 20 appliances (lights, desktops, laptops, TVs, monitors, smartphone chargers, accumulators, projectors) across 10 days. For preprocessing, we subtract the average value to remove the DC component.

Uniqueness. Firstly, we need to verify whether the PFC signal of the appliances are unique. The dataset is divided into two subsets: a subset of different models (different brands or different appliances) and a subset of the same model of the same appliance (including light, desktop, and monitor). We utilize t-SNE [56] to visualize the clustering relationship

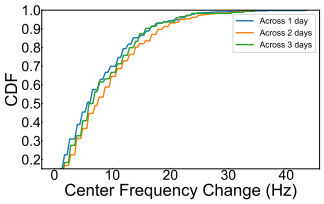
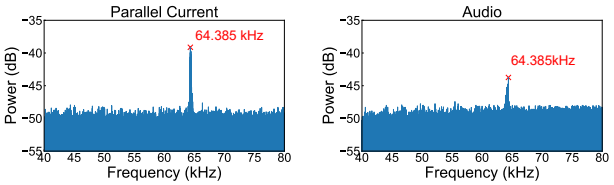


Figure 7: The CDF of PFC center frequency variation over time.



(a) Parallel Current (b) Audio

Figure 8: Spectrum of the current of a smart speaker and the audio received by a microphone. When a dell desktop is working, the current and the audio have the same frequency spike $64.385kHz$ which demonstrate that the audio signal performs consistent with the current.

between data. Fig. 5(a) plots the t-SNE between 15 different models we collected in 2-dimensional space. It can be seen from the figure that there exists obvious variance across different models. Fig. 5(b) shows the t-SNE between the same model of the appliances, including 5 lights, 4 desktops and 4 monitors. We can see that the difference between appliances is much greater than that between the intra-class. However, the PFC still performs differences between the same model due to variations in the manufacturing processes. These results reveal that the PFC frequency of both different appliances and the same model of appliances are distinguishable.

Stability. Secondly, we need to verify that the PFC signal are stable across times. Fig. 7 presents the CDF of PFC center frequency variation across days ($1day$,$2days$,$3days$). The result shows that the variation in PFC signal is within $40Hz$ which confirms that the PFC signal is stable.

In summary, the PFC signal is unique to each appliances and is stable in time. Therefore, it is possible to utilize PFC signal as the fingerprint to identify electrical appliances. In the next section, we will discuss the consistency in projecting the current to the sound.

3.2 Current-Acoustic Model

The PFC signal generated by electrical appliances will be attenuated through the powerline due to the line resistance. In this section, we will verify whether the the sound emitted by power supply is consistent with the branch current in smart speaker side, and whether it can be transmitted remotely.

Consistency. First, we need to verify whether the PFC signal can transmit to the smart speaker side and whether it can activate the power supply to emit sound. A dell desktop is working as the transmitter and a Tmall smart speaker power supply is used as the receiver. We use an an electret condenser microphone [3] to sample the sound at $192kHz$. Fig. 8 shows the spectrum of the current passing through smart speaker and the audio received, both of which have the same frequency spike $64.385kHz$. It demonstrates that the audio signal performs consistent with the current. Note that the SNR of the sound ($7dB$) is lower than the SNR of the current ($14dB$) due to component properties, it still can be received by a microphone.

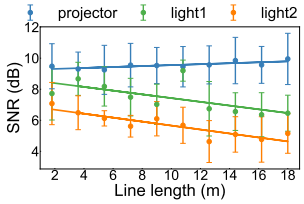
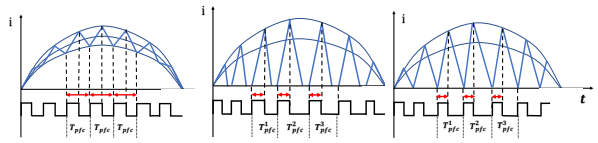


Figure 9: The SNR of the sound of three different models (a projector and two different lights) under distance attenuation (from 1.8m to 18m).

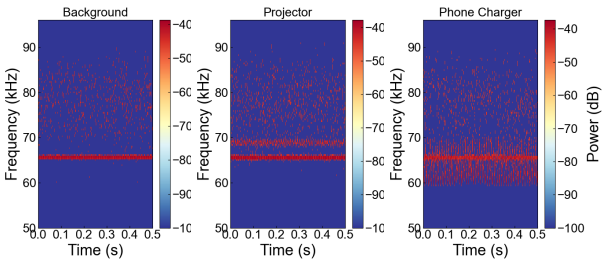
Distance attenuation. The resistance of line increases with distance, thereby probably weakens the amplitude of PFC signal. In this part, we verify whether the PFC signal can remotely activate sound in a house. We conduct experiments by adjusting the line length of the transmitter side, from $1.8m$ to $18m$, and record the audio signal. Fig. 9 shows the SNR of the sound for 3 model of appliances for parallel transmission. Interestingly, the SNR for the projector is slightly improved due to compensation of voltage loss [37]. That is, the line resistance will divide the voltage, resulting in voltage drops on transmitter side. In order to keep the power constant, the appliance automatically increases the current. For the other two different lights, the SNR decreases. We perform a linear regression to estimate the measurement. The rate of decline is approximate, which prove the principle in Sec. 2.2. The slope is about $-0.13dB/m$ which means the maximum distance for activating sound is about $53m$. It can meet the requirements for covering a common family. We also conduct experiment in real-world family environment, the results show the PFC signal can be effectively detected.

[3] The electret condenser microphone can effectively reduce eletromagnetic interference to ensure that the signal picked up is sound



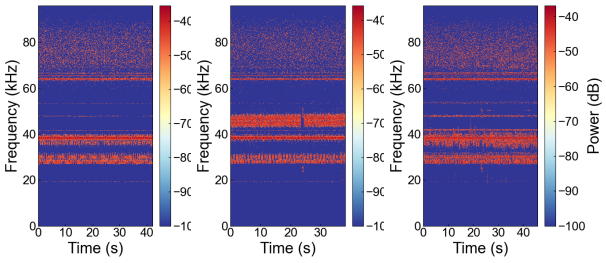
(a) CCM (b) DCM (c) CrCM

Figure 10: The illustrations of current waveform generated by PFC of CCM, DCM and CrCM, respectively. CCM performs constant frequency. DCM and CrCM performs changing frequency.



(a) Background (b) Projector (c) Phone charger

Figure 11: Audio signal of power supply that the projector performs constant frequency ($69kHz$) and the phone charger performs changing frequency ($59kHz - 69kHz$).



(a) Battery< 20% (b) Battery20% – 80% (c) Battery> 80%

Figure 12: Audio signal of power supply when a phone charger is of different battery levels.

4 PRE-ANALYSIS OF AUDIO SIGNAL FEATURE

The audio signal received by the smart speaker is closely related to the PFC component in the received current. However, PFC module varies from device to device. In this section, we highlight the pre-analysis of the received audio signal with various PFC component.

4.1 Impact of Conduction Modes

Typically, PFC modules can be divided into three types according to their conduction modes: Continuous Conduction Mode (CCM), Discontinuous Conduction Mode (DCM), and Critical Conduction Mode (CrCM) [46]. Fig. 10 illustrates the current waveform generated by PFC of different conduction modes.

CCM performs constant frequency, fits sinusoidal wave by adjusting the duty ratio in a cycle, as shown in Fig. 10(a). The duration of a cycle is constant, denoted by T_{pfc}. The frequency of a CCM appliances can derived as $1/T_{pfc}$, thereby resulting in a fixed frequency component in audio signal. CCM usually applies in high-power appliances ($> 300W$), such as desktop and projectors. Different from CCM, DCM adjusts the current by regulating to different cycle, which keeps the duty ratio constant, as shown in Fig. 10(a). Similar to DCM, CrCM also keeps the duty ratio (on-time) constant, but when the current drops to zero, the switch of PFC immediately turns on, as shown in Fig. 10(c). Therefore, both DCM and CrCM perform changing switching frequency, resulting in changing frequency in audio signal with wide bandwidth. These two conduction modes usually applies in low power appliances ($> 300W$), such as laptops and phone chargers. Fig. 11 shows the audio signal emitted by a power supply of while a projector ($330W$) is working and a phone charger ($65W$) is working. Results show that projector performs constant narrow band frequency ($69kHz$), the phone charger performs changing frequency across a wide band ($59kHz - 69kHz$). Note that even if the appliance may perform frequency changing, it is still a periodic signal with a constant center frequency.

4.2 Impact of Charging States

Charging is a common behavior in nowadays daily life. Monitoring the charging states is important for protecting lithium batteries and preventing fires. Specifically, the processing for charging can be divided into three modes according to battery level: pre-charge mode for battery level less than 20%, constant current charging mode from 20% to 80%, constant voltage charging mode above 80% [25]. In different stages of charging, its internal charging power will change, thereby affects PFC signal, and finally affects the audio signal received. It provides the possibility for AudioSense to detect charging states. Fig. 12 shows the audio signal received when a phone charger is working of different battery levels. Result shows that the battery level will influence the PFC signal, therefore we can use the audio signal to identify charging states.

5 SYSTEM OVERVIEW

We design a single-point electrical appliance identification system, AudioSense, which leverages the sound of power

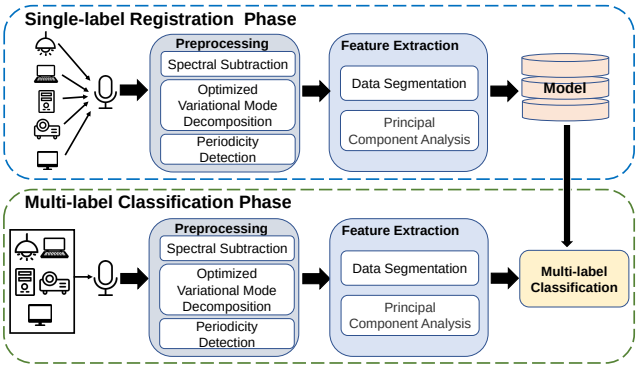


Figure 13: The system overview of AudioSense.

supply without any other additional devices. Fig. 13 illustrates the architecture of the proposed AudioSense system, comprising two parts: *single-label registration phase* and *multi-label classification phase*.

In single-label registration phase, AudioSense collects the data when the user first accesses an electrical appliances. The user first needs to turn on the smart speaker to record the a background sound of smart speaker itself. For a newly appliance, for every possible state, AudioSense collects a 2-minute data for registration. After the signal is collected, it first goes through a high-pass filter to discard vocal and ambient sound bands below $20kHz$ and processes a spectral subtraction to boost SNR. Then, AudioSense utilizes an optimized Variational Mode Decomposition (OVMD) and periodicity detection to extract the PFC signal of the appliance. After that, AudioSense segments the data into parts and then utilize PCA for feature extraction. The model will be leveraged for multi-label classification phase.

In multi-label classification phase, our target is to classify multi-label data from only single-label data. AudioSense first captures the sound of power supply, and preprocessing in the same way as in register phase. AudioSense then utilizes VMD and periodicity detection to extract the characteristic band. Then, it also extracts the features using PCA. Comparing the acquired dimentsionaliy reduction component by PCA with the model in feature matching library by calculating Euclidean distance. The distance below the threshold means that the appliance is working.

6 SYSTEM DESIGN

The audio signal received by smart speaker contains the same frequency component as the PFC signal in current. In order to support identification of more devices at a single point, we need to solve the problem of signal feature extraction and overlapping by multiple appliances. In this section, we introduce the detailed design of AudioSense.

6.1 Preprocessing

The collected audio signal consists of four parts: environmental noise, sound card circuit noise, sound of the SMPS from the internal transformer itself and sound produced by external PFC signal that we required. Extracting the band of PFC frequency component cleanly is crucial to the performance for feature extraction and classification. The pipeline for preprocessing is presented as follows:

6.1.1 Spectral subtraction. In order to boost SNR, our target is to weaken the background noise and extract the band of PFC component. We first utilize a high-pass filter to remove the audible noise ($< 20kHz$) produced by music, vocal and environment. The left signal received Y_r (shown as Fig. 14(b)) can be denoted as

$$Y_r = S_{pfc}(\omega_0) + N_1(\omega_1) + N_2(\omega_2)$$

where $S_{pfc}(\omega_0)$ performs the external PFC signal,ω_0 is the central frequency of PFC component. $N_1(\omega_1)$ denotes the sound card circuit noise, $N_2(\omega_2)$ denotes the sound of internal transformer itself. Similarly, the background signal Y_{bg} (shown as Fig. 14(a)) can be denoted as

$$Y_{bg} = N_1(\omega_1) + N_2(\omega_2)$$

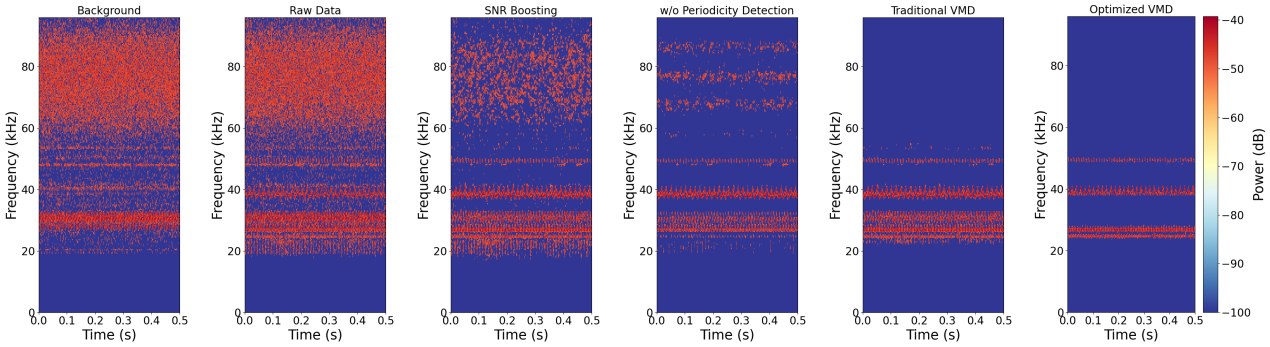
Fortunately, the noise from hardware circuitry $N_1(\omega_1)$ is time invariant and the sound produced by the transformer in smart speaker itself obeys a certain distribution. Therefore, a spectral subtraction scheme [3] can be used to weaken the background noise, the enhanced set of PFC component $\hat{S}_{pfc}(\omega_0)$ can be obtained as follows:

$$\hat{S}_{pfc}(\omega_0) = F^{-1}[[|Y_r| - E[|Y_{bg}|]]e^{j\theta}] = S_{pfc}(\omega_0) + \epsilon$$

where $E[|Y_{bg}|]$ is the spectral estimation of background noise, F^{-1} is inverse fast Fourier transform (IFFT),θ is the original phase, ϵ is estimation error. Fig. 14(c) shows the signal with SNR boosting. We need to emphasize that spectral subtraction can only weaken the noise but does not completely remove it. Thus, we need to further extract the band of PFC component without noise.

6.1.2 Optimized Variational Mode Decomposition. After the SNR boosting process, the weakened background noise still exists. As mentioned in Sec. 3.1, the PFC signal operates stably in unchanged central frequency. To this end, we use Variational Mode Decomposition (VMD) [14] to decompose the audio signal into several band components, which is called Intrinsic Mode Functions (IMFs) and then extract the PFC component in decomposed IMFs. VMD is a nonrecursive, intrinsic and adaptive signal decomposition method, which is robust to noise and sampling. The process of the VMD algorithm is described as follows.

Traditional VMD. For a 1D signal $\hat{S}_{pfc}(\omega_0)$, VMD separates it into k narrow band signals, which is denoted by u_k,



(a) background noise (b) Raw data (c) SNR boosting (d) Traditional VMD without periodicity detection (e) Traditional VMD with periodicity detection (f) Optimized VMD with periodicity detection

Figure 14: Spectrograms of the acoustic emission from a power supply during the preprocessing pipeline. (a) is the background noise contains sound card noise $60kHz - 90kHz$ and noise of smart speaker itself. (b) is the raw data when two lights and a lenovo laptop is working. (c) weakens the background noise by spectral subtraction to boost SNR. (d) shows a traditional VMD without periodicity detection, containing random noise band ($60kHz-90kHz$) (e) shows the traditional VMD with periodicity detection, appearing broadband with noise ($25khz-32khz$). (f) performs the best by utilizing optimized VMD with periodicity detection, the PFC frequency are correctly distinguished.

with different estimated central frequency ω_k. For each u_k, shift its analytic signal from the pass band to base band and then compute its penalty term. The optimization objective in frequency domain can be represented as:

$$\min_{u_k(\omega_0)} \{|| \sum_k u_k(\omega_0) - \hat{S}_{pfc}(\omega_0)||_2^2 + \alpha \sum_k ||j\omega u_k(\omega_0 - \omega_k)||_2^2\}$$

Note that, in frequency domain, $||j\omega u_k(\omega_0 - \omega_k)||_2^2$ is equal to $||j(\omega_0 - \omega_k)u_k(\omega_0)||_2^2$. To guarantee the fidelity of decomposition, the constrained variational model that corresponds to the decomposition process of the response signal $\hat{S}_{pfc}(\omega)$ is given as

$$\min_{u_k(\omega_0)} \{|| \sum_k u_k(\omega_0) - \hat{S}_{pfc}(\omega_0)||_2^2 + \alpha \sum_k ||j(\omega_0 - \omega_k)u_k(\omega_0)||_2^2\}$$
$$s.t. \sum_k u_k(\omega_0) = \hat{S}_{pfc}(\omega_0)$$

The optimal solution can be achieved by using Lagrange multiplier, followed by ADMM algorithm to find the saddle point. Please refer to [14] for more details.

Parameters optimization. However, the VMD algorithm decomposes the original signal based on pre-defined parameters (i.e. k and α). If the parameter k is too large, it will possibly lead to spurious modes (consisting of noise content) or mode splitting (the same component is shared by several IMFs). Conversely, a small k will result in mode loss or mode mixing. For the parameter α, a small value leads to a larger bandwidth (contains noise) and a large value will cause mode splitting. Back to our experiments, the bandwidth of PFC component varies from device to device. An universal parameter design cannot restore the original PFC waveform well. To tackle this problem, we design a bandwidth-based parameter optimization to obtain a better component.

Take a decomposed narrow band signal $u_k(\omega)$ as an example, the bandwidth is equivalent to the width of the rectangle whose area is equal to that of the power spectrum of $u_k(\omega)$, and whose height is the amplitude at the center of power spectrum [4]. The bandwidth (bw) can be derived as follows

$$\Delta_f^2 = \int_0^\infty (\omega - \omega_k)^2 \frac{|\hat{u_k}(\omega)|^2}{\int_0^\infty |\hat{u_k}(\omega)|^2 d\omega} d\omega$$
$$bw_k = 2\Delta_f$$

where $\hat{u_k}(\omega)$ denotes the amplitude spectrum of $u_k(\omega)$, bw_k denotes the bandwidth of decomposed signal $u_k(\omega)$. We marked the total bandwidth (BW) of decomposed signal as $BW = \sum_{k=1}^k bw_k$ .In theory, the more the decomposed mode fits the real signal, the smaller the BW is. The inappropriate k or α will both result in a larger BW, and vice versa. Therefore, we initialize parameters through a couple of set from $k = 2:1:40, \alpha = 1000:500:100000$, which means the range for k and α is $[2, 40]$ and $[1000, 100000]$, the step of parameters adjustment for k and α is 1 and 500, respectively. It can distinguish the difference at most $2.4kHz$ between two modes. We choose the combination of k and α to achieve the smallest BW of as the optimal solution.

Periodicity Detection. The signal is decomposed into several IMFs, we only concentrate on the sound produced by PFC. Note that in order to fit curve of a s sinusoidal, the PFC component is periodic. By contrast, the noise floor of

the soundcard ($N_1(\omega_1)$) is non-periodic (i.e., random noise). In this part, we managed to extract the external PFC signal from the IMFs.

For each IMF, we first compute the upper envelope, and then derive the corresponding auto-correlation coefficients. Next, we segment the IMF into multiple parts, where the start point is determined by the peak of auto-correlation coefficients. The duration of the parts should exceed the fundamental period of each IMF, empirically we set $0.05s$. We computed the Pearson Correlation Coefficient (PCC) for pairs of segments, and we utilize a threshold to select which IMF is periodic. Fig. 14(f) shows the best result of OVMD with periodicity detection by comparison with traditional VMD with/without periodicity detection. We can see that the characteristic frequency band is well extracted.

6.2 Feature Extraction

In above subsection, we extract the PFC component from the total audio signal. The audio signal is closely related to the current signal. Due to the characteristic of the current with linear addition, the audio signal thence performs the same characteristics. Our target is to identify multi-label appliance combination using only single-label data. Traditional methods such like neural networks needs to train with the multi-label data to learn by utilizing data augmentation. However, if a new appliance is registering, it needs the combination with any other appliances as the cost is exponential growth. To this end, we solved the problem by utilizing the linear addition characteristics of audio signal. Specifically, we utilized Principal Component Analysis (PCA) [54] to extract linear additive features as the model. PCA is effictive in our system because it maintains an additivity property. Given the PCA measures in the interpretation of maximum variance direction by $\frac{1}{n}\sum_{i=1}^{n}(V^T x_i)^2 = \frac{1}{n}v^T XX^T v$ In signal processing, it is considered that the signal has a large variance and the noise has a small variance. The best k-dimensional feature is that after the n-dimensional sample points are converted to k-dimensional, the sample variance on each dimension is very large. Therefore we can transform the equation above into Lagrangian form $\mathcal{L}_v = v^T XX^T v + \lambda(1 - v^T v)$

Note that when we deploy the Eigen decompsition, we need to use a linear kernel function to secure its additivity property.

6.3 Multi-label Classification

After feature projection, we will perform distance metrics on the obtained dimensionality-reduced data. We divided it into different classes and perform multi-label classification. One challenge we have in this section is that the device information obtained by our register part is single label. Therefore, we need to use the model trained on single-label data for multi-label classification.

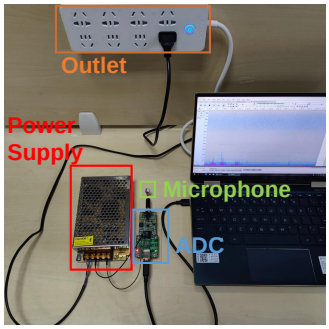


Figure 15: The prototype of AudioSense system.

To achieve the target, first, we need to verify the distance metrics method. Considering that there is linear additivity between our different types of equipment, there is no need to consider the correlation between the equipment, only their characteristics can be reasonably separated. In this paper we deploy Euclidean distance as our distance metric method , which is defined as $d(X, Y) = \sqrt{\sum_i (x_i - y_i)^2}$

Next, our equipment identification process is similar to a checking list process. According to the device data obtained during the registration phase, we look for the characteristics of the device in the sound signal, so as to detect the registered devices one by one. In the entire classification process, the data we need is only single-label data, and we are dealing with multi-label complex environments. In the classification process, with the help of the linear additivity we mentioned earlier, we can easily process the existing signals or remove the checked signals to complete the identification process.

7 EVALUATION

In this section, we evaluate the performance of our proposed AudioSense

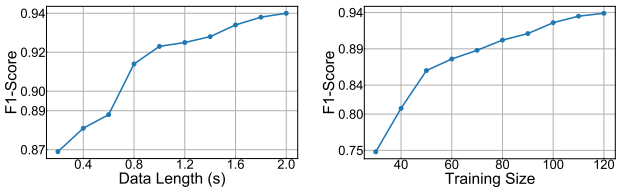
7.1 Experiment Setup

Prototype: Due to privacy protection issues, we cannot obtain microphone sampling information from COTS smart speaker devices. Therefore, we built a prototype for verifying the efficacy of AudioSense. Fig. 15 presents the prototype schematic. A power supply is plugged into an outlet, for receiving the PFC signal from remote outlets. It meanwhile provides power to an analog-to-digital converter (ADC) (i.e. sound card) which continuously sample the audio signal at $192kHz$ received by an electret condenser microphone. The working principle of the prototype is the same as nowadays smart speakers. By analyzing the collected audio signal, we can identify states of remote electrical appliances.

Dataset: In our evaluation, two datasets were used in our experiment. *1) Laboratory dataset.* With a laboratory environment, we collected data from over 26 appliances using the

Table 1: List of 26 Appliances collected in the experiments. For each kind of appliance, the corresponding manufacturer, quantity of same model, different models are listed.

Appliance	Company	#Same	#Different	#Total
Light	Xiaomi,PHILIPS,etc.	5	2	7
Monitor	Dell, PHILIPS,etc.	4	1	5
Desktop	Dell, HP, etc.	4	2	6
Laptop	Hp, Lenovo, etc.	0	2	2
Projector	Canon, Sony	0	2	2
Phone charger	Huawei, Xiaomi, etc.	0	2	2
Induction cooker	Supor	0	1	1
TV	Skyworth	0	1	1



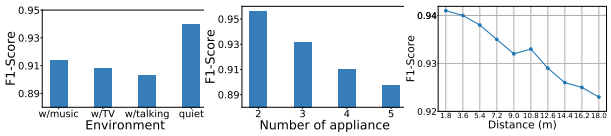
(a) Data length (b) Training size

Figure 16: Performance under training with different data length and different training size.

prototype mentioned above across 10 days. The laboratory dataset can be divided into single-label dataset (for training and testing) and multi-label dataset (only for testing). For single-label dataset, we collect data for each appliance, for $2-minutes$ as registration. Table. 1 lists the appliances used in the experiments. Particularly, we also collected data on load change of induction cooker and the charging status of smartphones. For the multi-label dataset, we randomly combine these appliances from the number of 2 to 5. *2) Real-world Dataset.* We also implemented our prototype in real world. We collected the data from 5 real home environment, whose area from $80m^2$ to $140m^2$. The prototype can be placed in different rooms of the house. In this way, we can effectively evaluate the efficacy of AudioSense in real-world.

Train-test split.: We divided the dataset into training set and testing set according to the time. Specifically, the forward 5 days used as training set and afterward 5 days used as testing set.

Performance metrics. Classifier performance was indicated by the average of F1 scores, computed as: $F1_i = \frac{2\times p_i \times r_i}{p_i + r_i}$ and p_i and r_i as the precision and recall of type i of appliances. The macro F1 score can then be computed as: $F1 = \frac{1}{N}\sum_{i=1}^{N} F1_i$ where N denotes the number of different appliances or working states. We used the macro-averaged F1 score as a metric by which to compare the performance of the various models.



(a) Envrionment (b) Number of appliance (c) Distance

Figure 17: The robustness of AudioSense against environmental noise, the number of stacked devices, and the transmitted distance.

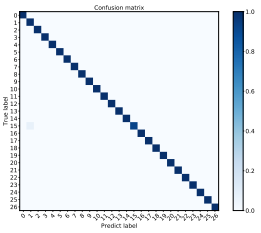


Figure 18: Confusion matrix of classification on single label datasets.

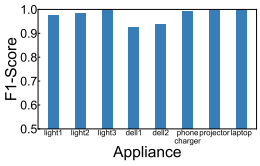


Figure 19: F1-score for multi-label classification

7.2 Microbenchmark

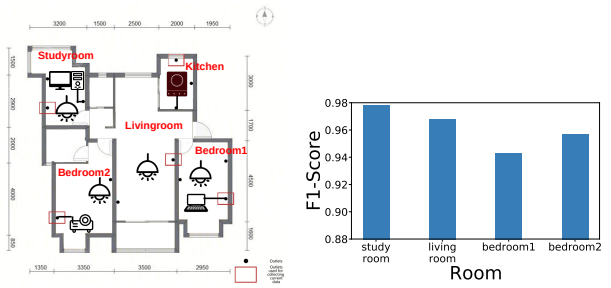
7.2.1 Data length. Data length is crucial to achieving a good balance between response time and prediction accuracy. We segmented data samples of various duration (from 0.2s to 2.0s). According to Fig. 16(a), the F1-score increases as the data length increases, and performs convergence at 2.0s. So we select 2.0s as the data length for training.

7.2.2 Training size. We trained our classification model using training data of various sizes. From Fig. 16(b), it shows the performance that with the training size increases, F1-score increases. Considering about the experience of registration phase, we select $2-minutes$ data, and split it for 120 parts of data length of 2.0s with overlap of 1.0s.

7.3 Overall performance

In this section, we evaluate the overall performance of AudioSense on single-label classification and multi-label classification,respectively.

7.3.1 Single-label classification. We evaluate the overall performance over 26 electrical appliances. Fig. 18 shows the confusion matrix on single-label dataset. It achieves overall performance of 98.7% accuracy, which demonstrates AudioSense can work well in single-label classification.



(a) Floor plan (b) Performance

Figure 20: A case study of implementation of AudioSense in real-world home environment.

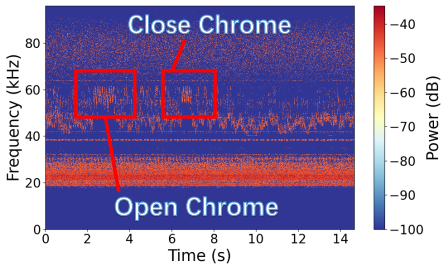


Figure 21: The spectrogram of a laptop when a chrome browser is opening and closing.

7.3.2 Multi-label classification. We also evaluate the performance on multi-label classification. We select 8 appliances for pairwise stacking. Among the appliances, 3 lights and 2 dell desktops are of the same model. Fig. 19 shows AudioSense achieve good performance on multi-label classification.

7.4 Robustness

In this subsection, we evaluate the robustness of AudioSense.

7.4.1 Impact of environmental interference. First, we investigate the impact of environmental interference. As shown in Fig. 17(a), the environmental noise performs slight decreases to the F1-score. It is because that we removed the noise below than $20kHz$, and we did a well-done work on extract PFC component.

7.4.2 Impact of the number of appliances. Second, we sought to study the impact of multiple electrical appliances superimposed on multi-label classfication. As shown in Fig. 17(b), F1-score decreases as the number of appliances increases, but still performs F1-score higher than 0.89. It shows AudioSense is robust against impact of number of appliances.

7.4.3 Impact of distance. Second, we sought to determine the impact of the transmit distance. Fig. 17(c) shows that the F1-score decreases as the distance decreases. It is because the line resistance increases as the distance increases, resulting in lower magnitude of PFC signal. The result shows AudioSense still performs over 0.92 F1-score, which means it is robust to the distance.

7.5 Case Study

We also implemented AudioSense in real-world home environment across 5 houses. Fig. 20 shows a case study of the implementation of a house with four rooms whose area is $136m^2$ (Fig. 20(a) is the floor plan). We evaluate the placement of our prototype in four rooms: study room, living room and two bedrooms. Fig. 20(b) shows F1-score can achieve over 0.93 in all of the rooms, which demonstrates that AudioSense can work in real-world situation.

8 RELATED WORK

8.1 Electrical Appliance Detection

Electrical appliance detection has attracted lots of attention in recent years with the growth of IoT devices. Prior works have given number of different problem-solving approaches and practicable applications. They mainly lies in three categories, namely close-range sensing, non-intrusive load monitoring (NILM) and infrastructure-mediated sensing (IMS).

Close-range sensing: A common idea for detecting electrical appliance is to install distributed sensors on each appliance. [33, 44, 51, 52]. Device-level detection performs straightforward sensing results but requires time consuming and costly installation and maintenance. There also exists lots of indirectly sensing technologies such as visual-based [34, 63, 68], audio-based [8, 42, 67], RF-based [5, 18, 48, 60], and radiated-emission based approaches [1, 23, 29, 30]. However, these approaches must be used with limited distance (within a room) and limited number of devices. Our approach on contrary can detect multiple electrical events with a single-point detection.

Non-intrusive load monitoring: Non-intrusive load monitoring (NILM) [22] primarily detect electrical appliances based on their energy usage patterns form a power meter of a household. State-of-the-art solutions disaggregate power consumption data of individual appliances and extract the consumption characteristics from the total energy usage [12, 32, 62]. However, these works are unsuccessful in identifying complex time-varying consumption devices (e.g. desktops, projectors) and low-power equipments (e.g. CFLs, LEDs), distinguishing appliance of the same model. Our approach utilizes time-invariant feature of each appliances which makes up the shortcomings of conventional NILM techniques.

Infrastructure-mediated sensing: Infrastructure-mediated sensing (IMS) [40] is another single point technology which detects events that affect house infrastructure including gas [11], water [16, 17, 53], and electricity [9, 15, 19–21, 41].

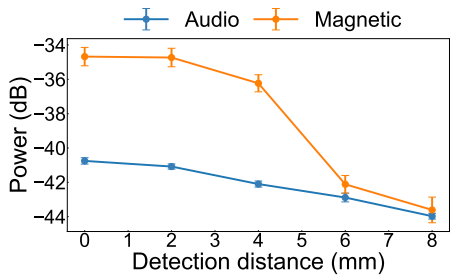


Figure 22: The power attenuation of the audio signal and the magnetic field strength against detection distance.

VibeSense [50] used a Doppler vibrometer to capture subtle vibrations caused by various devices from house's interior surface. A series of works by Patel, et al. [15, 21, 41] measure electromagnetic interference (EMI) using expensive systems (e.g. USRP and spectrum analyzers.) Gulati, et al. [19, 20] developed a customized EMI sensor to sense common mode EMI from electrical appliances. NoDE [47] and OutletSpy [65] sense powerline voltage for communication and information theft using an oscilloscope. Our approach also belongs to IMS technologies but it does not need any additional device or customized sensing system unlike the approaches mentioned above.

8.2 Acoustic-based Appliaction

Acoustic-based sensing [6] has been widely employed for a couple of applications with the explosive growth of speakers and microphones on commodity devices, ranging from indoor localization [26, 58], gesture tracking [59, 61],health sensing [45, 64], authentication [7], lip-reading [35, 66], and so on. For the electrical events, there exists many techniques by hearing audiable motor sound [8, 42, 67]. Howerver, these approaches may endanger users' privacy because human voice may also be recorded and they require a close-range detection. On contrary, our approach focus on inaudible sound ($> 20kHz$) without violating privacy. Choi, et al. [10] couple switching EM leakages with MSoCs to realize an audio eavesdropping threat using an USRP. By comparison, our approach makes use of the sound produced by switching noise in the current coupled with electronic components in adapter as features to identify electrical appliances without any additional device.

9 DISCUSSION

9.1 Sound vs. Magnetic

In this section, we want to discuss the difference between sound signals and magnetic signals in propagation, and explain why sound signals can stand out as our transmission medium.

Spread distance. During the propagation process, the sound vibrates and propagates in the air, and the attenuation speed is slower than that of the magnetic field. Specifically, the sound decays linearly, and the magnetic field strength is cubic decay. Compared to magnetism, without the help of additional enhancement equipment (e.g., attenna), the sound can be received at a greater distance, so our microphone does not need to be very close to the circuit. Fig. 22 shows the power attenuation of the audio signal and the magnetic strength received by DRV425 Hall sensor against detection distance ($0 - 8mm$). It shows the magnetic field strength decays faster than the sound.

Magnetic shield. In industrial production and life, magnetic signals are often regarded as a kind of pollution to the signal due to their influence on devices and metal components. Also, electromagnetic interference (EMI) disrupts the operation of electrical and electronic equipment, creating malfunctions and faulty readings [55]. Therefore, almost all electronic equipment will be equipped with magnetic shield, such as installing Faraday cage, adding copper pieces and so on. In comparison, high-frequency inaudible acoustic noise attracts less attention. Most electrical appliances will not be specially equipped with sound shield so that the transmission of audio signals will not be affected.

9.2 PFC vs. EMI

In this section, we want to discuss the relationship between PFC signal and EMI, and explain that AudioSense can be used in a wider range of applications.

Electromagnetic interference (EMI) is a complex mechanism that takes place at different levels including the chassis, board, component, and finally, the device level. PFC signal is also one kind of EMI with signal frequency between $40 - 150kHz$. As for EMI in current, once their bandwidth is lower than 96khz, they can be identified by AudioSense and applied to our system. Therefore, our AudioSense has a wider range of applications.

9.3 App-level identification

In this section we will discuss the extension future work and other possible application scenarios of the system. We noticed that the audio signal performs a variation when a program is launching. (e.g., open a chrome browser.) However, it only appears at the moment of launching or closing the program, as shown in 21. Particularly, [65] also observed this phenomenon in voltage waveform, and further utilized the characteristics of the voltage of PFC to detect which application is launching when using a desktop. Our work focus more on steady state feature now, but can also detect app-level identification in future work using audio signal.

REFERENCES

[1] Ammar Ahmed Alkahtani, Farah Hani Nordin, ZAM Sharrif, Nur Badariah Bte, and Ahmad Mustafa. 2012. Analysis on RF emission of electrical appliances. In *2012 IEEE International Conference on Control System, Computing and Engineering*. IEEE, 539–543.

[2] Henno Allik and Thomas JR Hughes. 1970. Finite element method for piezoelectric vibration. *International journal for numerical methods in engineering* 2, 2 (1970), 151–157.

[3] Steven Boll. 1979. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on acoustics, speech, and signal processing* 27, 2 (1979), 113–120.

[4] Ronald Newbold Bracewell and Ronald N Bracewell. 1986. *The Fourier transform and its applications*. Vol. 31999. McGraw-Hill New York.

[5] Michael Buettner, Richa Prasad, Matthai Philipose, and David Wetherall. 2009. Recognizing daily activities with RFID-based sensors. In *Proceedings of the 11th international conference on Ubiquitous computing*. 51–60.

[6] Chao Cai, Rong Zheng, and Menglan Hu. 2019. A survey on acoustic sensing. *arXiv preprint arXiv:1901.03450* (2019).

[7] Jagmohan Chauhan, Yining Hu, Suranga Seneviratne, Archan Misra, Aruna Seneviratne, and Youngki Lee. 2017. BreathPrint: Breathing acoustics-based user authentication. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*. 278–291.

[8] Jianfeng Chen, Alvin Harvey Kam, Jianmin Zhang, Ning Liu, and Louis Shue. 2005. Bathroom activity monitoring based on sound. In *International Conference on Pervasive Computing*. Springer, 47–61.

[9] Ke-Yu Chen, Sidhant Gupta, Eric C Larson, and Shwetak Patel. 2015. DOSE: Detecting user-driven operating states of electronic devices from a single sensing point. In *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 46–54.

[10] Jieun Choi, Hae-Yong Yang, and Dong-Ho Cho. 2020. TEMPEST Comeback: A Realistic Audio Eavesdropping Threat on Mixed-signal SoCs. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*. 1085–1101.

[11] Gabe Cohn, Sidhant Gupta, Jon Froehlich, Eric Larson, and Shwetak N Patel. 2010. GasSense: Appliance-level, single-point sensing of gas activity in the home. In *International Conference on Pervasive Computing*. Springer, 265–282.

[12] Wang Dan, Huang Xiao Li, and Ye Shu Ce. 2018. Review of non-intrusive load appliance monitoring. In *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*. IEEE, 18–23.

[13] Rohan Dayal, Suman Dwari, and Leila Parsa. 2010. A new design for vibration-based electromagnetic energy harvesting systems using coil inductance of microgenerator. *IEEE Transactions on Industry Applications* 47, 2 (2010), 820–830.

[14] Konstantin Dragomiretskiy and Dominique Zosso. 2013. Variational mode decomposition. *IEEE transactions on signal processing* 62, 3 (2013), 531–544.

[15] Miro Enev, Sidhant Gupta, Tadayoshi Kohno, and Shwetak N Patel. 2011. Televisions, video privacy, and powerline electromagnetic interference. In *Proceedings of the 18th ACM conference on Computer and communications security*. 537–550.

[16] Jon Froehlich, Eric Larson, Tim Campbell, Conor Haggerty, James Fogarty, and Shwetak N Patel. [n.d.]. Infrastructure-Mediated Single-Point Sensing of Whole-Home Water Activity. ([n. d.]).

[17] Jon E Froehlich, Eric Larson, Tim Campbell, Conor Haggerty, James Fogarty, and Shwetak N Patel. 2009. HydroSense: infrastructure-mediated single-point sensing of whole-home water activity. In *Proceedings of the 11th international conference on Ubiquitous computing*. 235–244.

[18] Tobias Grosse-Puppendahl, Sebastian Herber, Raphael Wimmer, Frank Englert, Sebastian Beck, Julian Von Wilmsdorff, Reiner Wichert, and Arjan Kuijper. 2014. Capacitive near-field communication for ubiquitous interaction and perception. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 231–242.

[19] Manoj Gulati, S Sundar Ram, ANGSHUL Majumdar, and Amarjeet Singh. 2016. Detecting IT and lighting loads using common-mode conducted EMI signals. In *Proc. NILM2016 3rd International Workshop on Non-Intrusive Load Monitoring, 5p.*

[20] Manoj Gulati, Shobha Sundar Ram, Angshul Majumdar, and Amarjeet Singh. 2016. Single point conducted EMI sensor with intelligent inference for detecting IT appliances. *IEEE Transactions on Smart Grid* 9, 4 (2016), 3716–3726.

[21] Sidhant Gupta, Matthew S Reynolds, and Shwetak N Patel. 2010. ElectriSense: single-point sensing using EMI for electrical event detection and classification in the home. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*. 139–148.

[22] George William Hart. 1992. Nonintrusive appliance load monitoring. *Proc. IEEE* 80, 12 (1992), 1870–1891.

[23] Hua Huang and Shan Lin. 2020. MET: a magneto-inductive sensing based electric toothbrushing monitoring system. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*. 1–14.

[24] IECEE. 2018. IEC61000-3-2;Electromagnetic compatibility (EMC)- Part 3-2. https://webstore.iec.ch/publication/67329

[25] Texas Instruments. 2017. Flash Battery Charging Pushes the Boundary of Charging Current. https://www.ti.com/lit/an/slua820/slua820.pdf?ts=1629466986523&ref_url=https%253A%252F%252Fwww.google.com%252F

[26] Ruoxi Jia, Ming Jin, Zilong Chen, and Costas J Spanos. 2015. SoundLoc: Accurate room-level indoor localization using acoustic signatures. In *2015 IEEE International Conference on Automation Science and Engineering (CASE)*. IEEE, 186–193.

[27] Pui Yee Kong, JA Aziz, MR Sahid, and Low Wen Yao. 2014. A bridgeless PFC converter for on-board battery charger. In *2014 IEEE Conference on Energy Conversion (CENCON)*. IEEE, 383–388.

[28] Abdallah Kouzou. 2018. Power Factor Correction Circuits. In *Power Electronics Handbook*. Elsevier, 529–569.

[29] Anand Sunil Kulkarni, Cindy K Harnett, and Karla Conn Welch. 2014. EMF signature for appliance classification. *IEEE Sensors Journal* 15, 6 (2014), 3573–3581.

[30] Gierad Laput, Chouchang Yang, Robert Xiao, Alanson Sample, and Chris Harrison. 2015. Em-sense: Touch recognition of uninstrumented, electrical and electromechanical objects. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. 157–166.

[31] Shih-Hsiung Lee and Chu-Sing Yang. 2017. An intelligent power monitoring and analysis system for distributed smart plugs sensor networks. *International Journal of Distributed Sensor Networks* 13, 7 (2017), 1550147717718462.

[32] Jian Liang, Simon KK Ng, Gail Kendall, and John WM Cheng. 2009. Load signature study—Part I: Basic concept, structure, and methodology. *IEEE transactions on power Delivery* 25, 2 (2009), 551–560.

[33] Joshua Lifton, Mark Feldmeier, Yasuhiro Ono, Cameron Lewis, and Joseph A Paradiso. 2007. A platform for ubiquitous sensor deployment in occupational and domestic environments. In *Proceedings of the 6th international conference on Information processing in sensor networks*. 119–127.

[34] Beth Logan, Jennifer Healey, Matthai Philipose, Emmanuel Munguia Tapia, and Stephen Intille. 2007. A long-term evaluation of sensing

modalities for activity recognition. In *International conference on Ubiquitous computing*. Springer, 483–500.

[35] Li Lu, Jiadi Yu, Yingying Chen, Hongbo Liu, Yanmin Zhu, Yunfei Liu, and Minglu Li. 2018. Lippass: Lip reading-based user authentication on smartphones leveraging acoustic signals. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 1466–1474.

[36] Abdul Majid, Jawad Saleem, and Kent Bertilsson. 2012. EMI filter design for high frequency power converters. In *2012 11th International Conference on Environment and Electrical Engineering*. IEEE, 586–589.

[37] SI MALAFEEV and SS MALAFEEV. 2018. COMPENSATION OF VOLTAGE LOSS IN THE POWER LINE. *DAAAM International Scientific Book* (2018).

[38] James Clerk Maxwell. 1865. VIII. A dynamical theory of the electromagnetic field. *Philosophical transactions of the Royal Society of London* 155 (1865), 459–512.

[39] Allegro microsystems. 2019. ACS712 datasheet. https://www.allegromicro.com/-/media/files/datasheets/acs712-datasheet.ashx

[40] Shwetak N Patel. 2008. *Infrastructure mediated sensing*. Georgia Institute of Technology.

[41] Shwetak N Patel, Thomas Robertson, Julie A Kientz, Matthew S Reynolds, and Gregory D Abowd. 2007. At the flick of a switch: Detecting and classifying unique electrical events on the residential power line (nominated for the best paper award). In *International Conference on Ubiquitous Computing*. Springer, 271–288.

[42] Nilavra Pathak, Md Abdullah Al Hafiz Khan, and Nirmalya Roy. 2015. Acoustic based appliance state identifications for fine-grained energy analytics. In *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 63–70.

[43] Luis Pérez-Lombard, José Ortiz, and Christine Pout. 2008. A review on buildings energy consumption information. *Energy and buildings* 40, 3 (2008), 394–398.

[44] Matthai Philipose, Kenneth P Fishkin, Dieter Fox, Henry Kautz, Donald Patterson, and Mike Perkowitz. 2003. Guide: Towards understanding daily life via auto-identification and statistical analysis. In *Proc. of the Int. Workshop on Ubiquitous Computing for Pervasive Healthcare Applications (Ubihealth)*. Citeseer.

[45] Kun Qian, Chenshu Wu, Fu Xiao, Yue Zheng, Yi Zhang, Zheng Yang, and Yunhao Liu. 2018. Acousticcardiogram: Monitoring heartbeats using acoustic signals on smart devices. In *IEEE INFOCOM 2018-IEEE conference on computer communications*. IEEE, 1574–1582.

[46] ON Semiconductor. 2011. Power factor correction (PFC) handbook. *Choosing the right power factor controller solution. Denver: ON Semiconductor* (2011).

[47] Zhihui Shao, Mohammad A Islam, and Shaolei Ren. 2020. Your noise, my signal: exploiting switching noise for stealthy data exfiltration from desktop computers. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 4, 1 (2020), 1–39.

[48] Joshua R Smith, Kenneth P Fishkin, Bing Jiang, Alexander Mamishev, Matthai Philipose, Adam D Rea, Sumit Roy, and Kishore Sundara-Rajan. 2005. RFID-based techniques for human-activity detection. *Commun. ACM* 48, 9 (2005), 39–44.

[49] Energy Star. 2018. Computers specification version 7.0. https://www.energystar.gov/products/spec/computers_specification_version_7_0_pd

[50] Wei Sun, Tuochao Chen, Jiayi Zheng, Zhenyu Lei, Lucy Wang, Benjamin Steeper, Peng He, Matthew Dressa, Feng Tian, and Cheng Zhang. 2020. VibroSense: Recognizing Home Activities by Deep Learning Subtle Vibrations on an Interior Surface of a House from a Single Point Using Laser Doppler Vibrometry. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–28.

[51] Emmanuel Munguia Tapia, Stephen S Intille, and Kent Larson. 2004. Activity recognition in the home using simple and ubiquitous sensors. In *International conference on pervasive computing*. Springer, 158–175.

[52] Emmanuel Munguia Tapia, Stephen S Intille, Louis Lopez, and Kent Larson. 2006. The design of a portable kit of wireless sensors for naturalistic data collection. In *International Conference on Pervasive Computing*. Springer, 117–134.

[53] Edison Thomaz, Vinay Bettadapura, Gabriel Reyes, Megha Sandesh, Grant Schindler, Thomas Plötz, Gregory D Abowd, and Irfan Essa. 2012. Recognizing water-based activities in the home through infrastructure-mediated sensing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. 85–94.

[54] Michael E Tipping and Christopher M Bishop. 1999. Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 61, 3 (1999), 611–622.

[55] E Unal, A Gokcen, and Y Kutlu. 2006. Effective electromagnetic shielding. *IEEE Microwave magazine* 7, 4 (2006), 48–54.

[56] Laurens Van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of machine learning research* 9, 11 (2008).

[57] Voicebot.ai. 2020. Smart Home Consumer Adoption Report December 2020. https://research.voicebot.ai/report-list/smart-home-consumer-adoption-report-us-2020/

[58] Anran Wang and Shyamnath Gollakota. 2019. Millisonic: Pushing the limits of acoustic motion tracking. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–11.

[59] Wei Wang, Alex X Liu, and Ke Sun. 2016. Device-free gesture tracking using acoustic signals. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*. 82–94.

[60] Yan Wang, Jian Liu, Yingying Chen, Marco Gruteser, Jie Yang, and Hongbo Liu. 2014. E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures. In *Proceedings of the 20th annual international conference on Mobile computing and networking*. 617–628.

[61] Sangki Yun, Yi-Chao Chen, Huihuang Zheng, Lili Qiu, and Wenguang Mao. 2017. Strata: Fine-grained acoustic-based device-free tracking. In *Proceedings of the 15th annual international conference on mobile systems, applications, and services*. 15–28.

[62] M Zeifinan, Craig Akers, and Kurt Roth. 2011. Non intrusive appliance load monitoring (nialm) for energy control in residential buildings. *Energy Etliciency in Domestic Appliances and Lighting 20 II* (2011), 24–26.

[63] Chenyang Zhang and Yingli Tian. 2012. RGB-D camera-based daily living activity recognition. *Journal of computer vision and image processing* 2, 4 (2012), 12.

[64] Fusang Zhang, Zhi Wang, Beihong Jin, Jie Xiong, and Daqing Zhang. 2020. Your Smart Speaker Can" Hear" Your Heartbeat! *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 1–24.

[65] Juchuan Zhang, Xiaoyu Ji, Yuehan Chi, Yi-chao Chen, Bin Wang, and Wenyuan Xu. 2021. OutletSpy: cross-outlet application inference via power factor correction signal. In *Proceedings of the 14th ACM Conference on Security and Privacy in Wireless and Mobile Networks*. 181–191.

[66] Yongzhao Zhang, Wei-Hsiang Huang, Chih-Yun Yang, Wen-Ping Wang, Yi-Chao Chen, Chuang-Wen You, Da-Yuan Huang, Guangtao Xue, and Jiadi Yu. 2020. Endophasia: Utilizing Acoustic-Based Imaging for Issuing Contact-Free Silent Speech Commands. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 1 (2020), 1–26.

[67] Ahmed Zoha, Alexander Gluhak, Michele Nati, Muhammad Ali Imran, and Sutharshan Rajasegarar. 2012. Acoustic and device feature fusion for load recognition. In *2012 6th IEEE International Conference Intelligent Systems*. IEEE, 386–392.

[68] Nadia Zouba, Bernard Boulay, Francois Bremond, and Monique Thonnat. 2008. Monitoring activities of daily living (ADLs) of elderly based