VEH-COM: Demodulating Vibration Energy Harvesting for Short Range Communication

Guohao Lan, *Student Member, IEEE*, Weitao Xu, *Student Member, IEEE*, Sara Khalifa, *Member, IEEE*, Mahbub Hassan, *Senior Member, IEEE*, and Wen Hu *Senior Member, IEEE*

Abstract—This paper investigates the possibility of using a vibration energy harvesting (VEH) device as a communication receiver. By modulating the ambient vibration energy using a transmitting speaker, and demodulating the harvested power at the receiving VEH, we aim to transmit small amounts of data at low rates between two proximate devices. The key advantage of using VEH as a receiver is that the modulated sound waves can be successfully demodulated directly from the harvested power without employing the power-consuming digital signal processing (DSP), which makes a VEH receiver significantly more power efficient than a conventional microphone-based decoder. To address the extremely narrow bandwidth of VEH, we design a simple ON-OFF keying modulation, but optimized for VEH hardware. Experiments with a real VEH device shows that, at a distance of 2 cm, a laptop speaker with the proposed modulation scheme can achieve 30 bps communication for a target bit error rate of less than 1%, which would enable many emerging short range applications, such as mobile payment. The communication range of a laptop can be extended to 80 cm for 5 bps, allowing a range of other audio-based device-to-device communications, such as a web advertisement on a laptop browser transferring tokens to a nearby smartphone. We also demonstrate that the proposed VEH-based sound decoding is resilient to background noise, thanks to its extremely narrow power harvesting bandwidth, which works as a natural noise filter.

1 Introduction

Recently, there is a growing interest in using speaker and microphone to transfer small amounts of data via sound waves between proximate devices. Examples include Google Tone [1] for sharing URLs between nearby devices, sound-based authentication from SlickLogin [2], Shopkick's [3] in-store voucher-delivery using sound beacons, SilverPush's viewer tracking for TV commercials or web-based ads by using audio-based cookies [4], and so on. Commercial interests in speaker-microphone-based data communication are equally matched by academic research. For example, in a recent study, Ka et al., [5] proposed advanced modulation and coding scheme to realize robust audio communication for TV's *second screen services* that enable TV programs to push notifications to nearby smartphones. Combining Fast Fourier Transform (FFT) with Hilbert transform, Lee et al., [6] designed an efficient signal processing method for microphones that can compensate for Doppler effect due to mobility of the devices. There are several other recent efforts [7], [8], [9], [10], [11] dedicated to increase the data rate and range of audio communication using built-in microphones.

While previous works differ in their modulation, coding, and signal processing techniques, they all rely on complex DSP, such as FFT, to obtain frequency-domain information from the time-domain microphone samples. Microphone-based audio communication therefore consumes significant power of the mobile device. Our measurements with different smartphones show (see Section 5) that FFT can consume up to several hundreds of milliwatts.

In this paper, we investigate the possibility of using a VEH device as a power-efficient audio communication receiver. Our motivation for VEH-based audio communication is two-fold. First, the recent advancements in VEH technology has probably brought the tipping point for it to be embraced in mobile devices, which is evidenced by its successful integration in many wireless sensors nodes [12], [13], [14] and realistic considerations for personal mobile devices [15], [16], [17]. Second, in principle, a VEH device should be able to harvest energy from ambient sound vibrations [18], [19], making it a candidate technology for audio communications receiver. By modulating the ambient vibration energy using a transmitting speaker, and demodulating the harvested energy at the receiving VEH, it may be possible to transmit small amounts of data at low rates between two proximate devices. The overwhelming advantage of using VEH as a receiver is that the modulated sound waves can be demodulated directly from the harvested energy without employing the power-consuming DSP. As a matter of fact, it will be shown later in the paper that FFT is the main contributor to power consumption for audio receivers. A VEH receiver without FFT therefore can significantly more power efficient than a conventional microphone-based decoder that relies on FFT.

A key challenge for using VEH for communication is its extremely narrow bandwidth. Power harvested from VEH has a peak at a particular frequency, called resonance frequency. If the vibration frequency deviates from this by a small amount, the power harvesting capability is significantly reduced. Usable power harvesting bandwidth for typical VEH devices can be as narrow as 1Hz [20], [21]. Consequently, we propose the Vibration-ON-OFF-Keying (VOOK) modulation that uses a single carrier frequency matched to the resonance frequency of the VEH. We study the power generation shape of the VEH and further optimize the signal power detection that improves the demodulattion performance by several folds. The novelty and

contributions in this paper can be summarized as follows:

- We propose the use of VEH as a audio communication receiver which eliminates FFT-related power consumption in acoustic receivers. To the best of our knowledge, this is the first attempt to use VEH as a audio communication receiver.
- We propose a OOK-based (de)modulation scheme and optimize its performance for VEH. Comparing to a naiive approach, the optimized method improves the data rate by 200% for short distance communications, and improves the communication range by 100% for low data rate communications.
- We evaluate the proposed (de)modulation techniques using a real VEH device. Our results show that at a distance of 2cm, a laptop speaker with the proposed modulation scheme can achieve 30bps communication for a target bit error rate of less than 1%, which would enable many emerging short range applications, such as mobile payment. The communication range of a laptop can be extended to 80cm for 5bps, allowing a range of other audio-based device-to-device communications, such as a web advertisement on a laptop browser transferring tokens to a nearby smartphone.
- Finally, we analyze the robustness of VEH communication against noise in typical environments. We show that the extremely narrow power harvesting bandwidth of VEH works as a natural noise filter making it immune to typical noise in indoor environments.

The rest of the paper is structured as follows. In Section 2 we review the basics of vibration energy harvesting. In Section 3 we present the system design of VEH-COM, followed by its performance evaluation in Section 4. FFT-related power measurement study is presented in Section 5. Related works are reviewed in Section 6 before concluding the paper in Section 7.

2 Vibration Energy Harvesting

In this section, we introduce the basics of vibration energy harvesting (VEH) and piezoelectric energy harvester (PEH). Vibration energy harvesting is the process of generating electrical energy from ambient vibration sources. The most widely applied VEH energy conversion techniques are the piezoelectric [22], electromagnetic [12], and electrostatic [23]. Among them, piezoelectric has been widely shown the greatest potential to achieve better performance (higher voltage and power density levels) in harvesting energy [23].

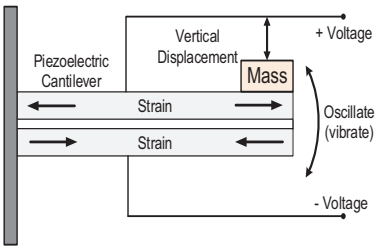


Fig. 1: The principle of piezoelectric energy harvester.

PEH converts mechanical energy into electric by straining a piezoelectric material. The stress applies to the piezomaterial will cause the piezoelectric effect which converts the external mechanical strain into electric voltage. The strain may come from a variety of ambient sources, such as human motions [16], [17], mechanical vibrations from infrastructures [24], engine vibration [15], and even acoustic waves [18], [19]. Figure 1 shows the principle of PEH. The piezoelectric vibrational energy harvesters are usually modeled as an inertial oscillating system consisting of a cantilever beam attached with two piezoelectric outer-layers. One end of the beam is fixed to the device, while the other is set free to oscillate (vibrate). An inertial mass is usually attached to one end of the piezoelectric cantilever to provide heavier strain or deformation. When the piezoelectric material is subjected to a mechanical stress, it expands on one side and contracts on the other, the induced piezoelectric effect will generate an alternating voltage (AC voltage) output as the beam oscillates around its neutral position. In this study, we choose the V25W piezoelectric vibration energy harvester from MIDÉ Technology [25] to build our VEH receiver.

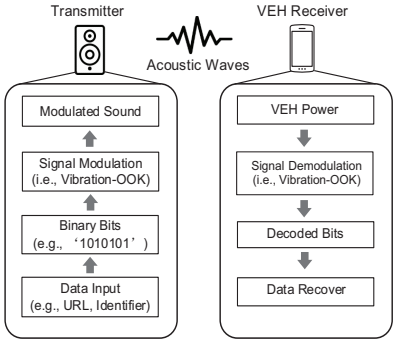
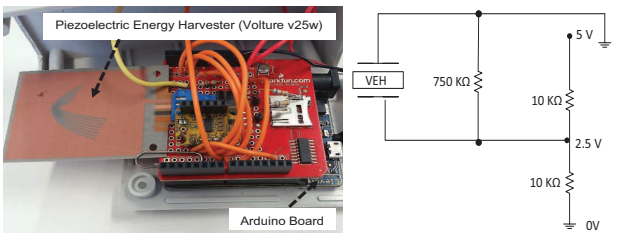


Fig. 2: An overview of VEH-COM.

3 VEH-COM System Design

Figure 2 depicts the proposed system, which consists of a transmitter modulating sound waves using a speaker, and a VEH acting as a receiver attempting to demodulate the energy harvested from the sound vibrations. When the user tries to send a piece of information, such as a short URL, the input data are first encoded into a string of binary bits. We designed a modulation scheme called vibration On-Off keying (VOOK) to modulate the binary bit string into acoustic signal. The modulated sound is then played through the host device's voice speaker at a specific sound frequency which ensures the best signal (de)modulation performance for the VEH. At the receiver, the VEH employs our proposed symbol detection techniques that enable demodulation of information bits by processing the harvested energy levels at periodic intervals.

Design of sound modulation and demodulation that works effectively with VEH requires good understanding of energy harvesting properties of the VEH hardware. We, therefore, build a custom VEH data recorder capable of recording energy harvested from any vibration sources, including sound. We then use this data logger to design and evaluate the proposed modulation and demodulation techniques.



(a) VEH data recorder. (b) Circuit diagram.

Fig. 3: Custom-made VEH data recorder.

3.1 VEH Data Recorder

Figure 3(a) shows the design of our VEH data recorder. A V25W piezoelectric vibration energy harvester is mounted on the Arduino UNO board which is an open source development platform based on the ATmega328P microcontroller (MCU). The VEH is connected to an external resistor of 750KΩ as shown in the circuit diagram in Figure 3(b). We do not physically store energy on the device, but the output AC voltage from the harvester is sampled by the MCU via its onboard 10-bit analog-to-digital converter (ADC) at 1KHz frequency and the sampled AC voltage data are stored in a microSD card for further analysis. This way, we are able to estimate harvested energy over any given time interval at a resolution of 1 ms. For example, given an AC voltage sample of V volt, the harvested power (in Watt) at that instant can be estimated as $P = \frac{V^2}{750K\Omega}$. The whole system is powered by an external 9V battery.

3.2 VEH Frequency Response

We conduct a 'sine sweep' test to explore the response of the VEH for a given sound frequency generated by a nearby speaker. We implement a MATLAB program at a laptop to generate a sinusoid sound wave at a particular frequency for 200ms. The sound is played through an external speaker at a volume of 60dBSPL. During this 200ms, we calculate the peak AC voltage recorded by the VEH recorder. We repeat this at each frequency from 50Hz to 450Hz. We conduct the sine sweep test for two different distances, 5cm and 10cm, between the speaker and the VEH.

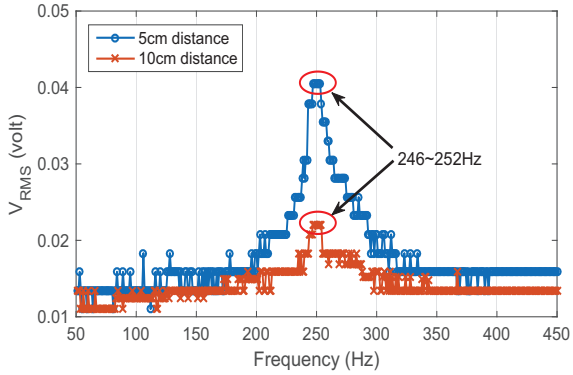


Fig. 4: VHS frequency response test for different distances.

Figure 4 presents the results of our sine sweep tests where the horizontal numbers (X-axis) indicate transmitted sound frequencies and the vertical numbers (Y-axis) represent root-mean-square voltage (V_{RMS}) generated by the VEH. It is interesting to observe that the VEH has an extremely narrow response bandwidth of only a few Hertz between 246Hz-252 Hz. The harvested power peaks within this narrow bandwidth and falls rapidly outside this band, which is consistent with the VEH literature [20], [21]. It is also interesting to note that, although the peak harvested power decreases with the distance between the speaker and the VEH, the response bandwidth is not affected by the distance. This property is very useful as it enables a transmitter to select the right frequency without having to estimate the distance.

Given the extremely narrow energy harvesting bandwidth of VEH, it is difficult to employ frequency shift keying (FSK) modulation techniques, such as OFDM, that employ a large number of different frequencies to achieve high data rates. It is also not desirable to employ phase shift keying (PSK) as it would require complex circuitry to detect phases. Instead, we propose to use the well-known ON-OFF keying, which is a binary version of amplitude shift keying (ASK), that can be readily implemented using a single frequency and would allow demodulation by simply observing the amount of harvested energy.

ON-OFF keying is a general concept, which stipulates that a signal is transmitted for bit '1' (transmission is turned ON) and no signal is transmitted for '0' (transmission is turned OFF). Thus a single frequency is sufficient. It is relatively straightforward to implement this simple modulation in radio frequency (RF) based systems, which has a very fast response for turning a signal ON and OFF. A VEH, however, contains mechanical elements, which has significant inertia. To achieve optimal performance, design of any ON-OFF keying modulation and demodulation for VEH-based receivers, therefore, must consider such inertia. In the following section, we present the proposed ON-OFF keying for VEH, which we call vibration ON-OFF keying (VOOK).

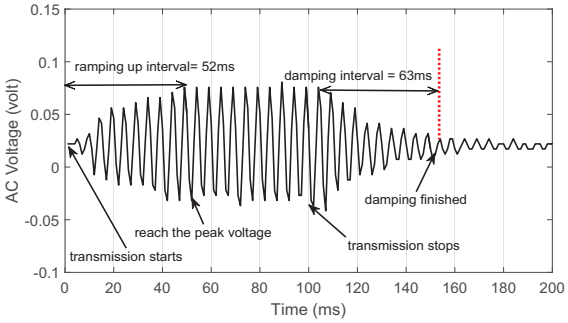
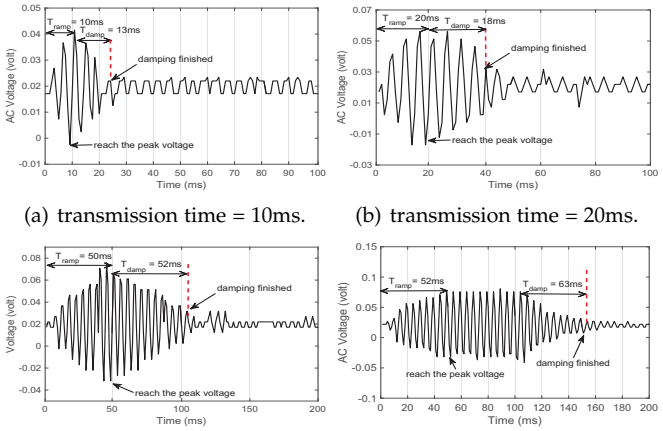


Fig. 5: Ramping up and damping intervals for 'ON' symbol.

3.3 VOOK Modulation

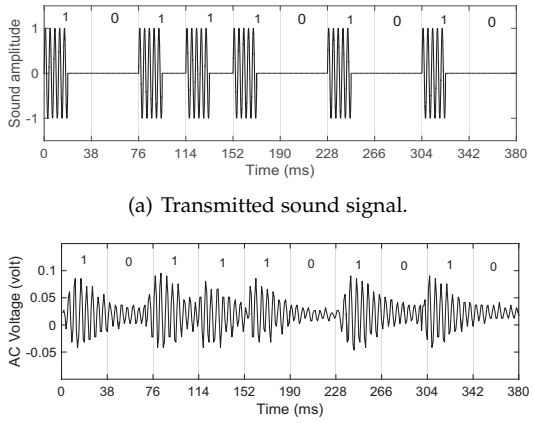
To profile the effect of inertia on VEH energy harvesting, we conduct a simple test of transmitting a sound wave at the resonance frequency of 247Hz for 100ms (transmission time, T_S, is 100ms) and record the harvested energy over time. Figure 5 plots the AC voltage output of the VEH for 200ms. We can clearly see that the VEH starts to harvest energy immediately after the start of transmission, but there is a



(a) transmission time = 10ms. (b) transmission time = 20ms.

(c) transmission time = 50ms. (d) transmission time = 100ms.

Fig. 6: Effect of T_S on T_{ramp} and T_{damp}.



(a) Transmitted sound signal.

(b) VEH AC voltage signal.

Fig. 7: VOOK transmission at 26.3 bps ($T_B = 38$ms)

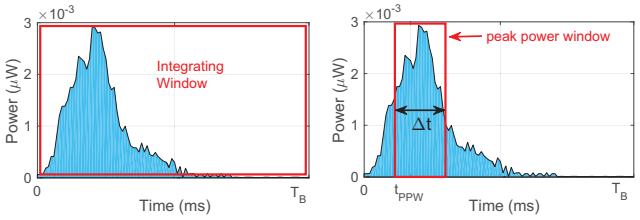
TABLE 1: Data rates for different transmission times.

T_S (ms)	Power (μW)	T_{ramp} (ms)	T_{damp} (ms)	T_B (ms)	Data Rate (bps)
100	0.0052	52	63	163	6.1
50	0.0039	50	52	102	9.8
20	0.0021	20	18	38	26.3
10	0.0014	10	13	23	43.4

'ramp up' phase, T_{ramp}, of about 52ms before the cantilever can reach the peak vibration amplitude and produce the peak power expected at the modulated frequency. We also observe that the vibration does not stop immediately when the transmission stops at 100ms, but instead it starts to reduce its vibration amplitude slowly and it takes about 63ms of 'damping time', T_{damp}, before the vibration and the consequent power generation stop completely.

From this simple test, we learn that the total bit interval, T_B, should include the damping time, i.e., we must have $T_B = T_S + T_{damp}$ to avoid inter-symbol interference when a '1' is followed by a '0'. It is obvious that longer the T_S, the higher the accumulated harvested energy and the easier it is to detect the signal or bits. However, given that the data rate is reciprocal of T_B, increasing transmission time would decrease the data rates. As such, we have a motivation to seek ways to reduce the transmission time as much as possible without making it difficult to detect signals.

Since it takes maximum 52ms for the VEH to reach the voltage peak, for $T_S < 52ms$, it results in a lower peak power but also shorter time for ramp up and damping. Figure 6 shows how the ramp up time, the peak harvested energy, and the damping times are affected by the choice of T_S. The results are summarised in Table 1. We see that for $T_S < 52ms$, $T_S = T_{ramp}$. This table can be used to configure VOOK for different data rates, which require different bit intervals imposing different requirements for the transmission times. For example, for a transmission time of 20 ms, a damping time of 18 ms is required, yielding a bit interval of 20+18=38 ms or a data rate of 26.3 bps. Figure 7(a) shows an example of VOOK transmitting a 10 bits data ('1011101010') at 26.3 bps. The corresponding AC voltage generated by the VEH during the demodulation is shown in Figure 7(b).



(a) BASIC (b) Peak Power

Fig. 8: Averaging windows for symbol power estimation.

3.4 VOOK Demodulation

VOOK demodulation for a given bit interval k is achieved by first determining the average power harvested, $\bar{P}(k)$, during the bit interval k and then applying a decision threshold η^* to obtain the received symbol $s(k)$ as follows:

$$s(k) = \begin{cases} 0 & \text{if } \bar{P}(k) \leq \eta^*, \\ 1 & \text{otherwise.} \end{cases} \quad (1)$$

The challenges are how to determine $\bar{P}(k)$ and the decision threshold η^* that minimize bit detection errors or bit error rates (BERs) of the communication. Optimization of η^* requires knowledge of probability density functions (pdfs) of $\bar{P}(k)$ conditioned on the transmitted symbol. For example, if $\bar{P}(k)$ for both '1' and '0' are assumed Gaussian distributed with mean μ_1 and standard deviation σ_1 for symbol '1' and, μ_0 and σ_0 for '0', then the optimal threshold η^* for maximum likelihood (ML) demodulation is obtained by solving the equation $f(\eta^*|\mu_1, \sigma_1^2) = f(\eta^*|\mu_0, \sigma_0^2)$, which yields [26]:

$$\eta^* = \frac{(\mu_0\sigma_1^2 - \mu_1\sigma_0^2) + \sigma_0\sigma_1\sqrt{(\mu_1-\mu_0)^2 + 2(\sigma_1^2-\sigma_0^2)ln(\frac{\sigma_1}{\sigma_0})}}{\sigma_1^2 - \sigma_0^2}. \quad (2)$$

On the other hand, if the $\bar{P}(k)$ for '0' and '1' are arbitrarily distributed, the η^* can be determined by exhaustive search.

The remaining challenge is how to determine $\bar{P}(k)$. The most basic method is to consider the average harvesting power during the entire bit interval as shown in Figure 8(a). This method, which will be referred to as BASIC from now on, is very simple to implement. However, due to

the significant guard interval, the VEH generates a long low-power tail in each bit interval, which would increase the noise floor. The increased noise floor would reduce the signal to noise ratio making it harder to correctly detect a bit symbol. To address the poor performance issue of the BASIC approach, we propose to determine $\bar{P}(k)$ as the harvesting power over a much shorter interval that contains only the peak power generation periods of the VEH. We call the proposed method the *peak power window* (PPW) method, which illustrated in Figure 8(b). As shown, PPW starts at time t_{PPW}, and with a window length of Δt. We optimize those two parameters by solving the optimization problem:

$$(t^*_{PPW}, \Delta t^*) = \underset{t^*_{PPW}, \Delta t^*}{\arg\max}\ \bar{P}(k), \tag{3}$$

subjects to the constraints that:

$$0 \leq t_{PPW} \leq T_{ramp}, \tag{4}$$

$$T_S \leq \Delta t \leq T_B, \quad \text{for } T_S \leq 50ms, \tag{5}$$

$$T_S \leq t_{PPW} + \Delta t \leq T_{ramp} + T_{damp}, \tag{6}$$

in which,

$$\bar{P}(k) = \frac{1}{\Delta t} \int_{t_{PPW}}^{t_{PPW}+\Delta t} p(t)dt, \tag{7}$$

where, $p(t)$ is the instantaneous power generated by the VEH over time t in the given bit interval T_B. We aim to maximum the average harvesting power $\bar{P}(k)$ generated by the VEH receiver during the peak power window, such that we can achieve a higher signal discriminability when compare against the noise. Constraint (4) is derived from the fact that, the power signal can increase to the peak power during the climbing duration, thus, the peak power window should start before T_{ramp}. For constraint (5), it requires the peak power window should be no shorter than the symbol time (T_S), but no longer than the bit interval (T_B). We required a lower bound of T_S to ensure sufficient amount of energy has been accumulated in the capacitor to read. Otherwise, for an extremely short Δt, the harvested energy could be too small to be read by using the capacitor. Constraint (6) requires that the peak power window should finish before the end of the damping (at time $T_{ramp} + T_{damp}$). Lastly, Equation (7) calculates the average power $\bar{P}(k)$, note that the capacitor can automatically accumulate the harvester energy within Δt.

By using brute-force search on the power trace samples, we find the averaged optimal values of t^*_{PPW} and Δt^* given different transmission time T_S. The results are averaged over 5 traces for each T_S, and stated in Table 2. Further, we can generalize the parameters as:

$$t^*_{PPW} = \begin{cases} T_{ramp} & \text{if } T_S \geq 100ms, \\ 0.46 \times T_{ramp} & \text{otherwise.} \end{cases} \tag{8}$$

and,

$$\Delta t^* = \begin{cases} T_S - T_{ramp} & \text{if } T_S \geq 100ms, \\ 0.53 \times (T_{ramp} + T_{damp}) & \text{otherwise.} \end{cases} \tag{9}$$

From the results, we can notice that, $T_S = 100$ms is a special case, for which the generated power can maintain at its peak value for a long time till the end of the transmission.

TABLE 2: The average parameters of PPW method given different transmission times.

T_S (ms)	T_{ramp} (ms)	T_{damp} (ms)	t^*_{PPW}	Δt^*
100	52	63	52	**48**
50	50	52	23	53
20	20	18	9	21
10	10	13	5	13

Thus, for $T_S = 100$ms, the lower edge of the PPW is at the end of the ramp up time where it reaches the power peak, and the upper edge of PPW is at the end of transmission, with $\Delta t^* = T_S - T_{ramp}$. For cases with $T_S \leq 50$ms, in which the power generated by VEH will decrease right after T_{ramp}, the lower and upper edge of the PPW is approximately at $0.46 \times T_{ramp}$, and $0.53 \times (T_{ramp} + T_{damp})$, respectively.

4 Performance Evaluation of VEH-COM

In this section, we present the experimental setup used for the performance evaluation of the proposed VEH-COM and analyse the BER and communication range under various environments for both the BASIC and the optimized demodulation schemes (using the peak power window with the optimized parameters). We also analyse the robustness of VEH-COM against different types of noise, including sound noise and mechanical vibration noise, which reveals that the extremely narrow power harvesting bandwidth acts as a natural protection against noise.

4.1 Experiment Setup & Data Collection

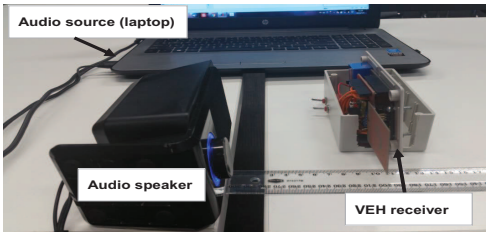


Fig. 9: Experiment setup.

Figure 9 presents the experiment setup used for performance evaluation. A laptop generates random bit streams, which are first converted to ON-OFF sound signal following the proposed VOOK modulation, and then played through a speaker connected to the laptop. As the speaker causes the table to vibrate when playing out sound, we place the VEH receiver on a different table to eliminate the effect of such mechanical vibrations. We conduct a series of experiments by varying the following parameters:

- **Distance**: the distance between the speaker and VEH receiver changes from 2cm to 100cm.
- **Data Rate**: we consider data rates from 5bps to 35bps by adjusting the transmission times in the VOOK modulation (see Table 3 for the details).
- **Sound Volume**: the speaker volume is configured to either 60dBSPL or 64dBSPL, which are comparable to a normal conversation [27].

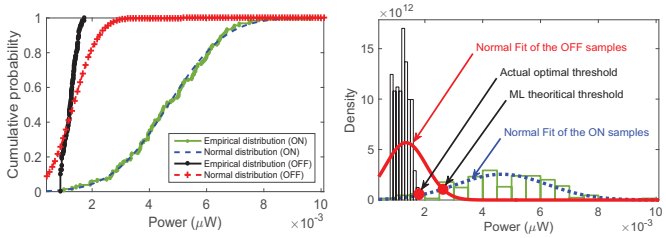
For each of these experiments, we use MATLAB to generate five random bit traces, each containing 50 bits. As each

TABLE 3: The transmission time used for different data rates.

Data Rate (bps)	5	10	20	25	30	35
Transmission Time (ms)	100	20	20	20	10	10

trace is played out, the VEH receiver records the generated AC voltage sampled at 1KHz. Therefore, we obtain 5 AC voltage traces for each experiment setting. The AC voltage traces are demodulated offline using the BASIC as well as the optimized demodulation schemes and compared against the transmitted bit streams to compute the BER.

4.2 Selection of Decision Threshold



(a) The CDF. (b) The PDF and the threshold.

Fig. 10: Example of decision threshold selection.

Before presenting the evaluation results, in this section, we compare the theoretical decision threshold obtained by applying Equation 2, and the actual decision threshold we find by using exhaustive search. Figure 10(a) gives an example of how the power samples for '1' and '0' are distributed. As it shows, comparing to the '1' samples, the '0' samples are actually not normally distributed. Consequently, as shown in Figure 10(b), the standard deviation σ_0 of '0' samples has been overestimated, which results in a large overlap between the pdfs of '1' and '0'. As indicated in Figure 10(b), the ML theoretical threshold is actually larger than the actual optimal threshold we found using exhaustive search. Thus, in the evaluation, we apply the exhaustive search to obtain a better threshold to minimize the errors.

4.3 Performance Metrics

For every experiment, we consider the following metrics:

- **Discriminability index (d_a)**: in signal detection theory [28], [29], d_a is a measure of how discriminable the signal is from the noise (similar to signal-to-noise ratio) given by:

$$d_a = \frac{\mu_S - \mu_N}{\sqrt{\frac{1}{2}(\sigma_S^2 + \sigma_N^2)}}, \tag{10}$$

  where, μ_S and σ_S, and μ_N and σ_N, are the mean and standard deviation for the signal and noise, respectively. The larger the value of d_a, the better the performance in distinguishing the modulated signal from the noise, and vice versa. For our traces, we calculate d_a by using the means and standard deviations of the average powers of '1' (signal) and '0' (noise) in the peak power window.

- **Bit Error Rate (BER)**: For a given trace, BER is obtained by dividing the number of bits in error by the total number of bits in the trace.

These two metrics, however, are highly correlated. It is obvious that the higher the discriminability, the lower the expected BER, and vice versa. As the signal power attenuates with distance, the distance between the speaker and the VEH receiver, which we will also refer to as the *communication range*, would directly influence d_a and BER. Similarly, as the signal transmission time and consequently the signal power is affected by the choice of data rates, BER is also expected to be influenced by the selected data rate.

For a target BER, performance of VEH-COM is therefore evaluated in terms of both data rates as well as the communication range. We consider two different application scenarios with different targets for data rates and communication range:

- **Mobile Payments**: widely used in stores or transportation services, in which VEH enabled mobile devices/reader could communicate within a close distance. In this scenario, we consider a short distance of 2cm, and investigate the highest possible data rate we can achieve, as the latency of a transaction is a significant performance factor for this type of application.
- **Data Sharing**: similar to Google Tone, which aims to transmit small amount of data to nearby mobile devices. Such applications can tolerate a lower data rate but expect a longer communication range. Therefore, we fix the data rate to 5bps, but explore the maximum communication range that can be achieved under different demodulation schemes.

In the following subsections, we first evaluation the performance of VEH-COM for these two applications in a quiet (low noise) office environment where the background noise mainly comes from the central air-conditioning system (HVAC). Later, we carry out additional experiments to evaluate the robustness of VEH-COM against different types of noises, including sound noise as well as physical vibration noise.

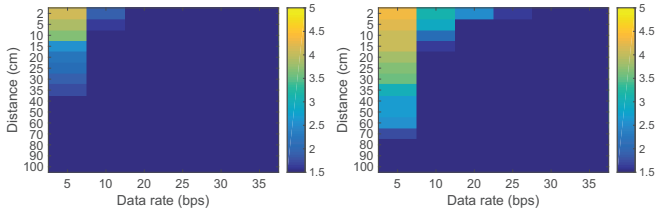
4.4 Performance in Office Environment

Figures 11 and 12, respectively, show the achieved d_a and BER as a function of the distance and data rate for two different sound volumes and modulation schemes. We consider a maximum BER of 0.01, hence do not show BERs larger than 0.01 in Figure 12. These figures clearly demonstrate that, for a given data rate and distance, higher discriminability and lower BER can be achieved by (1) increasing the speaker volume, and (2) by using the optimized demodulation, which captures the peak signal power during '1' intervals.

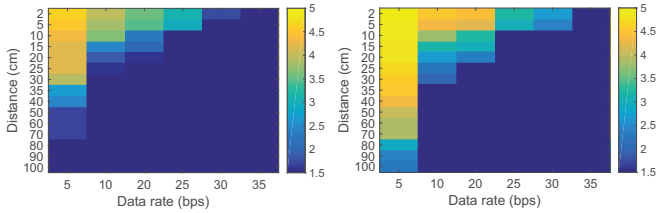
Recall that for a target BER, different applications have different performance objectives. More specifically, while Mobile Payment seeks to achieve high data rates, Data Sharing aims for longer communication range. For a target BER of 0.01, we extract the data rate and communication range performance data from Figures 11 and 12 for these two applications and summarize them in Table 4. We observe that, the optimized demodulation outperforms the Basic scheme in all cases. For mobile payment scenario, optimized scheme improves the data rate by 400% and 200%, given

TABLE 4: Performance comparison between BASIC and Optimized demodulations.

	Mobile Payment		Data Sharing	
	60dBSPL	64dBSPL	60dBSPL	64dBSPL
Basic	5bps	10bps	15cm	40cm
Optimized	25bps	30bps	30cm	80cm
Improvement	20bps (400%)	20bps (200%)	15cm (100%)	40cm (100%)

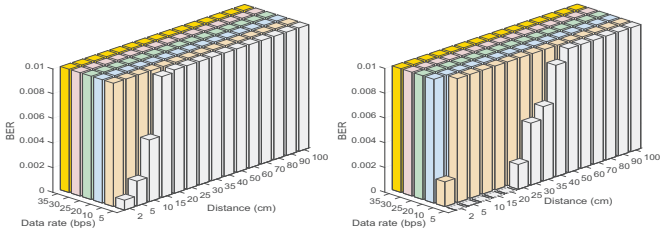


(a) Basic (60dBSPL) (b) Basic (64dBSPL)

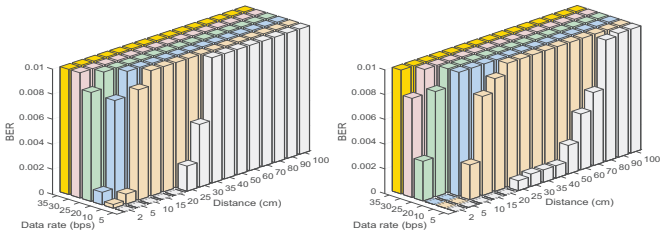


(c) Optimized (60dBSPL) (d) Optimized (64dBSPL)

Fig. 11: The heat-map shows the achieved d_a as a function of the transmission distance (Y-axis) and data rate (X-axis). The color indicates the value of d_a. (best view in color)



(a) Basic (60dBSPL) (b) Basic (64dBSPL)



(c) Optimized (60dBSPL) (d) Optimized (64dBSPL)

Fig. 12: The bar charts indicate the achieved BER (Z-axis) as a function of the transmission distance (Y-axis) and data rate (X-axis). As we assumed a BER requirement of 0.01, we apply 0.01 as cutoff bound and discard all values above it.

a volume of 60dBSPL and 64dBSPL, respectively. For data sharing scenario, it doubles the communication range that can be achieved by the Basic method under both sound volumes.

4.5 Impact of Noise

Because VEH-COM decodes the vibrations of the energy harvester to extract information, all vibration sources can potentially create noise interfering with the decoding process. In particular, we have two types of noises, acoustic noise and physical vibration noise. Acoustic noise is basically the background sound that may interfere with the vibrations of the energy harvester. Physical vibration is due to physically vibrating the receiver that houses the energy harvester. A typical source of physical vibration would be due to people holding the receiver in their hands while doing a mobile payment transaction, for example.

In this subsection, we analyze the impact of both acoustic and physical noise sources. Note that the proposed VOOK demodulation is based on distinguishing a signal from noise based only on the amount of power harvested during the bit intervals. As such, we study how different types of noise can cause vibrations for the energy harvester and how much power is harvested from each type of noise. A noise source would be considered detrimental to VOOK demodulation only if it is capable of causing large power harvesting.

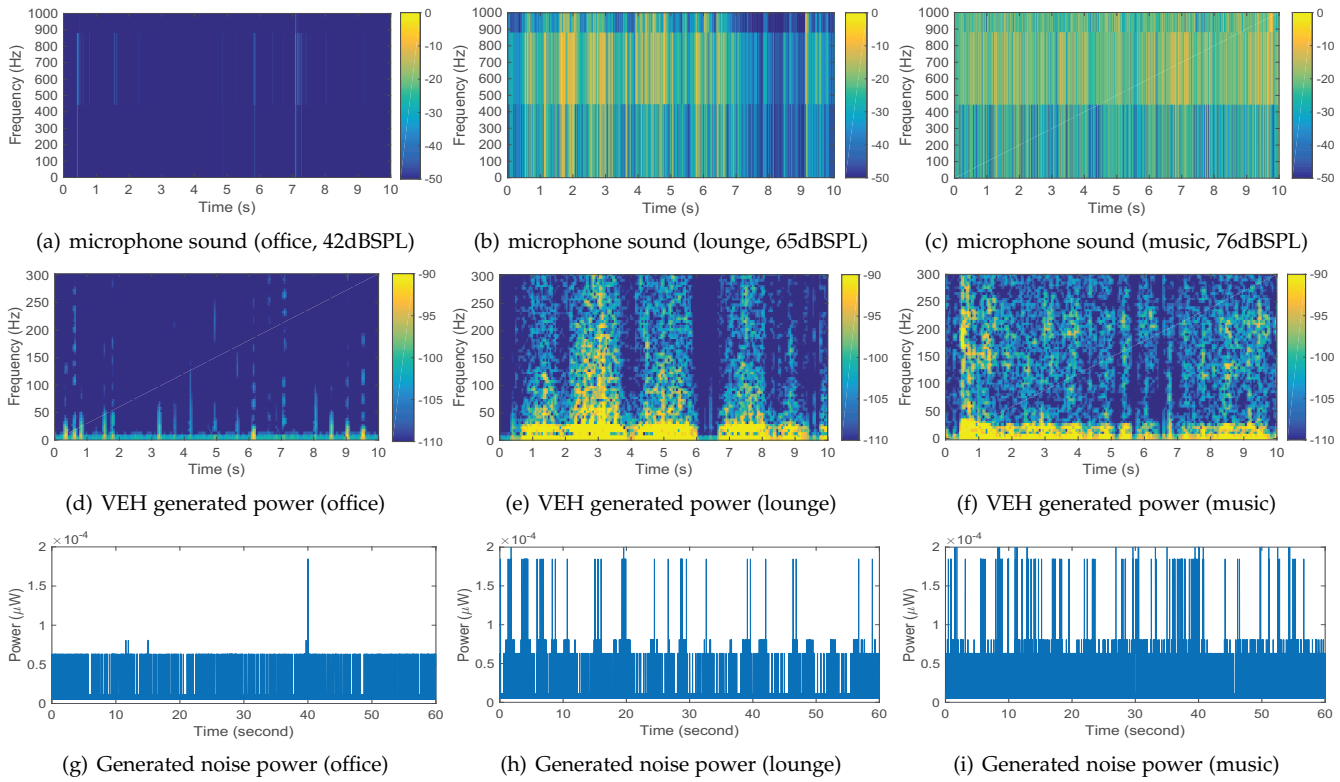
4.5.1 Acoustic Noise

We analyse three types of background acoustic noise: (1) an office environment with air condition running, (2) a lounge environment where people are chatting and laughing loudly, and 3) a nearby laptop playing music loudly. A microphone is used to record the noise sound for spectrogram analysis, and a SPL meter is used to record the noise sound level. We also record the AC voltage of the VEH for spectrogram analysis of its vibration frequencies.

We recorded noise sound levels of 42dBSPL, 65dBSPL, and 76dBSPL, respectively, for office, lounge, and music environments. Figure 13(a-c) show the spectrogram of noise sounds in these three environments. Although we can see that the noise energy is increasing with increasing sound levels in lounge and music environments, the total noise energy is actually spread over a large frequency band. In Figure 13(b-c), we can see that the noise energy is visibly more pronounced in the 450-850Hz band, which means that only a small fraction of the noise energy is contained in the VEH resonance band of 246-252Hz band. As a result, we can see in the VEH vibration spectrograms, i.e., in Figure 13(d-f), that these noise sources are unable to 'resonate' the VEH in its resonance frequencies. Instead, these high-frequency noise sources cause VEH vibrations in very low frequencies, which ultimately does not increase noise power much despite high sound levels (see Figure 13(g-i)). This analysis reveals a very interesting property of VEH-COM. It shows that the extremely narrow power harvesting bandwidth of the harvester is acting as a nature defence against typical environmental noise.

4.5.2 Noise from Human Vibration

In this experiment, the VEH receiver is held in hand while the subject is standing (simulating a mobile payment scenario). To analyze the vibration frequencies due to hand movements, Figure 14(a) plots the spectrogram of the AC voltage generated by the VEH receiver. We can notice that the vibration mostly lies in a low frequency band of 1-30Hz, which is far away from the VEH receiver's resonant band of 246-252Hz. Consequently, as shown in Figure 14(b), the generated power from these hand movements are still not



(a) microphone sound (office, 42dBSPL) (b) microphone sound (lounge, 65dBSPL) (c) microphone sound (music, 76dBSPL)

(d) VEH generated power (office) (e) VEH generated power (lounge) (f) VEH generated power (music)

(g) Generated noise power (office) (h) Generated noise power (lounge) (i) Generated noise power (music)

Fig. 13: (a-c) spectrogram of sound noise recored by microphone, for sound noises in the office (42dBSPL), lounge (65dBSPL), and music video (76dBSPL), respectively; (d-f) plot the corresponding spectrogram of the VEH harvested power from the sound noise; and (g-i) the harvested power from sound noise. (Picture view best in color)

much, although the power fluctuates and sometimes have peaks. The amplitude of the harvested power from such noise (during '0' bit interval) relative to the power harvested during a transmission of '1' is shown in Figure 14(b) for a transmission of 10 bits of data. It is clear that although the noise is dynamic in this case due to unpredictable hand movements (some '0' intervals have higher noise than others), they are all much lower than the power harvested during '1' intervals.

4.5.3 VEH-COM Performance in Noisy Environment

For the optimized demodulation, Figure 15 analyzes the effect of increasing noise level on BER performance. In the experiment, the sound volume, i.e., the volume of transmitted sound for bit '1', is fixed to 60dBSPL which is the lowest volume we considered. We can observe from the results that the system usability does not affect much by the environmental noise. Although, the values of d_a have slightly decreased in the lounge due to the increase of the noise floor, our system can still achieve a BER lower than 0.01 within 10cm distance for both 5 and 10 bps data rates. On the other hand, for 15 and 20 cm distances, with the weaker received signal power, the achieved BER in the lounge exceeds the 0.01 mark. However, this problem can be solved by increasing the sound volume.

5 Power Saving Opportunity with VEH-COM

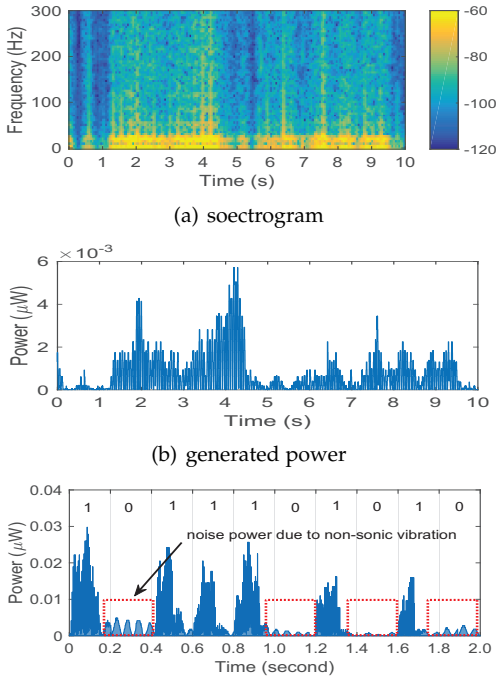
The power consumption of a microphone-based communication receiver mainly comes from audio sampling and FFT.

TABLE 5: Specifications of the smartphones used.

	CPU	Co-processor	RAM	Android version
HTC One X	Cortex-A9	No	1GB	v4.0.1
Moto E2	Cortex-A7	Yes	1GB	v5.0
Samsung S4	Cortex-A7	Yes	2GB	v4.4.2

In this section, we investigate the power consumptions of microphone-based audio sampling and FFT computations to gain insight to the potential power saving opportunity with the proposed system. We implement two smartphone applications, one for continuously sampling the microphone and the other for FFT calculations only. We run these applications at different times separately and measure the corresponding power consumptions by using the PowerTutor [30], which is widely used for android applications power measurement [31], [32], [33]. To avoid bias to any particular hardware, we conduct power analysis on three different types of smartphones (see Table 5).

Power consumption of audio sampling depends on the sampling rate. Similarly, FFT power consumption depends on the FFT size, which determines the frequency resolution of FFT obtained as the sampling rate divided by the FFT size. We repeat our experiments for many different sampling rates and FFT sizes to gain a comprehensive understanding of the power consumption. Figure 16(a) shows the power consumption of microphone sampling between 8KHz and 44.1KHz, which is the valid range of sample rates supported by most smartphones. We can observe that the power consumption increases slowly with sampling rate at the beginning, but after 22KHz it increases very rapidly. Given

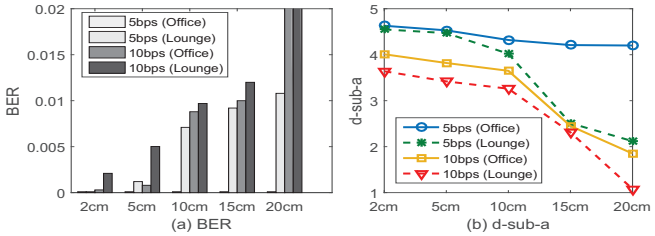


(a) soectrogram

(b) generated power

(c) generated power for data under non-sonic noise.

Fig. 14: (a) the spectrogram of the harvested power by VEH receiver from the non-sonic vibration; (b) the harvested power; and (c) the power generated by VEH receiver from the 10 bits sound signal under the noise from the non-sonic vibration.



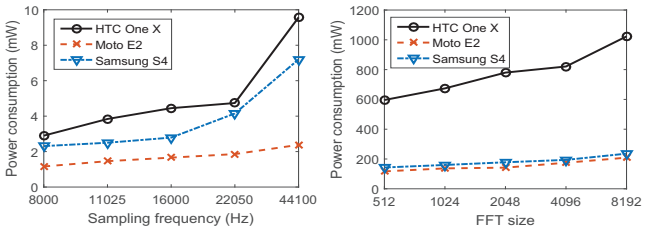
(a) BER (b) d-sub-a

Fig. 15: System performance in different environments.

that most of the microphone-based communication systems choose to sample at 44100Hz, the power consumption for sampling could be *several milliwatts*. The important observation, however, is that the FFT consumes power in the order of *several hundred milliwatts* as shown in Figure 16(a)(b), which plots power consumption as a function of FFT size for a sampling rate of 44.1Hz. FFT therefore consumes the Lion's share of the power and hence VEH-COM promises significant power saving as it can achieve demodulation without FFT. We should, however, note that VEH-COM, as implemented in this paper, achieves only modest data rates compared to microphone-based systems that can take advantage of a much larger communication bandwidth. How to improve the data rates for VEH-COM will be the focus of future works.

6 Related Work

In light of the ubiquity of microphone-enabled devices, researchers are now working on speaker-microphone based audio communications [5], [6], [7], [8], [9], [10], [34], [35],



(a) Sampling Microphone. (b) Running FFT.

Fig. 16: Power consumption analysis.

but all these works apply DSP to demodulate data from microphone samples. In [6], the authors proposed a ultrasound chirp-based communication system using commercial speakers. The system is able to transmit 16bps data up to 25m distance. A more recent work in [5], the authors developed a near-ultrasound chirp-based system as the second screen service between TVs and smart devices. Some researchers did consider mechanical vibrations as a source for low bit-rate data communication [10], [36], but they used the vibration motor as a transmitter and the decoding at the receiver is achieved by applying DSP.

To the best of our knowledge, there has been no prior work on using VEH as a potential receiver for audio communications. The only prior work related to VEH-based sound detection is our *own work* [37], in which we have demonstrated that VEH can be used to detect the utterance of the phrase "OK Google" from a close distance. Earlier before, we have also demonstrated that VEH can be used to detect a range of human activities and contexts, including activity recognition [38], [39], step counting [40], estimating calorie burning [41], transportation mode detection [42], and user authentication [43].

7 Conclusion and Future Work

In this paper, we explore the feasibility of using the VEH device as a communication receiver. By modulating the ambient vibration energy using audio speaker, and demodulating the power harvested from the acoustic sound by the VEH receiver, our system is able to achieve 30bps data rate within a short range. Moreover, VEH-COM is able to transmit as far as 80 cm for 5bps, which allows many device-to-device data sharing applications. While the achieved data rates in our current implementation is modest compared to microphone-based systems, we believe that the date rate can be improved with further research. A key advantage of VEH-COM is its potential power saving by not requiring complex FFT for demodulation used in conventional microphone-based system. This new explored capability of VEH therefore will serve as the building block for future VEH-based communication technologies and applications.

Current work can be extended in many directions. VEH-COM uses audible frequencies which may be annoying to some users. One may design solutions that embed signals in music/audio so that the users do not notice them. Actively dampening the VEH by transmitting specifically designed 'out of phase' signals, which may help improving the data rates, is another future direction. Designing *physical layer security* solutions for VEH-COM to resist eavesdropping at

some distance would be useful to achieve high level of security for applications such as mobile payments. It is also useful to analyse the effect of physical parameters of the VEH, such as area, size, resonance frequency, etc., on the communication performance. Finally, evaluating the technology in real application scenarios would be an ultimate exercise.

REFERENCES

[1] "Google tone," https://chrome.google.com/webstore/detail/google-tone/nnckehldicaciogcbchegobnafnjkcne, accessed: 2016-09-23.

[2] "Slicklogine," https://en.wikipedia.org/wiki/SlickLogin, accessed: 2016-09-23.

[3] "shopkick," https://www.shopkick.com/, accessed: 2016-09-23.

[4] "Silverpush," https://www.silverpush.co/, accessed: 2016-09-23.

[5] S. Ka, T. H. Kim, J. Y. Ha, S. H. Lim, S. C. Shin, J. W. Choi, C. Kwak, and S. Choi, "Near-ultrasound communication for tvs 2nd screen services," in *Proceedings of MobiCom*. ACM, 2016.

[6] H. Lee, T. H. Kim, J. W. Choi, and S. Choi, "Chirp signal-based aerial acoustic communication for smart devices," in *Proceedings of INFOCOM*. IEEE, 2015, pp. 2407–2415.

[7] H. Matsuoka, Y. Nakashima, and T. Yoshimura, "Acoustic communication system using mobile terminal microphones," *NTT DoCoMo Tech. J*, vol. 8, no. 2, pp. 2–12, 2006.

[8] R. Nandakumar, K. K. Chintalapudi, V. Padmanabhan, and R. Venkatesan, "Dhwani: secure peer-to-peer acoustic nfc," in *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4. ACM, 2013, pp. 63–74.

[9] G. E. Santagati and T. Melodia, "U-wear: Software-defined ultrasonic networking for wearable devices," in *Proceedings of MobiSys*. ACM, 2015, pp. 241–256.

[10] N. Roy and R. R. Choudhury, "Ripple ii: faster communication through physical vibration," in *Proceedings of NSDI*. USENIX, 2016, pp. 671–684.

[11] Q. Wang, K. Ren, M. Zhou, T. Lei, D. Koutsonikolas, and L. Su, "Messages behind the sound: real-time hidden acoustic signal capture with smartphones," in *Proceedings of MobiCom*. ACM, 2016, pp. 29–41.

[12] S. P. Beeby, R. Torah, M. Tudor, P. Glynne-Jones, T. O'Donnell, C. Saha, and S. Roy, "A micro electromagnetic generator for vibration energy harvesting," *Journal of Micromechanics and microengineering*, vol. 17, no. 7, p. 1257, 2007.

[13] S. Roundy, P. K. Wright, and J. Rabaey, "A study of low level vibrations as a power source for wireless sensor nodes," *Computer communications*, vol. 26, no. 11, pp. 1131–1144, 2003.

[14] J. Hester, L. Sitanayah, and J. Sorber, "Tragedy of the coulombs: Federating energy storage for tiny, intermittently-powered sensors," in *Proceedings of SenSys*. ACM, 2015, pp. 5–16.

[15] P. D. Mitcheson, E. M. Yeatman, G. K. Rao, A. S. Holmes, and T. C. Green, "Energy harvesting from human and machine motion for wireless electronic devices," *Proceedings of the IEEE*, vol. 96, no. 9, pp. 1457–1486, 2008.

[16] J. Yun, S. Patel, M. Reynolds, and G. Abowd, "A quantitative investigation of inertial power harvesting for human-powered devices," in *Proceedings of UbiComp*. ACM, 2008, pp. 74–83.

[17] M. Gorlatova, J. Sarik, G. Grebla, M. Cong, I. Kymissis, and G. Zussman, "Movers and shakers: Kinetic energy harvesting for the internet of things," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 8, pp. 1624–1639, 2015.

[18] L.-Y. Wu, L.-W. Chen, and C.-M. Liu, "Acoustic energy harvesting using resonant cavity of a sonic crystal," *Applied Physics Letters*, vol. 95, no. 1, p. 013506, 2009.

[19] W.-C. Wang, L.-Y. Wu, L.-W. Chen, and C.-M. Liu, "Acoustic energy harvesting by piezoelectric curved beams in the cavity of a sonic crystal," *Smart Materials and Structures*, vol. 19, no. 4, p. 045016, 2010.

[20] N. Jackson, F. Stam, O. Z. Olszewski, R. Houlihan, and A. Mathewson, "Broadening the bandwidth of piezoelectric energy harvesters using liquid filled mass," *Procedia Engineering*, vol. 120, pp. 328–332, 2015.

[21] D. Berdy, B. Jung, J. Rhoads, and D. Peroulis, "Increased-bandwidth, meandering vibration energy harvester," in *Proceedings of the 16th International Solid-State Sensors, Actuators and Microsystems Conference*. IEEE, 2011, pp. 2638–2641.

[22] A. Erturk and D. J. Inman, *Piezoelectric energy harvesting*. John Wiley & Sons, 2011.

[23] R. Vullers, R. van Schaijk, I. Doms, C. Van Hoof, and R. Mertens, "Micropower energy harvesting," *Solid-State Electronics*, vol. 53, no. 7, pp. 684–693, 2009.

[24] T. Galchev, J. McCullagh, R. Peterson, and K. Najafi, "Harvesting traffic-induced vibrations for structural health monitoring of bridges," *Journal of Micromechanics and Microengineering*, vol. 21, no. 10, p. 104005, 2011.

[25] "Mide," http://www.mide.com/, accessed: 2016-09-14.

[26] Z. Tian and B. M. Sadler, "Weighted energy detection of ultra-wideband signals," in *IEEE 6th Workshop on Signal Processing Advances in Wireless Communications, 2005.* IEEE, 2005, pp. 1068–1072.

[27] "Sound pressue level table," http://www.sengpielaudio.com/TableOfSoundPressureLevels.htm., accessed: 2016-09-04.

[28] H. Stanislaw and N. Todorov, "Calculation of signal detection theory measures," *Behavior research methods, instruments, & computers*, vol. 31, no. 1, pp. 137–149, 1999.

[29] A. J. Simpson and M. J. Fitter, "What is the best index of detectability?" *Psychological Bulletin*, vol. 80, no. 6, p. 481, 1973.

[30] L. Zhang, B. Tiwana, Z. Qian, Z. Wang, R. P. Dick, Z. M. Mao, and L. Yang, "Accurate online power estimation and automatic battery behavior based power model generation for smartphones," in *Proceedings of the eighth IEEE/ACM/IFIP international conference on Hardware/software codesign and system synthesis*. ACM, 2010, pp. 105–114.

[31] Y. Shen, W. Hu, M. Yang, B. Wei, S. Lucey, and C. T. Chou, "Face recognition on smartphones via optimised sparse representation classification," in *Proceedings of IPSN*. IEEE, 2014, pp. 237–248.

[32] W. Xu, G. Revadigar, C. Luo, N. Bergmann, and W. Hu, "Walkie-talkie: Motion-assisted automatic key generation for secure on-body device communication," in *Proceedings of IPSN*. IEEE, 2016, pp. 1–12.

[33] W. Xu, Y. Shen, N. Bergmann, and W. Hu, "Sensor-assisted face recognition system on smart glass via multi-view sparse representation classification," in *Proceedings of IPSN*. IEEE, 2016, pp. 1–12.

[34] B. Zhang, Q. Zhan, S. Chen, M. Li, K. Ren, C. Wang, and D. Ma, "Enabling keyless secure acoustic communication for smartphones," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 33–45, 2014.

[35] A. S. Nittala, X.-D. Yang, S. Bateman, E. Sharlin, and S. Greenberg, "Phoneear: interactions for mobile devices that hear high-frequency sound-encoded data," in *Proceedings of EICS*. ACM, 2015, pp. 174–179.

[36] N. Roy, M. Gowda, and R. R. Choudhury, "Ripple: Communicating through physical vibration," in *Proceedings of NSDI*. USENIX, 2015, pp. 265–278.

[37] S. Khalifa, M. Hassan, and A. Seneviratne, "Feasibility and accuracy of hotword detection using vibration energy harvester," in *Proceedings of WoWMoM*. IEEE, 2016.

[38] S. Khalifa, M. Hassan, and A. Seneviratne, "Pervasive self-powered human activity recognition without the accelerometer," in *Proceedings of PerCom*. IEEE, 2015, pp. 79–86.

[39] S. Khalifa, M. Hassan, A. Seneviratne, and S. K. Das, "Energy-harvesting wearables for activity-aware services," *IEEE Internet Computing*, vol. 19, no. 5, pp. 8–16, 2015.

[40] S. Khalifa, M. Hassan, and A. Seneviratne, "Step detection from power generation pattern in energy-harvesting wearable devices," in *2015 IEEE International Conference on Data Science and Data Intensive Systems*. IEEE, 2015, pp. 604–610.

[41] G. Lan, S. Khalifa, M. Hassan, and W. Hu, "Estimating calorie expenditure from output voltage of piezoelectric energy harvester: an experimental feasibility study," in *Proceedings of the 10th EAI International Conference on Body Area Networks*. ICST, 2015, pp. 179–185.

[42] G. Lan, W. Xu, S. Khalifa, M. Hassan, and W. Hu, "Transportation mode detection using kinetic energy harvesting wearables," in *Proceedings of PerCom Work-in-Progress*. IEEE, 2016, pp. 1–4.

[43] W. Xu, G. Lan, Q. Lin, S. Khalifa, N. Bergmann, M. Hassan, and W. Hu, "Keh-gait: Towards a mobile healthcare user authentication system by kinetic energy harvesting," in *Proceedings of NDSS*, 2017.