Towards Physical-Layer Vibration Sensing with RFIDs

Ping Li*†, Zhenlin An‡, Lei Yang‡, Panlong Yang*
*School of Computer Science and Technology, University of Science and Technology of China
†College of Communication Engineering, Army Engineering University of PLA
‡Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong
Email: pingli0112@gmail.com, {an, young}@tagsys.org, plyang@ustc.edu.cn

***Abstract*—Conventional vibration sensing systems, equipped with specific sensors (*e.g.*, accelerometer) and communication modules, are either expensive or cumbersome in deployment. In recent years, the community revisits this classic topic by taking advantage of off-the-shelf RFIDs. However, limited by lower reading rate and larger wavelength, current RFID based solutions can only sense low-frequency (*e.g.* below 100Hz) mechanical vibrations with larger amplitude (*e.g.* $>$ 5mm). To address this issue, this work presents TagSound, an RFID-based vibration sensing system that explores a tag's harmonic backscattering to recover high-frequency and tiny mechanical vibrations accurately. The key innovations are in two aspects: *harmonics based sensing* and *a new recovery scheme*. We implement TagSound with USRP platforms. Our comprehensive evaluation shows TagSound can achieve a mean error of 0.37Hz when detecting vibrations at frequencies below 100Hz, and a mean error of 4.2Hz even when the vibration frequency is up to 2500Hz.**

I. Introduction

Vibration is a common mechanical phenomenon whereby oscillations occur in the vicinity of an equilibrium point. In some time, vibration is useful. For example, the vibration of loudspeaker amplifies the sound, and the vibrating spear makes the concrete bonded densely. In many cases, however, vibration is undesirable and must be detected accurately. For instance, the unexpected downtime due to the undesirable vibration of rotating machineries has become more costly and ever before [1]. The common method of diagnosing such abnormalities is to detect the frequency of vibration.

Many efforts have been made to vibration sensing in past decades. Unfortunately, these traditional approaches all require to equip with specialized sensors, *e.g.* accelerometers, gyroscopes, lasers, cameras, *etc.* They are bulky, heavy, expensive and energy-consuming. For example, accelerometers and gyroscopes require wiredly connection to a control panel for power supply and signal communication. In particular, lasers [2] and high-speed cameras [3] have been demonstrated as attractive solutions for high-resolution and high-speed vibrations. Yet, they fail in none Line-of-Sight (NLOS) scenarios (*e.g.* inspecting tubes in centrifugal machine) since either infrared ray or visible light cannot penetrate middle obstacles.

Recently, the wireless community revisits vibration measurement through *wireless vibrometry* that explores wireless signals to make sense of the vibration. For example, [4] uses WiFi signals to track the human's vital signals, like breathing and heartbeats, which can be viewed as minute-level vibrations. [5] achieves the same goal with an RFID tag array. ART [6] eavesdrops through-wall loudspeakers remotely through the signal changes caused by vibrations. Most interestingly, Tagbeat [7] introduces RFID technology to the vibration sensing with the advantage of being cost-effective, applicable to occluded objects, and auto-associative with the spinning object (by the tag's ID). The follow-up work, TagTwins [8], further applies dual-tag solution in noisy environment.

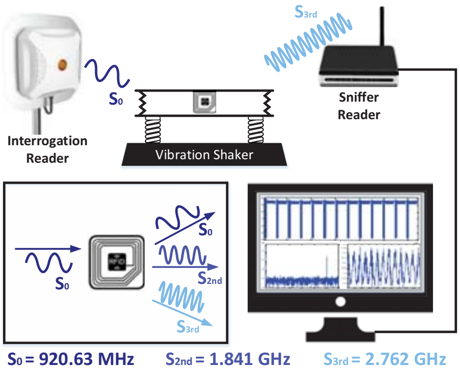


Fig. 1: TagSound Architecture

The core concept underlying vibration sensing using RFID is to inspect the vibration through the random and low-frequency readings of tag, where each reading is viewed as one sampling of the vibration. Vibration displaces the tag attached on the vibrating surface within a small range, resulting in a regular change pattern of backscatter signals. We can reveal the relevant vibration information by discerning such communication pattern without specialized sensors. Past RFID based solutions offer many attractive merits, however, still suffer from two major deficiencies:

- The first deficiency is from the limitation on reading rate. The key idea of vibration sensing is to detect the phase changes of the wireless signals backscattered from tags, which highly depends on the sampling rate (or called reading rate in RFID). The current typical RFID system can read tags about 70 times per second with a single antenna. According to the Nyquist sampling law, the upper bound of the frequency that the vibration signal contains must be less than 35Hz. Apparently, this cannot meet the demand of most cases in practice. Tagbeat [7] makes efforts to address this issue through compressive reading. However, the solution only works for the *periodic* vibrations, whose frequency-domain representations are extremely sparse. On the contrary, the *general solution* working for any kind of viberation is never studied yet.

- The second deficiency is from the minimum requirement on the amplitude of the vibration. RFID based solution usually discerns vibration through the phase, whose value changes as the vibration-induced displacement. The empirical study suggests that the phase would be drawn in thermal noise if its value falls below 0.1 radians. Suppose the frequency of carrier is 900MHz (*i.e.* the wavelength, $\lambda = 33$cm), the phase becomes detectable (> 0.1 radian) only when the tag is displaced by $33 \times 0.1/2\pi = 5.2$mm at least. Clearly, this threshold (*i.e.* 5.2mm) depends on the wavelength or the frequency of the carrier. Nevertheless, majority of practical vibrating exciter vibrates below this threshold [9], especially for high-frequency vibration.

This work introduces *TagSound*, an RFID-based vibration sensing system that pushes the sensing on mechanical vibration to the limit. It can measure high-frequency (*e.g.* $>$ 1kHz) and tiny vibrations (*e.g.* vibrating amplitude $<$ 2mm). TagSound fights off the above two issues as follows:

- Unlike the past solutions, which work at application layer across the readings reported by readers, TagSound operates at physical layer and directly extracts vibration information from physical-layer signals acquired form the air. It considers baseband signals instead of readings as the vibration signals. Consequently, the upper bound of perceptible vibration frequency is improved from a half of 70Hz (*i.e.* reading rate) to a half of 2MHz (*i.e.* ADC sampling frequency). Most importantly, TagSound can sense any kinds of vibrations without the assumption on their periodicities.
- The recent finding demonstrates that the tag backscatters not only at fundamental frequencies (*i.e.* 900MHz) but also at harmonics (*i.e.* $2 \times 900 = 1.7$GHz, $3 \times 900 = 2.7$GHz) due to the nonlinearity effect of rectifier. TagSound explores this underlying hardware property to fight off the second issue. Intuitively, the third-order harmonics at 2.7GHz has 11cm wavelength and thereby lowers the vibration threshold to $11 \times 0.1/2\pi = 1.7$mm, which is $1/3$ of the previous.

Fig. 1 shows the scenario of TagSound, where an RFID tag is attached on the vibrating surface. A reader is continuously interrogating the tag. Meanwhile, TagSound employs a sniffer-purpose reader nearby the tag to acquire the backscattering harmonics that carries vibration signals. TagSound extracts the vibration signal from the sniffed harmonics.

Transforming the above idea into a practical system, however, faces several challenges as follows: (*i*) the first challenge is from the carrier frequency offset (CFO). The today's commercial reader does not expose its physical layer to us. We have to employ a software defined reader (SDR) to acquire the harmonics, and meanwhile interrogate the tag by a commercial reader for the compatibility. Since two readers have different internal clock drifts, the sniffer reader experiences a CFO and sampling frequency offset with respect to the interrogation reader. The CFO causes an additional phase rotation, independent of the vibration, which unless eliminated can be accumulated over time. (*ii*) The second challenge comes from the modulation method of RFID. Since tag uses the On-Off keying (OOK) modulation, there are two clusters on the constellation diagram. As a result, the phase values jumps between two values. (*iii*) The third challenge is the drastic attenuation of harmonics. According to [10], the power of the third harmonic signals employed in TagSound is about 60dB lower than the fundamental frequency signal. Such weak signal is vulnerable to the ambient noise.

Contributions: To best of our knowledge, TagSound is the first work that explores the harmonics for the vibration sensing, pushing both target frequency and amplitude to the limit. In this work, three key innovations are made: first, we verify the harmonics induced from the tag's backscattering; second, we explore the CFO at harmonics when sniffing the vibration; third, we develops a prototype over the commercial reader and USRP N210, and evaluate it in real world.

II. Primer on Harmonic Backscattering

Since passive RFID tags do not carry power source, they harvest energy from the high-energy continuous wave transmitted from the RFID reader. To this end, a passive tag contains *rectifier*, consisting of many diodes, to convert the alternating current induced by the continuous wave (CW) to a direction current, thereby providing energy for the other parts (*e.g.* chip).

The nonlinearity effect of these diodes produces *harmonics signals* in addition to the *fundamental signal* [11]. Due to the absence of harmonics suppression in tags, the harmonic signals are also reflected to the air with the backscatter signals at fundamental frequency. Formally, suppose the input reader's single-tone CW is denoted by $S = \cos(2\pi f t)$. Then, the backscattered signals denoted by S_{out} is given by:

$$\begin{aligned} S_{\text{out}} =& A_1 \cos(2\pi t) + A_2 \cos^2(2\pi f t) + A_3 \cos^3(2\pi f t) + \cdots \\ =& \frac{1}{2}A_2 + (A_1 + \frac{3}{4}A_3)\underbrace{\cos(2\pi f t)}_{\text{1st-order}} + \frac{1}{2}A_2 \underbrace{\cos(2\pi(2f)t)}_{\text{2nd-order}} + \\ & \frac{1}{4}A_3 \underbrace{\cos(2\pi(3f)t)}_{\text{3rd-order}} + \cdots \end{aligned} \tag{1}$$

where A_k are the gains of the various components introduced by the rectifier. This equation indicates that the harmonic backscattering contains multiple frequency components, and the frequencies of each component are positive integral multiple of the fundamental frequency. Translating it to actual numbers, when $f = 920$ MHz, the backscattered signals could be detected at 920MHz, 1.84GHz, 2.76GHz, and 3.68 GHz, and so on.

Feasibility Study. We perform a group of empirical experiments to verify the feasibility of harmonic backscattering. In this experiment, an ImpinJ RFID reader (R420) is used to communicate with the test tag at 920.63MHz. A Keysight oscilloscope (MSOS404A)[12] is used to monitor the changes of the spectrum over the band of $0 \sim 5$GHz. In practice, we also observe harmonics leaked from the transmitter of the reader, which is supposed to be suppressed well but actually not. As an active device, these reader-induced harmonics may cause interferences to other electronic products. To ensure that the harmonic backscattering from tags is truly observed, we

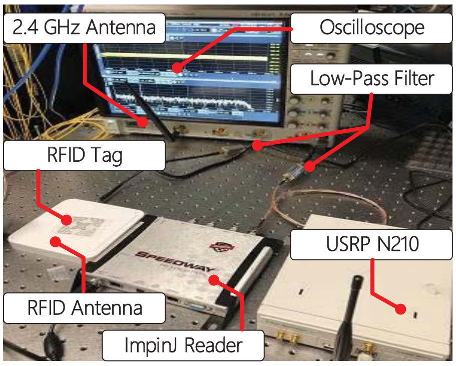


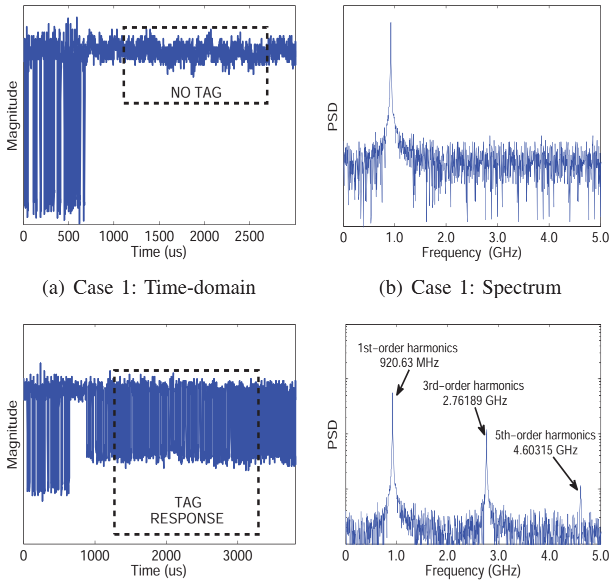
Fig. 2: Harmonic measurement setup

add a custom-made low-pass filter between the reader and its antenna. The experimental setup is shown in Fig.2. We acquire signals in two cases and obtain the following findings:

- **Case 1**: We intercept a sequence of the signals during the time window for a tag reply, without placing tag in front of the reader. Fig. 3(a) shows the time-domain signals. It can be seen from the figure that there still exist signals during the window even when no tag is present. This is because the reader's CW presents all the time. Fig. 3(b) shows the spectrum of the intercepted signals. A single spike at 920.63MHz is found in the spectrum because of the existence of CW.
- **Case 2**: We then intercept a sequence of the signals when a tag backscatters during the window. The time-domain signals are shown in Fig. 3(a), from which the changes of amplitude are greater than Case 1 due to the tag's reply. Fig. 3(d) shows the corresponding spectrum. In theory, the tag modulates its data by reflecting the CW, so no additional frequencies should be created in the reply. However, two additional spikes clearly appear at 2.76189GHz and 4.603GHz, which are exactly the 3× and 5× of 920.63MHz.

The above two case studies affirmatively confirm that the tag exactly backscatters at harmonic frequencies due to the nonlinearity of the rectifier. Our observations are consistent with the previous results reported in [13]. However, we cannot clearly observe the even-order harmonics. This is because the circuits of tag are more apt to create odd-order harmonics and the even-order harmonics are too weak to be detected. We refer [10] for more explanations.

These backscattered harmonics will not interfere the reader's receiver because creating a filter to reject signals above fundamental frequency is easy. This is also the reason why we cannot observe influence of the harmonics in the application layer when using commercial readers. Thus, we set up a nearby software defined reader to acquire the harmonics. TagSound, we choose the third-order harmonics as our media to sense the vibration since it is the strongest harmonics except the first-order. Finally, we would like to emphasize the reason why the third-order harmonics are explored for vibration sensing: its wavelength is far shorter than that of the fundamental frequency, so the it is much more sensitive to the displacement.



(a) Case 1: Time-domain

(b) Case 1: Spectrum

(c) Case 2: Time-domain

(d) Case 2: Spectrum

Fig. 3: Feasibility study on harmonic backscattering. (a) shows the time-domain signal of Case 1 where no tag reply is present; (b) shows the spectrum of the signals in Case 1; (c) shows the time-domain signals of Case 2 where a tag reply is present; (c) shows the spectrum of the signals in Case 2.

III. TagSound Design

TagSound is an RFID-based universal solution for inspecting vibration frequency of any objects. Although we present the system in the context of speaker in most of the time, TagSound technique could be applied to any modalities.

A. Problem Formalization

Suppose that a tag is attached on the vibrating surface. The tag's backscattered signal remains relatively stable without vibration; otherwise, it fluctuates as the vibrating. Formally, let $S(t)$ denote the harmonic signal backscattered from the tag. $S(t)$ is defined as follows:

$$S(t) = A(t)e^{\mathbf{J}(2\pi f_c t + 2\pi d/\lambda + \theta_0)} \tag{2}$$

where λ is the wavelength, d is the distance between the tag and the sniffer reader, and θ_0 is the initial phase of carrier signal. Since the initial phase θ_0 is a constant that makes no difference to our analysis, we omit this variable for clarity. f_c is the central frequency of the harmonic signal. $A(t)$ is also a complex that represents the reply data of the tag. Since the tag uses OOK to modulate its data, $|A(t)|$ follows into two states for representing bit '0' or bit '1' respectively.

Following the prior work [1], any kind of vibration can be abstracted into a simple harmonic motion. Further, the distance $d(t)$ between the sniffer reader and the tag is expressed as follows:

$$d(t) = d_0 \cos(2\pi f_v t) \tag{3}$$

where d_0 is the initial distance without vibration. f_v is the vibration frequency, *i.e.* the key parameter that we need to measure in practice. Substituting Eqn. 3 into Eqn. 2, we obtain the new form of the harmonic signal, which carries the vibration information as follows:

$$\begin{aligned} S(t) =& A(t)e^{\mathbf{J}(2\pi f_c t + 2\pi d(t)/\lambda)} \\ =& A(t)e^{\mathbf{J}(2\pi f_c t + 2\pi d_0 \cos(2\pi f_v t)/\lambda)} \\ =& A(t)e^{\mathbf{J}(2\pi f_c t + v(t))} \end{aligned} \tag{4}$$

where

$$v(t) = 2\pi d_0 \cos(2\pi f_v t)/\lambda \tag{5}$$

The $v(t)$ is considered as the *vibration signal*, which exactly characterizes the corresponding vibration. Meanwhile, we also define the complex representation of the vibration signal, denoted by $V(t)$, as below:

$$V(t) = e^{\mathbf{J}v(t)} \tag{6}$$

Putting them together, the final backscatter signal acquired by the sniffer reader is given by:

$$S(t) = A(t)e^{\mathbf{J}2\pi f_c t}e^{\mathbf{J}v(t)} = A(t)V(t)e^{\mathbf{J}(2\pi f_c t)} \tag{7}$$

This equation shows that two types of data are carried on the backscatter signal, *i.e.* $e^{\mathbf{J}(2\pi f_c t)}$, which centers at f_c. The data $A(t)$ comes from the tag and is modulated with OOK; the data $V(t)$ is from the vibration and modulated with PSK. Therefore, the final signal is modulated with Amplitude and Phase-Shift Keying (APSK). Our objective is to extract $V(t)$ or $v(t)$ from $S(t)$. Namely, our problem is formalized as follows:

Definition 1 (Vibration Problem)**.** *Given a backscatter signal $S(t)$ acquired by a sniffer reader, how could we demodulate the vibration signal $v(t)$ from $S(t)$?*

B. Solution Sketch

In this work, we propose a holistic solution, TagSound, to address the vibration problem. Querying the RFID tag attached on the vibrating surface, at a high level TagSound adopts a pipeline architecture for the signal processing. Briefly, it contains three components:

- *Persistent vibration modulation*: TagSound utilizes the retransmission mechanism defined in Gen2 protocol to acquire persistent modulation of the vibration.
- *Compensating frequency offset*: TagSound eliminates the additional phase rotations that are caused by CFO between the interrogation and the sniffer readers.
- *Extracting phase*: TagSound separates the phase values from the baseband signals downconverted from harmonic backscattering with CFO compensation.

The next few sections elaborate on the above components, providing the technical details.

IV. Persistent Vibration Modulation

An RFID reader uses a variant of framed ALOHA, called *Q-adaptive*, as its anti-collision protocol in link layer. For the sake of compatibility, TagSound also uses this protocol even though only a single tag is attached on the vibrating surface. Fig. 4 shows the anti-collision procedure, which highlights the long reply phase. Specifically, the reader divides time into several slots, which are further organized into frames. The reader starts a frame by broadcasting a `Query` command, which contains the parameter of frame length f. After receiving `Query`, each tag randomly selects an integer $\in [0, k-1]$ and stores it in the variable of `SC`. Afterwards, the reader starts a time slot by broadcasting `QueryRep`, making the tag decrease its `SC` by one. If the `SC` of the tag is equal to 0, then it immediately replies with a 16-bit random signal (*i.e.* `RN16`) for collision detection and avoidance. If only one tag replied in the slot, then the reader must acknowledge its reply with the `ACK` command. The acked tag immediately transmits a *long reply* including `PC` (Protocol Control), `EPC` and `CRC`; otherwise, the reader proceeds to the next time slot. In the end of each frame, if any collision slot exists, the reader needs to start a new frame for these tags, which chosen the collision tags, by issuing the `QueryAdjust` commands. This whole procedure repeats until no collision slots are found.

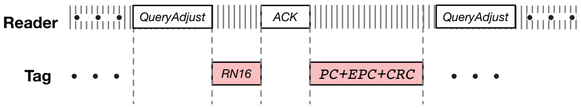


Fig. 4: Anti-collision procedure

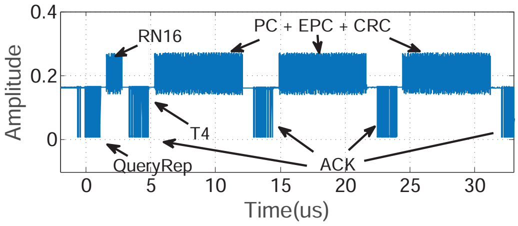
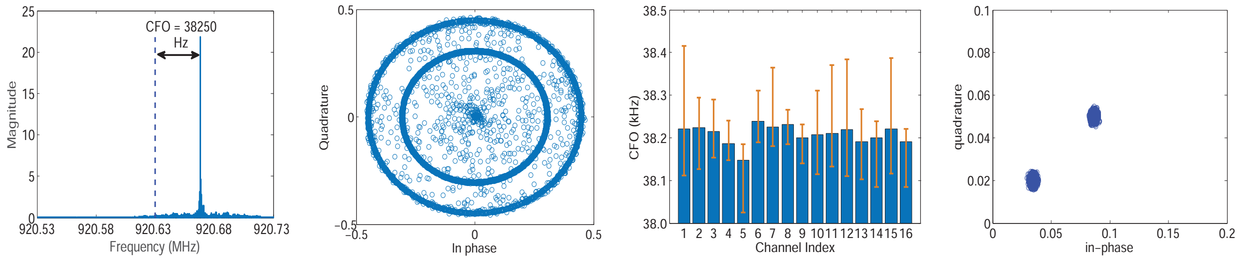


Fig. 5: Retransmission based modulation

Harmonic backscattering occurs only when the tag transmits its reply. Clearly, there are a large number of time intervals for broadcasting reader commands, such as `Query`, `QueryRep`, `ACK`, and so on. During these time, the tag does not backscatter its data, thereby no vibration information is carried out.

So, *how could TagSound enable the tag to backscatter or transmit persistently?* According to the Gen2 protocol, the reader is allowed to request a *retransmission* of the long reply by resending the `ACK` command followed with the last transmission. This mechanism is designed to avoid the data corruption. *Our trick is that we use the retransmission as an approach to acquiring the persistent vibration modulation.* Namely, TagSound always transmits an `ACK` command after each transmission no matter the reply is received correctly or not. Fig. 5 shows a the baseband signals acquired by the sniffer reader, where the reader acquires the long reply from the tag persistently via resending `ACK`. In this way, we can see that the tag almost backscatters in above 70% time. Note that there still exists 30% time for the `ACK`. We will discuss



(a) Spectrum (b) Constellation (c) CFO vs. Channels (d) Constellation

Fig. 6: The CFO of RFID Reader

this issue in Sec. VI.

V. Compensation of Carrier Frequency Offset

In this section, we introduce the CFO and how the CFO affects our vibration sensing. Finally, we introduce the compensation approach.

A. *CFO Background*

CFO is a usual phenomenon in a wireless communication system. It often occurs when the local oscillator signal for down-conversion in the receiver does not synchronize with the carrier signal contained in the received signal. The consequence is that the frequency of the received carrier signals does not match that of the transmitting ones, resulting in an offset in spectrum. CFO is normally caused by two main factors: 1) frequency mismatch due to the difference in clock drift rates between the oscillators of transmitter and receiver; 2) the Doppler effect when transmitter to receiver is moving.

B. *Impact of CFO on TagSound*

In TagSound, a commercial reader is used to generate the CW for interrogating tags, but another sniffer reader is employed to acquire the harmonics. The two readers have their own internal clocks with different drifts. Moreover, the tag is vibrating with high-frequency, which would cause Doppler effect at its backscattered signal. Thus, both factors would produce the CFO in TagSound.

To better understand the impact of CFO, we perform a group of empirical experiments on TagSound. We use the Speedway software provided by ImpinJ Company to configure an ImpinJ R420 to interrogate a static tag at 920.63MHz. Meanwhile, we use USRP N210 to estimate the frequency of the sniffed CW. The received signal is supposed to exactly spike at 920.63MHz, but actually appear at 920.6682MHz, leading to about 38.25kHz offset, as shown in Fig. 6(a).

Does this 38.25*kHz offset mater?* To answer this question, let us translate this number into the Eqn. 2 as follows:

$$\begin{aligned} S(t) =& A(t)\cos(2\pi(f_c+38250)t+2\pi\frac{d_0}{\lambda}) \\ =& A(t)\cos(2\pi f_c t+\underline{2\pi 38250t}+2\pi\frac{d_0}{\lambda}) \end{aligned} \tag{8}$$

where $f_c = 920.63$MHz. Clearly, an additional phase rotation is introduced when we downconvert the carrier using the supposed central frequency at 920.63MHz. Worsely, the rotation grows as time! For example, phase becomes $38250\times 1/2\pi \bmod 2\pi = 4.25$ radian larger than before after one second, even though the distance remains unchanged. Fig. 6(b) shows the distribution of baseband samples in the I-Q constellation. It is supposed to distribute in two clusters corresponding to bit '0' and bit '1'. The fact is the samples are distributed on two circles. This is because the phase caused by the offset linearly increases as time and is wrapped every 2π radians. We also investigate the CFO in other 15 channels that EPC Gen2 specifies. The results are shown in Fig. 6(c). We can see that the CFO varies across channels.

In summary, CFO affirmatively exerts serious influence on TagSound since the vibration information is modulated through the phase. Especially, the CFO would be enlarged by 3× at the third-order harmonics, *e.g.* $38.25\times 3 = 114$kHz.

C. *CFO Estimation*

To compensate the offset after the downconversion, we should estimate the CFO firstly.

Fine-Grained Estimator. Because of the CFO, the backscattering signal is actually transmitting at frequency f_c' but received at f_c where $\Delta f = |f_c - f_c'| > 0$. The backscattering signal is rewritten as follows:

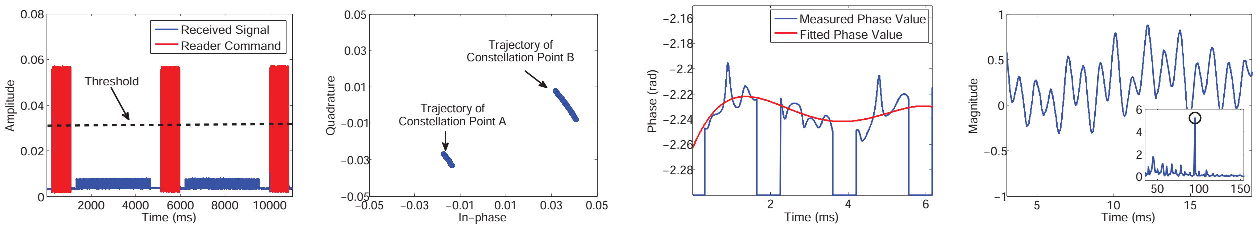
$$S(t) = B(t)e^{\mathbf{J}2\pi f_c' t} \tag{9}$$

where $B(t) = A(t)V(t)$. For clarity, we combine two data together and consider it as a whole baseband signal. The sniffer reader attempts to extract the baseband signal through the downconversion, *i.e.* multiplying carrier signal with a conjugated single-tone at f_c as follows:

$$\begin{aligned} S_{\downarrow}(t) &= S(t)e^{-\mathbf{J}2\pi f_c t} = B(t)e^{\mathbf{J}2\pi f_c' t}e^{-\mathbf{J}2\pi f_c t} \\ &= B(t)e^{\mathbf{J}2\pi(f_c'-f_c)t} = B(t)e^{\mathbf{J}2\pi\Delta f t} \end{aligned} \tag{10}$$

Suppose that we can find out two identical sequences with an interval of Γ that are received in different time. The second sequence is received after the first one with a delay of τ. Namely, $S_{\downarrow}(t) = S_{\downarrow}(t+\tau)$ for $0 \le t < \Gamma$. Then we compute the self-correlation of the downconverted signal with respect to the delay τ as follows:

$$\begin{aligned} Z =& \int_0^{\Gamma} S_{\downarrow}(t)S_{\downarrow}^{*}(t+\tau)dt = \int_0^{\Gamma} B(t)e^{\mathbf{J}2\pi\Delta f t}B(t+\tau)e^{-\mathbf{J}2\pi\Delta f(t+\tau)}dt \\ =& e^{-\mathbf{J}2\pi\Delta f\tau}\int_0^{\Gamma}|B(t)|^2 dt \end{aligned} \tag{11}$$



(a) (b) (c) (d)

Fig. 7: Phase extraction from baseband signals. (a) shows the comparison between the tag backscatter and reader command; (b) shows the samples of baseband signal in the constellation; (c) shows the comparisons between the measured phase and filtered phase; (d) the refined vibration signal.

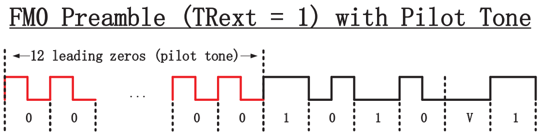

FM0 Preamble (TRext = 1) with Pilot Tone
12 leading zeros (pilot tone)
0
0
...
0
0
1
0
1
0
V
1

Fig. 8: FM0 extended preamble

Interestingly, Z is a complex constant. Then, the frequency offset Δf can be estimated using the following estimator:

$$\Delta f = -\frac{1}{2\pi\tau}\angle Z = -\frac{f_s}{2\pi K}\angle Z \qquad (12)$$

where $K = \tau f_s$ and f_s is the sampling frequency.

In RFID system, the long reply of a tag contains four fields: 6-bit preamble, 2-bit PC, 96-bit EPC and 16-bit CRC. Clearly, the preamble is an ideal field for us to estimate the CFO. Fig. 8 shows the FM0 extended preamble, which contains 12 consecutive sequences of data “0”, each of which has a duration of $25\mu s$. Removing the two sequences at the beginning and the end, we use the middle 10 sequences for the self-correlation. These 10 sequences are divided into two overlapping segments, each with 9 sequences, *i.e.* the samples number $\Gamma = 50 \times 9 = 450$.

Coarse-Grained Estimator. It is worth noting that the angle of $\angle Z$ follows within $[-\pi, \pi]$ and thereby it might be unambiguous due to the periodicity. To uniquely determine the angle, the maximum allowable frequency offset is given by:

$$\Delta f_{\max} = \frac{f_s}{2K} \qquad (13)$$

In TagSound, our sampling frequency $f_s = 2$MHz and each bit has an interval of $24\mu s$ (or 50 samples). Translating these two numbers into the above equation, we have $\Delta f_{\max} = 20$kHz. Cleary, it fails to estimate the 38.25kHz offset that we have measured in the preview section. Instead, the offset through this estimate approach equals $38.25 \bmod 20 = 18.25$kHz, which is wrapped by a period.

To eliminate the ambiguity due to the periodicity, we firstly identify the offset in a coarse range with the FFT. EPC Gen2 specifies that $125 \sim 150\mu$ time window (called T4) is reserved after the query command and before the tag's reply (see Fig. 5), during which only CW is present. With respect to 2MHz sampling, there are $200 \sim 250$ samples during T4. We use these samples to perform the FFT. It is known that the resolution of FFT is given by the division of the f_s to the number of samples. In our case, the resolution is $2\text{MHz}/200 = 10$kHz. Therefore, by taking FFT on the samples during T4, we can quickly identify the central frequency into a 10kHz-wide bin.

A natural question is that why not to perform FFT on a longer CW to improve frequency resolution? If we want to estimate the offset at Hz-level, then the FFT requires at least one second or 2M samples. Unfortunately, the vibration may induce varying Doppler frequency offset. One second time is too long to keep track of the real-time vibration.

Put It Together. Now, let us put all pieces together to sketch the algorithm. Firstly, we use the samples acquired during the T4 window to perform the FFT. Consequently, the offset should be located in a 10kHz wide frequency bin by comparing the supposed carrier frequency. Secondly, we use the preamble of the tag's long reply to continue to estimate the offset at Hz level within 20kHz. Third, the final Hz-level offset is obtained by combining the two estimation results.

D. CFO Compensation

Suppose the estimated CFO is $\Delta\widehat{f}$, we can compensate the offset by multiplying the baseband signals, downconverted by the sniffer reader already, with a single tone at $\Delta\widehat{f}$ as follows:

$$\begin{aligned}\widehat{B}(t) &= B(t)e^{-\mathbf{J}2\pi\Delta\widehat{f}t} = B(t)e^{\mathbf{J}2\pi\Delta f t}e^{-\mathbf{J}2\pi\Delta\widehat{f}t} \\ &= B(t)e^{\mathbf{J}2\pi(\Delta f-\Delta\widehat{f})t} \approx B(t)\end{aligned} \qquad (14)$$

It can bee seen that the signals caused by the CFO can be exactly removed from the above equation. The remained baseband signal is the multiplication of the tags' reply and our vibration signal. Fig. 6(d) shows the phase in the constellation after the compensation. The two circles shown in Fig. 6(b) collapse into two clusters, which correspond to bit ‘0’ and bit ‘1’ respectively.

VI. Phase Extraction

So far, we have successfully separated the persistent baseband signal (*i.e.* $A(t)V(t)$) from the backscattering signal, by removing the errors caused by the CFO. Our final step is to extract the vibration signals from the baseband. Fortunately, although the baseband signal contains two types of data, they have two independent modulation mechanisms, *i.e.* the OOK does not affect the PSK. The procedure contains three steps as follows:

Step 1: Preprocessing. Since the tag remains silent when the reader broadcasts commands, the sniffer reader should never receive the harmonics of reader commands. Unfortunately, as we pointed, the commercial reader does not suppress the harmonics well. It is possible that the final baseband signal still contains the harmonics of commands. Thus, our first step is to remove these noise. Reader's signals are normally much stronger than the backscattering signal, we remove these noise by a threshold. An example is shown in Fig. 7(a).

Step 2: Interpolation. Although OOK and PSK are independent, OOK would produce two clusters in the constellation, obscuring the phase values. Fig. 7(b) shows an example of baseband signal in the constellation. We can see two clusters, which have similar angles (*i.e.* phase) but different amplitudes. Thus, in the second step, we select one clusters to measure their phase change, and estimate the phase value of the other clusters by linear interpolation. The same method also used to estimate the phase value of the interval between two tag signals, as shown in Fig. 7(c).

Step 3: Refining. Due to the huge gap between the sampling frequency and vibration frequency, , TagSound uses an sliding average filter to cancel the ambient noise, and output the vibration signal. In order to avoid the sudden interference from the surrounding, another Kalman filter is used to smooth the final vibartion signal. Fig. 7(d) shows the final output vibration signal in both time and frequency domain.

VII. Implementation and Evaluation

In this section, we introduce the implementation and evaluate our prototype from various perspectives.

A. Implementation

We implement a prototype of TagSound using an ImpinJ COTS UHF reader R420 [14] for the interrogation operation and a USRP N210 for the monitoring. The USRP is connected to the a ThinkPad T450P with $2.2Ghz$ CPU. We utilize a $4dBi$ RFID antenna on $920MHz$ for COTS reader and a $3dBi$ directional antenna on $2.4GHz$ for USRP. Fig. 9 illustrates the detailed set up. The ImpinJ reader is set to interrogation at the carrier frequency of $920.63MHz$, while the USRP monitors the harmonic backscatter signal at the frequency of $2761.89MHz$ with the sampling frequency of $2MHz$. The signal will received by running a GNU Radio monitor program and than send to Matlab for offline processing. The transmitting power of RFID reader is $31dBm$.

B. The effect of CFO estimation

Since the performance of CFO estimation and compensation will directly determine whether TagSound works properly, we first test the performance of our proposed CFO estimator. As illustrated in Figure 6(c), the frequency offset on the commercial RFID reader is not stable, thus it is difficult to provide the ground truth for our test. Therefore, we use the USRP device to perform the test of our CFO estimator. We implemented the RFID reader on one USRP device[**?**] and allowed it to communicate with the tag, while another USRP is used to listen the entire RFID communication in the harmonic band. In order to ensure that the test is not interfered by the CFO between USRPs, we use an external clock source to synchronize both USRPs. We introduce a known CFO by adjusting the carrier frequency of the USRP reader, and the CFO ranges from $30kHz$ to $40kHz$ with a step of $100Hz$. The CFO estimation error is defined as $|CFO_{est} - CFO_{real}|$. As shown in Fig. 10, about 90% of the CFO estimation error is less than $5.5Hz$, and 50% of the CFO estimation error is below $2Hz$. This means that our proposed CFO estimator is sufficient to guarantee TagSound works properly.

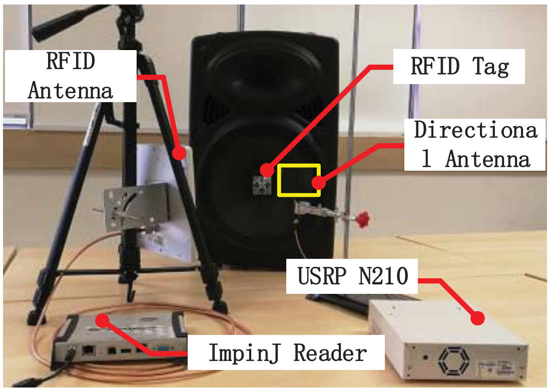


Fig. 9: The Experiment Scenario

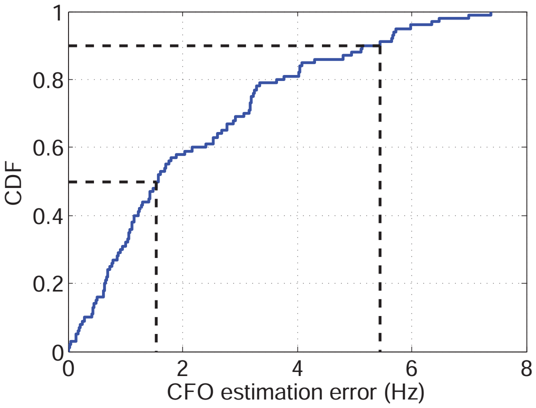
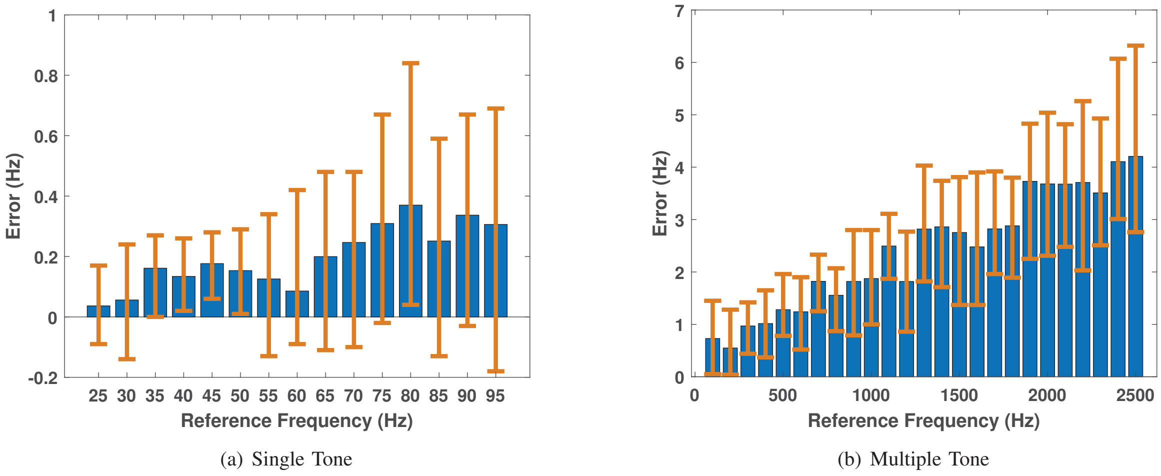


Fig. 10: The performance of CFO estimation algorithm

C. The impact of frequency

In this section, we will test the accuracy of frequency recognition of TagSound. We put the RFID tag on the surface of a loudspeaker, and use an USRP N210 to sniffer the harmonic copy of the communication between RFID reader and tag nearby. (See Figure 9) In order to obtain vibrations with various frequencies, we generate the sound files with different frequencies and drive the loudspeaker to play these acoustic files. In this experiment, the sound frequency is set from $25Hz$ to $2500Hz$. and divided into two sections of low frequency and high frequency. In the low frequency section,



(a) Single Tone

(b) Multiple Tone

Fig. 11: System Performance Evaluation. Both figures are the result of frequency recognition.(a) The performance of low frequency tone recognition (b) The performance of high frequency tone recognition

we play single-tone signals from $25Hz$ to $95Hz$, with a step of $5Hz$, while in the high frequency section, we play multi-tone signal from $100Hz$ to $2500Hz$, with a step of $100Hz$. For each frequency setting, we repeat the experiment for 25 times. Fig. 11(a) shows the results of the experiment. It can be seen that, TagSound can achieve high accuracy on vibration frequency identification in both low frequency and high frequency, with the average error below $0.37Hz$ and $4.2Hz$ respectively.

D. The impact of sampling frequency

The sampling frequency of TagSound is much larger than the frequency of the vibration signal. The higher sampling frequency can better eliminate the interference caused by environmental noise and improve the accuracy of vibration frequency estimation. Fig. 12 shows the error of the vibration frequency estimation of TagSound at different sampling frequencies. With the increases of sampling frequency, the estimation error does not decrease significantly. This is because even the sampling frequency is $2MHz$, it is much higher than the frequency of vibration. With the help of averaging filter, TagSound can eliminate most of the noise interference. Therefore, applying a higher sampling frequency does not improve TagSound's performance. By considering both estimate accuracy and time cost, we choose $2MHz$ as our default sampling frequency.

E. The impact of tag diversity

Finally, we test the performance of TagSound among different types of tags. We first measure the freuqency estimation performance of TagSound with different tags. Four types of tags are used, which are ImpinJ QT4 with the size of $4.8cm \times 4.8cm$, ImpinJ BT45 with the size of $2cm \times 2cm$, Alien 9629 with the size of $2.55cm \times 2.55cm$ and Alien 9640 with the size of $15.9cm \times 1.5cm$. We utilize the TagSound to estimate the vibration frequency for 50 times and calculate the estimation error. As shown in Fig. 13, the tag ImpinJ QT4 achieves the best performance with the mean error of $0.37Hz$ and the performance of tag Alien 9640 and Alien 9629 are closely followed. The tag ImpinJ BT45 does not work in this

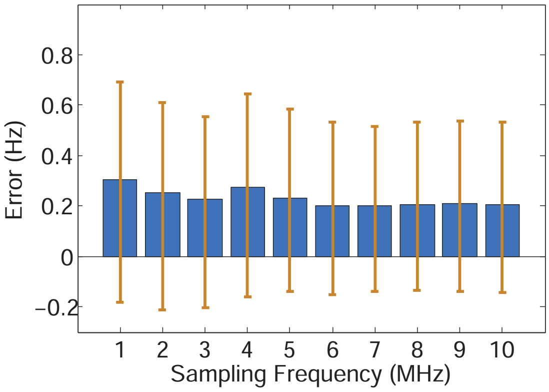


Fig. 12: Error *vs.* sampling frequency

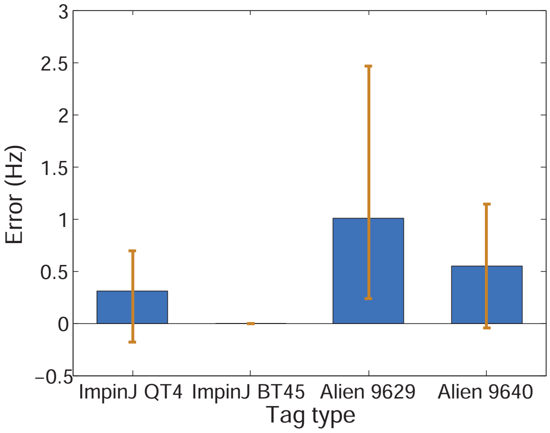


Fig. 13: Error*vs.* types of tags

experiment since the power of its harmonic signal is too weak to be detected by our USRP N210 platform. This experiment confirms that the tag with a larger antenna can absorb more energy from the reader. Bigger antenna makes tag's harmonic backscattered signal stronger and thereby results more precise.

VIII. Related Work

In this section, we review literatures related to our work.

RFID-based vibration methods: Although there are lots of works on RFID-based positioning algorithms[15], there are few results in applying RFID to vibration detection. TagBeat[1] makes the first attempt to measures vibration via RFID technology, with the advantage of being low-cost and device identification capability. TagTwins[8] has been improved on the basis of TagBeat to make it immune to the interference of ambient noise. However, these methods are subject to the reading rate of RFID tag and can't detect high frequency vibrations. Even with the help of the compressed sensing technology, they can not measure the vibration frequency above $100Hz$. TagSound, on the other hand, breaks through the limitation of reading rate of RFID tag by mining the characteristics of harmonic backscatter, and can accurately detect high frequency vibration signals. Moreover, [16] can detect and track tag movements to infer customer's behavior, but it failed to slight tag movement. TagSound takes the advantage of the short wavelength of harmonic signal, can easily capture the vibration with small magnitude.

Radar-based method: Using the principle of radar to replace the special sensor to achieve remote vibration detection is a trend in the academic and industrial world. LADAR[17] investigates the development of such methods. Inspired by LADAR, Teng Wei *et al.* studied the relationship between mechanical vibration and the wireless signal in both RSS and phase, and proposed ART[6] to recover the loudspeaker sound by using a co-located WiFi transmitter. [18] leverages FMCW technique to let the drone track user's movement accurately. In contrast, TagSound monitors the phase fluctuation of the signal caused by mechanical vibration in the harmonic frequency band, and skillfully avoids the interference caused by the leakage of the transmitted signal, which exists in algorithms such as ART. Moreover, with the help of RFID tag, TagSound can directly identify the vibration source easily.

IX. Conclusion

This work presents an RFID-based vibration sensing system which can break the limitation of reading rate of RFID tags. Our key innovations lie in leveraging the harmonic backscatter signals caused by the non-linearly of RFID tag, and a new form of phase extraction algorithm. Experimental result demonstrate that even with high frequency, TagSound still can sense the vibration frequency to an accuracy of sub-hertz. We believe our system will promote more possibilities of RFID-based sensing solution in practical deployments.

Acknowledgment

The corresponding author of this paper is Panlong Yang. The research is supported by NSFC General Program (No. 61572282, No. 61625205, No. 61632010, No. 61772546, No. 61751211, No. 61772488, 61520106007), UGC/ECS (No. 25222917), Shenzhen Basic Research Schema (No. JCYJ20170818104855702), National key reseach and development plan (No. 2017YFB0801702), Key Research Program of Frontier Sciences, CAS (No. QYZDY-SSW-JSC002), NSFC with No. NSF ECCS-1247944, NSF CNS 1526638, NSF of Jiangsu For Distinguished Young Scientist (No. BK20150030) and Alibaba Innovative Research Program. We thank all the anonymous reviewers and shepherd for their valuable comments and helpful suggestions.

References

[1] "Tagbeat," https://github.com/tagsys/tagbeat.

[2] "Laser microphone," https://en.wikipedia.org/wiki/Laser_microphone.

[3] A. Davis, M. Rubinstein, N. Wadhwa, G. J. Mysore, F. Durand, and W. T. Freeman, "The visual microphone: passive recovery of sound from video," 2014.

[4] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng, "Tracking vital signs during sleep leveraging off-the-shelf wifi," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. ACM, 2015, pp. 267–276.

[5] C. Wang, L. Xie, W. Wang, Y. Chen, Y. Bu, and S. Lu, "Rf-ecg: Heart rate variability assessment based on cots rfid tag array," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 2, p. 85, 2018.

[6] T. Wei, S. Wang, A. Zhou, and X. Zhang, "Acoustic eavesdropping through wireless vibrometry," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 2015, pp. 130–141.

[7] L. Yang, Q. Lin, and Y. Li, "Making sense of mechanical vibration with cots rfid systems," in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*. ACM, 2016, pp. 487–488.

[8] C. Duan, L. Yang, H. Jia, Q. Lin, Y. Liu, and L. Xie, "Robust spinning sensing with dual-rfid-tags in noisy settings," *Energy*, vol. 1, pp. 1–1.

[9] Y. Lei, Z. He, and Y. Zi, "Application of an intelligent classification method to mechanical fault diagnosis," *Expert Systems with Applications*, vol. 36, no. 6, pp. 9941–9948, 2009.

[10] P. V. Nikitin and K. Rao, "Harmonic scattering from passive uhf rfid tags," in *Antennas and Propagation Society International Symposium, 2009. APSURSI'09. IEEE*. IEEE, 2009, pp. 1–4.

[11] D. Allane, G. A. Vera, Y. Duroc, R. Touhami, and S. Tedjini, "Harmonic power harvesting system for passive rfid sensor tags," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 7, pp. 2347–2356, 2016.

[12] "KEYSIGHT Technologies," https://www.keysight.com/us/en/home.html.

[13] G. A. Vera, Y. Duroc, and S. Tedjini, "Analysis and exploitation of harmonics in wireless power transfer (h-wpt): passive uhf rfid case," *Wireless Power Transfer*, vol. 1, no. 2, pp. 65–74, 2014.

[14] "Impinj, Inc," http://www.impinj.com/.

[15] L. Yang, Y. Chen, X.-Y. Li, C. Xiao, M. Li, and Y. Liu, "Tagoram: Real-time tracking of mobile rfid tags to high precision using cots devices," in *Proc. of ACM MobiCom*, 2014.

[16] J. Han, H. Ding, C. Qian, W. Xi, Z. Wang, Z. Jiang, L. Shangguan, and J. Zhao, "Cbid: A customer behavior identification system using passive tags," *IEEE/ACM Transactions on Networking*, vol. 24, no. 5, pp. 2885–2898, 2016.

[17] P. Castellini, M. Martarelli, and E. P. Tomasini, "Laser doppler vibrometry: Development of advanced solutions answering to technology's needs," *Mechanical Systems and Signal Processing*, vol. 20, no. 6, pp. 1265–1285, 2006.

[18] W. Mao, Z. Zhang, L. Qiu, J. He, Y. Cui, and S. Yun, "Indoor follow me drone," in *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 2017, pp. 345–358.