IET Wireless Sensor Systems

Special issue Call for Papers

Be Seen. Be Cited.
Submit your work to a new IET special issue

Connect with researchers and experts in your field and share knowledge.

Be part of the latest research trends, faster.

Read more

The Institution of Engineering and Technology

IET Wireless Sensor Systems

Research Article

IET Journals
The Institution of Engineering and Technology

Hybrid indoor location positioning system

ISSN 2043-6386
Received on 6th January 2019
Revised 11th March 2019
Accepted on 15th April 2019
E-First on 22nd May 2019
doi: 10.1049/iet-wss.2018.5237
www.ietdl.org

Shuo Li[1] ✉, *Rashid Rashidzadeh*[1]

[1]*Department of Electrical and Computer Engineering, University of Windsor, 401 Sunset Ave, Windsor, ON N9B 3P4, Windsor, Canada*
✉ *E-mail: li1c8@uwindsor.ca*

Abstract: Indoor location positioning techniques have experienced a significant growth in recent years. This work presents a hybrid indoor positioning system with fine and coarse modes. It utilises acoustic signals for fine positioning and received signal strength (RSS) for coarse location estimation. Acoustic positioning systems require a line-of-sight connection for accurate positioning which may not be available due to obstacles in indoor environments. A new solution is presented to overcome this problem using RSS as a reference to validate the line-of-sight connection. Moreover, a new digital signal processing algorithm using a matched filter is presented to enhance the system's robustness in indoor environments with low a signal-to-noise ratio. Experimental measurement results in an indoor environment show that the proposed solution can accurately determine indoor locations with <6 cm positioning error on average.

1 Introduction

Location-aware applications and services are becoming more attractive than ever particularly in indoor environments. GPS is widely used for location positioning however GPS requires a line-of-sight connection between satellite transmitters and a receiver to ensure positioning accuracy. Implementation of a GPS positioning system for indoor areas is costly and unreliable [1]. Many cost-effective solutions have been developed for indoor location estimation in recent years. Indoor positioning systems using received signal strength emerged as low-cost systems achieving acceptable accuracy for many applications [2, 3]. Ultra-wideband (UWB) technology [4] has also appeared for accurate and precision indoor location estimation. However, UWB-based systems are relatively costly. Location estimation systems using audio signals support a high-accuracy positioning at low cost. Also, due to the widespread use of the cell phones, it is highly desired to design an indoor positioning system that only requires smartphone at the user end.

The proposed system is developed to support off-line cell-phones for location positioning without the need to Internet connection. Such a positioning system preserves user privacy as the location information is not shared online. In the proposed solution, bluetooth low energy (BLE) tags with sub-1 GHz RF communication are used where the BLE nodes serve as audio receivers. Their power consumption is lower compared to systems using such nodes as audio transmitters. The BLE nodes RF circuit operate in the sleep mode >50% of the time during the operation cycle, they are only activated in certain time points. The chirp spread spectrum [5] technique is utilised in the audio mode to increase the rage by reducing the effective noise level. Frequency modulated audio waves together with a matched filter are used to support indoor location estimation over a wide range. Moreover, an audio data clustering technique is used to enhance the system's robustness in dense multipath environments. A non-line-of-sight (NLoS) identification method is developed to discard the outliers. Moreover, the proposed system supports a coarse mode and a fine mode for location positioning. In the coarse mode, the approximate location with meter-level accuracy is determined based on RSS from BLE nodes while in the fine mode audio signals are utilised to locate a position with centimetre-level accuracy. As a result, the proposed system is useful in indoor scenarios where both meter and centimetre level accuracies are required. The accuracy levels for positioning systems depend on the application scenarios. Take the shopping mall as an example, to find a store a meter-level positioning accuracy is required, but for locating an item in the store centimetre-level positioning is necessary. The proposed system can switch between the two accuracy levels depending on the user's needs.

The rest of the paper is organised as follows: Section 2 covers the available solutions in the literature, Section 3 presents an overview of the proposed solution. Section 4 explains the details of the implemented DSP algorithms in the acoustic mode. Section 5 presents the experimental measurement results and section 6 summarises the results and presents the conclusions.

2 Related work

There are several reported indoor positioning systems in the literature using Bluetooth RSS technology. A commercial system is presented in [6], which is able to determine a user's proximity to a beacon. Another system is presented in [7], which utilises the fingerprinting method and supports indoor location positioning with <5 m error in 95% of the case. Although the fingerprinting algorithm can improve the positioning accuracy, it requires intensive work before the positioning phase. There are also light weight systems for indoor positioning. Zaki and Rashidzadeh proposed a BLE-RSS-based indoor positioning system in [8], which utilises the weighted average of the received signal strengths which can achieve an average accuracy of <2.1 m. Acoustic indoor location positioning systems can be classified into two categories of (a) systems using ultrasound and (b) systems using near-ultrasound. Cricket [9] and active bat [10] are the two pioneering solutions using ultrasound to localise indoor positions supporting centimetre level accuracy. However, they require custom-designed devices at the user side and they do not support smartphone platforms. Beep [11] uses audible signals with off-the-shelf devices, but its positioning accuracy is low due to the latency at the sound card. Assist [12] utilises near-ultrasound pulses generated by a smartphone speaker to achieve centimetre level positioning accuracy, however, the use of Wi-Fi NTP protocol for time synchronisation increases power-consumption considerably [13]. Akkurate [14] uses smartphones as audio receivers and supports a high positioning accuracy in 2D environments, but the accuracy drops dramatically for a real-world 3D location positioning. The BLE and ultrasound technologies have been used in ALPS [15] for localisation in which BLE nodes transmitter acoustic signals. An algorithm is also presented using machine learning techniques to

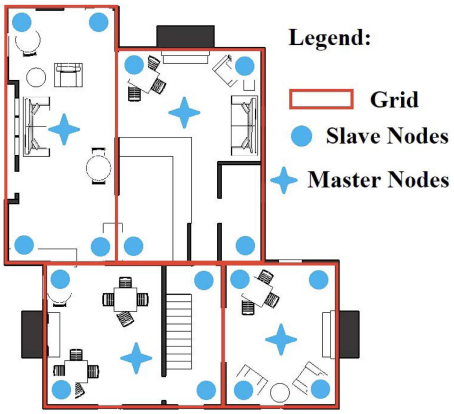

Fig. 1 *Indoor environment covered with the system*

enhance the system's robustness. Murata *et al.* proposed a high accuracy near-ultrasound positioning system compatible with smartphones, however the positioning accuracy drops considerably in noisy environments [16]. Pérez *et al.* proposed an android application called LOCATE-US. The system implements code-division multiple access technique to overcome the problems at the acoustic frequency range between 20–22 kHz [17]. Zhang *et al.* proposed MAIDLOC and RA2LOC systems along with DSP algorithms to enhance the system's performance in noisy and dense multipath locations [18]. They also presented an NLoS identification algorithm using machine learning to discard the outliers due to NLoS propagations [19].

3 System overview

3.1 Dual mode positioning

The proposed system combines the RSSI-based positing algorithm developed by the research team [8] for coarse positioning, and the acoustic positioning method for fine positioning. The RSSI-based positioning method can achieve a meter-level accuracy while the acoustic mode using the time difference of arrival (TDoA) algorithm can achieve centimeter level accuracy. The algorithms are implemented on BLE nodes and supported by a smartphone app. The indoor area, as shown in Fig. 1, is divided into grids. Each grid is covered with the BLE nodes consisting master and slave nodes which can receive audio pulses and broadcast BLE packets. Mater nodes directly communicate with the smartphone via BLE. Slave nodes receive inter-node commands from master nodes through sub-1 GHz RF, this architecture consumes less power compared to the method in which each node receives commands directly from smartphone via BLE. In the coarse positioning mode, BLE packets are transmitted by all BLE nodes and received by smartphones. The user location in this mode is determined using the RSS-based algorithm in [8]. In the acoustic mode, the smartphone transmits short audio pulses, the BLE nodes receive the pluses and send the results back to the smartphone after applying a DSP algorithm. Acoustic mode positioning is commonly affected by (a) indoor environment audio interferences and (b) lack of line-of-sight propagation. The RSSI-based method is utilised to filter out outliers from acoustic positioning mode as explained in Section 4.5.

3.2 Multi-lateration positioning

Smartphone operating systems have audio-latency problems which leads to inaccurate time-of-flight (ToF) estimation between audio receivers and smartphones. As a result range-based positioning such as trilateration cannot be implemented.

A range-free positioning algorithm called multi-lateration is utilised to overcome the audio latency problem. The difference between the time-of-arrivals (ToAs) at the audio receivers are subtracted to get corresponding TDoAs, as shown in Fig. 2. The implemented BLE nodes support sub-1 GHz RF signals used to synchronisation all audio receiver nodes [20]. In a room with audio

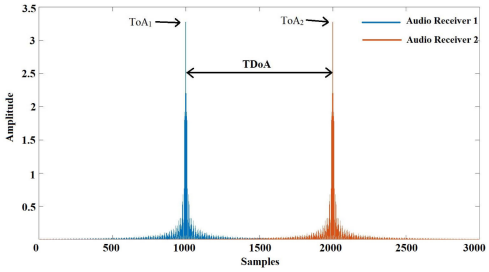

Fig. 2 *Time-of-arrivals (ToAs) for two nodes and the corresponding TDoA*

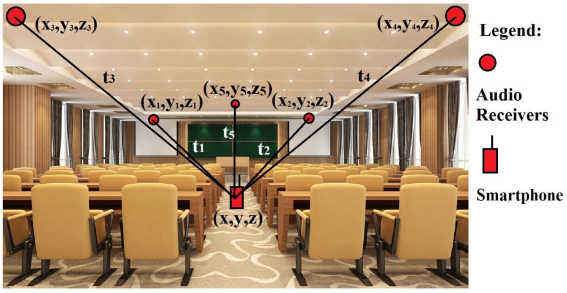

Fig. 3 *Smartphone and audio receivers*

receivers and a smartphone shown in Fig. 3, the user coordinates, x, y and z can determined from

$$\sqrt{(x-x_2)^2+(y-y_2)^2+(z-z_2)^2}-\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2}=v(t_2-t_1) \quad (1)$$

$$\sqrt{(x-x_3)^2+(y-y_3)^2+(z-z_3)^2}-\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2}=v(t_3-t_1) \quad (2)$$

$$\sqrt{(x-x_4)^2+(y-y_4)^2+(z-z_4)^2}-\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2}=v(t_4-t_1) \quad (3)$$

$$\sqrt{(x-x_5)^2+(y-y_5)^2+(z-z_5)^2}-\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2}=v(t_5-t_1) \quad (4)$$

where t_1–t_5 indicate the time-of-flight to different audio receivers plus the audio latency time. To cancel out the effect of audio latency, the time difference of arrival is determined. Each ToA is the summation of the time-of-flight and the audio latency time. By pairwise subtracting the ToAs, the effect of unknown latency time is nullified. If we assume the first audio receiver's is positioned at the origin, we can use the linear least-square algorithm to achieve an optimal solution [21]. We have the following matrices:

$$\boldsymbol{A}=\begin{bmatrix} d_{21} & x_2 & y_2 & z_2 \\ d_{31} & x_3 & y_3 & z_3 \\ d_{41} & x_4 & y_4 & z_4 \\ d_{51} & x_5 & y_5 & z_5 \end{bmatrix} \quad \boldsymbol{b}=\frac{1}{2}\begin{bmatrix} b_2^2-d_{21}^2 \\ b_3^2-d_{31}^2 \\ b_4^2-d_{41}^2 \\ b_5^2-d_{51}^2 \end{bmatrix}$$
$$\boldsymbol{y}=\begin{bmatrix} \boldsymbol{X} \\ x \\ y \\ z \end{bmatrix} \quad \boldsymbol{X}=\begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (5)$$

where $d_{ij}=v(t_i-t_j)$, $b_i^2=x_i^2+y_i^2+z_i^2$ and $\boldsymbol{X}$ is the Euclidean vector norm [21]. We then use the linear least squares to achieve the optimal estimation of smartphone coordinate as follows:

$$\tilde{\boldsymbol{y}}=\left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{b} \quad (6)$$

The above algorithm can provide accurate positioning results. There are also non-linear least square algorithms such as Gauss–Newton, Steepest Descent and Levenberg–Marquardt proposed to further improve the positioning accuracy [22]. These non-linear algorithms first expand (1)–(4) at the point $\boldsymbol{X}$, approximate these equations with the first-order Taylor series, then they use $\boldsymbol{X}$ as the initial guess to implement iterative process to get the final estimated positions. However, these algorithms require the initial point $\boldsymbol{X}$ to be close to the real position, otherwise the iterative process may fail to converge [23]. Using the above-mentioned linear least-square method the estimated z coordinates are often far away from their real values. Though the estimation error in z coordinate will not corrupt the horizontal positioning performance, the above-mentioned algorithms may fail to work. Moreover, an iterative process on a smartphone app may take long time to converge, so instead of the non-linear least-square methods we only utilise linear least square method with an optimal node placement, which will be discussed later. For accurate positioning, the accuracy of ToA estimations is of great importance since TDoAs are generated by pair-wise subtracting of ToAs. Moreover, an accurate inter-node time synchronisation is vital to eliminate the added latency by the nodes.

4 ToA estimation

4.1 Linear frequency modulated chirp & matched filtering

Linear frequency modulated (LFM) chirp pulse is a sinusoidal wave pulse whose frequency, f, increases or decreases linearly with time, $f = f_0 + kt$. The signal in time domain can be represented by

$$x(t) = \cos\left(2\pi f_0 t + k\pi t^2\right) \tag{7}$$

where f_0 is the initial frequency and k is the frequency changing rate, in our system the signal's length is 150 ms, the initial frequency f_0 is 15 kHz and the frequency changing rate k is 20 kHz/s, the transmitted signal is an up-chirp with the bandwidth of 3 kHz.

We used a matched filter [24] to determine the chirp's ToA. Ideally, an audio signal, $x(t)$, received as

$$y(t) = \begin{cases} x(t - T_{\text{ToA}}) & T_{\text{ToA}} \le t \le T' \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where T' is the chirp's ending time. We cross-correlated the received signal with the reference signal, this is essentially an auto-correlation. The cross-correlation function can be written as

$$c(\tau) = \int_0^T y(t) \cdot \text{Ref}(t - \tau)\text{d}t \quad \tau \ge 0 \tag{9}$$

where $y(t)$ is the received signal, $\text{Ref}(t)$ is the reference chirp signal and T is the chirp signal's length. The cross-correlation result can be written as [25]

$$c(t) = \sqrt{BT} \cdot \sin c[\pi B(t - T_{\text{ToA}})]\cos[2\pi f_0(t - T_{\text{ToA}})] \tag{10}$$

where B is the chirp signal's bandwidth, $c(t)$ reaches its peak at $t = T_{\text{ToA}}$ which is the arrival time of the transmitted audio signal.

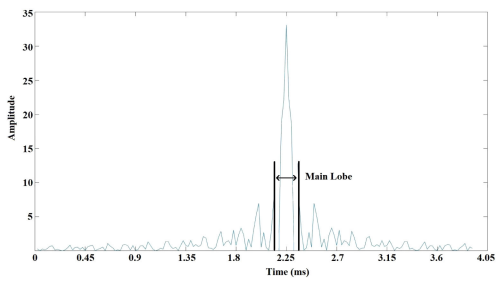

Fig. 4 *Cross-correlation function main lobe*

The energy of the whole received chirp pulse is compressed into the main lobe of $c(t)$ after the cross-correlation, as shown in Fig. 4. Compared to the original sinc function, its amplitude is increased by the ratio of $\sqrt{BT}$, in our system this ratio is ~21. Sound energy attenuates dramatically as the transmission range increases, however with chirp signal and a matched filter, cross-correlation peak corresponding to the ToA point can still be detected even after a propagation range of 20 m. Moreover, a chirp signal has a good noise-resistance performance and Doppler effect tolerance. As a result, the combination of a chirp signal and a matched filter is well suited for ToA estimation of acoustic signals.

In the proposed solution, the algorithm to determine the ToAs runs on the BLE nodes. We compute the cross-correlation in the frequency domain instead of the time domain to reduce the power consumption. The cross-correlation function can be written as

$$c(t) = \text{IRFFT}(Y \cdot \text{REF}^*) \tag{11}$$

where IRFFT is the inverse real Fourier transform. Y is the real Fourier transform of the received signal $y(t)$, and REF is the Fourier transform of the reference chirp signal $\text{Ref}(t)$, REF^* is the conjugate of REF. It takes ~20 ms to calculate the cross-correlation for 16,384 sample points in the frequency domain, which supports real-time location estimation for many applications.

4.2 Indoor acoustic environment modelling

In an ideal case, the received audio signal can be represented by

$$y(t) = \sum_{i=1}^{N} a_i x(t - \tau_i) + n(t) \tag{12}$$

where N is the number of propagation paths, a_i indicates different paths' fading coefficients and τ_i represents different paths' time delay and $n(t)$ is white noise. The signal from the direct propagation path takes the shortest transmission time. The received signal from all paths can be considered as the superposition of replicas of $x(t)$ with different time delays and energy attenuations. The direct path signal has the strongest energy contribution compared to signals from other paths creating the highest correlation peak. There is no correlation between the forwarded chirp signal and white noise. Thus the highest correlation peak is not masked by the audio noise in the environment as long as the noise is not too strong compared to the signal.

In a real-world scenario, the received signals are not the exact replicas of the reference chirp signal $x(t)$. The performance parameters of the propagation channels may vary significantly. An acoustic channel's frequency response is ideally considered flat for all in-band frequencies and represented by (13)

$$H(\omega) = \begin{cases} 1 & \omega_1 \le \omega \le \omega_2 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

In a real case, the in-band frequency response is not flat and the received chirp signal may get distorted as indicated in Fig. 5. Since the chirp signal's frequency changes linearly with time, the corresponding acoustic channel's frequency response is uneven for all in-band frequencies. The acoustic channel's is commonly a

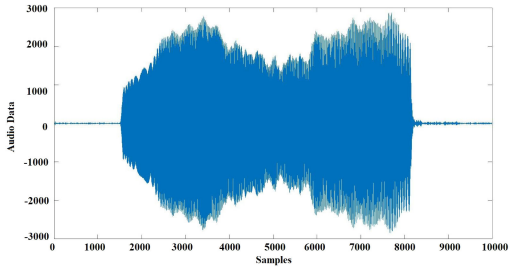

Fig. 5 *Received chirp signal transferred through a propagation channel with non-flat frequency response*

smooth curve in the frequency domain. Thus, it can be expressed by a polynomial using Taylor series. As a result, for every location in a room, no matter how different the acoustic channels are, the frequency responses can be always written as

$$H(\omega) = \begin{cases} a_n\omega^n + a_{n-1}\omega^{n-1} + \cdots + a_0 & \omega_1 \le \omega \le \omega_2 \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

As mentioned above, the cross-correlation algorithm can be performed in the frequency domain. Therefore, for an arbitrary acoustic channel, the cross-correlation function of (11) can be modified as

$$c(t) = \text{IRFFT}(H \cdot Y \cdot \text{REF}^*) \quad (15)$$

It can be expanded as

$$\begin{aligned} c(t) = {} & \text{IRFFT}(a_n\omega^n) * \text{IRFFT}(Y \cdot \text{REF}^*) \\ & + \text{IRFFT}(a_{n-1}\omega^{n-1}) * \text{IRFFT}(Y \cdot \text{REF}^*) \\ & + \cdots + a_0 \cdot \text{IRFFT}(Y \cdot \text{REF}^*) \end{aligned} \quad (16)$$

where IRFFT is the inverse real Fourier transform. The inverse Fourier transform of ω^n is the n_{th} derivative of the Dirac delta function which will not generate an output after convolution with the received signal and the reference signal. Therefore, the non-ideal frequency response of the propagation channel will not affect the final cross-correlation output.

Due to multipath effects, in some indoor environments, as shown in Fig. 6, the direct path signal's correlation peak is not the highest. Moreover, in some cases the line-of-sight (LoS) channel cannot be guaranteed. However, the signal from multi-paths takes longer propagation time compared to the direct path. As a result, we can conclude that in an indoor environment, as long as (a) there exists a direct path between the smartphone speaker and the audio receiver and (b) the SNR is above a certain level, no matter how severe the multi-path effects are, the first detectable correlation peak location is always the ToA of the transmitted audio signal. If the LoS condition cannot be guaranteed and the whole received signal only consists of NLoS path signals, the first correlation peak may not represent the ToA. The NLoS case is discussed in the following section.

4.3 Noise filtering

Environmental noise is an important factor in an acoustic channel. There are many high frequency signals generated by different activities such as door slamming, high-heeled shoes hitting the ground and tableware collisions. The chirp signal in our system sweeps from 15 to18 kHz. A window-based band-pass filter with the passband from 15 to 18 kHz is implemented to filter out the out-of-band noise. However, there are still in-band noises added to the received signal. Using a matched-filtering approach, the received signal generates outputs only when it matches the reference signal. Pearson correlation coefficient [26] can be used to measure the similarity between the received and the reference signals. It is defined as

$$\rho(A, B) = \frac{\text{cov}(A, B)}{\sigma_A \sigma_B} \quad (17)$$

where A and B are the two signals, $\text{cov}(A, B)$ is the covariance of A and B, σ_A is the standard deviation of A and σ_B is the standard deviation of B.

We sampled noise in different indoor environments and computed their Pearson correlation coefficients with a reference signal to estimate the possibilities that they can alter the matched filter's output. A correlation coefficient between -0.1 and $+0.1$ indicates a negligible correlation [27]. The tests results in Table 1 indicate that in 99.9% cases the correlation coefficients are between -0.1 and $+0.1$. As a result, it was concluded that the environmental audio noise is unlikely to correlate with a reference chirp signal to alter the matched filter's output.

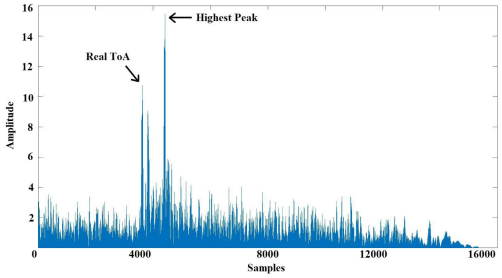

Fig. 6 *ToA peak and highest peak*

Table 1 Noise correlation coefficients tests

Scenarios	No or very weak correlation, %
shopping mall	99.97
restaurant	99.89
university hallway	99.96
indoor parking lot	99.99
noisy office	99.85

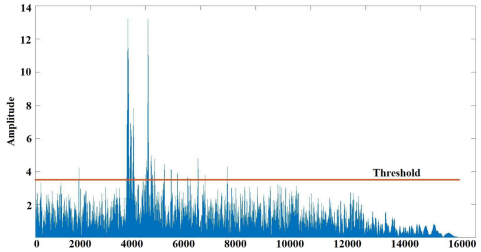

Fig. 7 *Calculated threshold for the propagation channel to filter out the noise*

To further filter out the noise, we used a threshold-based method. We set the threshold according to the average positive amplitude of the cross-correlation output. We first computed this average value and then set the threshold to be four times of the average value. Fig. 7 shows the calculated threshold for the propagation channel. It can be seen that the noise in the cross-correlation output is mainly below the threshold. The solution using a chirp signal with a matched filter presents a good performance in noisy environments.

4.4 ToA point selection

After filtering the noise, in the remaining data blocks all non-zero values are either generated by the matched filter's output or by the noise in the channel. If they are indeed from the matched-filter's output, they have to fit in (18) which represents the valid output of the matched filter

$$c[n] = \sqrt{\frac{BN}{F_s}} \cdot \text{sinc}\left[\pi B\left(\frac{n - N_{\text{ToA}}}{F_s}\right)\right] \cdot \cos\left[2\pi f_0\left(\frac{n - N_{\text{ToA}}}{F_s}\right)\right] \quad (18)$$

where n indicates sampling points index, F_s is the sampling rate, N is the signal's length or the total number of sampling points. For $N_{\text{ToA}} = 6615$, $F_s = 44.1$ kHz, the positive part of $c[n]$ is shown in Fig. 8. Three features can be extracted from Fig. 8, (a) the main lobe of $c[n]$ has nine positive audio data points, (b) the two adjunct positive audio data in the main lobe are separated by two or three sampling points from each other and (c) the peak audio data corresponding to the ToA is the local maximum in the main lobe. Therefore, the remaining non-zero values should fit the above-mentioned features if they are generated by the matched filter.

We process the remaining positive audio data by: (a) separating the first audio data block containing the ToA point and (b) discarding the noise after matched-filtering. In the remaining positive data, as shown in Fig. 9 we can have two intuitive

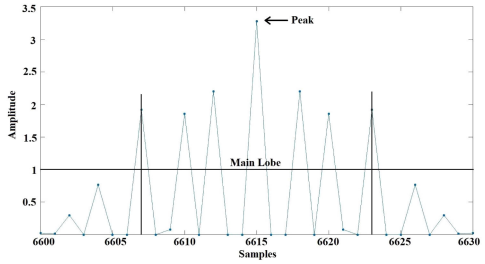

Fig. 8 *Positive part of c[n] with* $N_{\mathrm{ToA}} = 6615$, $F_{\mathrm{s}} = 44.1$ kHz

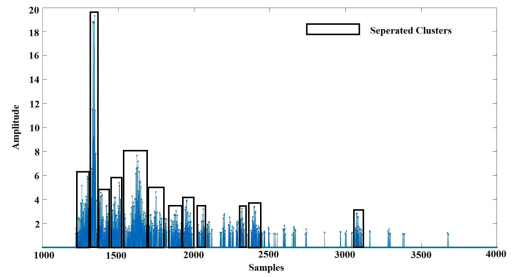

Fig. 9 *Separated audio data clusters*

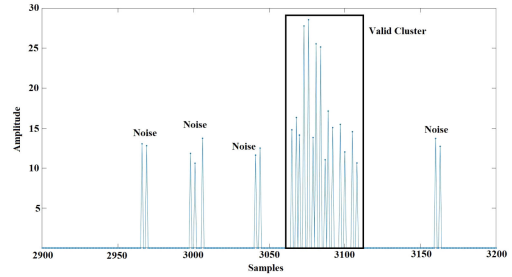

Fig. 10 *Valid cluster and noise*

observations: (a) within certain segments there are audio data which are very close to each other to form a cluster, some clusters have more data while others have less, (b) two clusters are separated from each other with a relatively large interval.

As previously demonstrated, due to the fact that natural environmental noise is not likely to correlate with the reference chirp signal, it will not generate a high amplitude signal at the matched-filtering's output. Since $c[n]$ in (19) has the main lobe with nine positive audio data, as shown in Fig. 10, a data cluster with more than 9 positive audio data can be considered as a valid data cluster which may contain the ToA point, otherwise they are generated by noise and should be discarded. Moreover, if two blocks of audio data are generated from the target audio signal transmitted through two different paths, due to the propagation time difference, the two blocks of data should be apart from each other with a larger interval compared to the intervals between two neighbouring data in the main lobe of $c[n]$.

To summarise, a group of audio data with the minimum of nine positive audio data points can be considered as a valid audio data cluster otherwise it will be discarded. If the two positive audio data points are separated from each other with the interval larger than three sampling points, they should be considered as two different clusters. The algorithm initially searches the first positive audio data point, and then it tries to find the next positive one within three sampling points' interval. If it can find the next positive one within three sampling points' interval, the search for the next positive one will continue. If after three sampling points it still cannot find a positive value, it considers the last positive audio data as boundary of the cluster. The total amount of positive audio data is used to determine whether a valid cluster is found or not. Fig. 9 shows the separation of all the positive audio data based on the above procedure. In this way, the first cluster can be identified to search for the ToA point, and the remaining noise data will not affect the result as they are discarded.

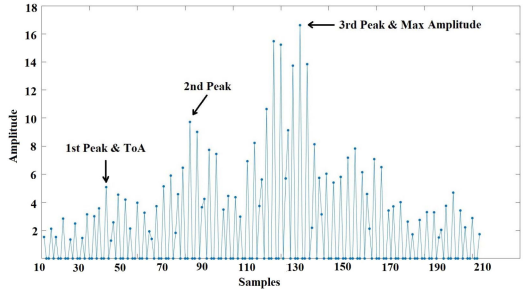

Fig. 11 *First local maxima as ToA*

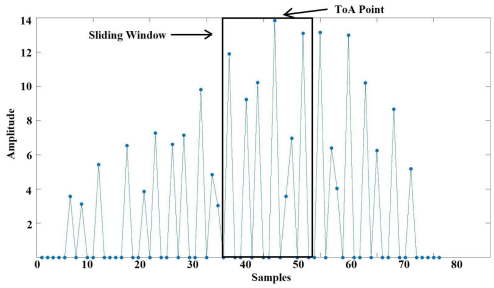

Fig. 12 *Local maxima indicating the ToA point*

After separating out the first cluster, we search for the ToA point within the cluster. If the output is indeed the ToA, it should fit $c[n]$. In a real scenario, especially in a dense multipath environment, we may get close echoes. In this case, the global maxima may not be the true ToA, but a ToA point is definitely the local maxima according to the third feature of $c[n]$. Fig. 11 shows the matched-filter's output in an environment with severe multipath effects. The first local maxima have a lower amplitude compared to the second and the third local maxima but it corresponds to the real ToA point. We used a sliding window with the size of $d = 9$ to search for the first local maxima of all positive values, as shown in Fig. 12. The point where the first local maxima locates is the estimated ToA point. In an environment with severe multipath effect, the window size d can be decreased to distinguish between close echoes.

4.5 NLoS identification

As shown in Fig. 13, although there are multipath propagations, there is a direct path channel. Thus, the first detectable cross-correlation peak is the ToA. However, if there is an obstacle between the audio receiver and the smartphone speaker, LoS propagation channel is blocked, as indicated in Fig. 14. In this case, the received signal consists of reflected, refracted and diffracted signals. The first local maxima under this circumstance do not represent the ToA. Therefore, we have to identify the NLoS propagation condition and discard the corresponding sampled audio data. When there is no line-of-sight, the number of correlation peaks will increase. Moreover, the correlation peak amplitude decreases compared to LoS propagation. Figs. 15*ab* show an experiment conducted under the same conditions, except in the first one there is a line-of-sight connection while in the second one there is not. We can clearly see that the cross-correlation for the NLoS has more correlation peaks, the output is 'blurred' and the correlation peak amplitudes are lower than the one with the LoS path.

When an audio signal hits an obstacle, the obstacle refracts or reflects the signal generating more correlation peaks at the matched filter's output compared to the LoS condition where the correlation peak is generated by only the direct path signal, as shown in Fig. 14. Moreover, reflected audio signals experience energy loss which causes the cross-correlation peak amplitude to decrease.

In addition to the audio-based positioning, the proposed system also supports RSSI-based location estimation. The RSSI positioning algorithm is less affected by NLoS propagation, it supports location positioning with a metre-level accuracy. The

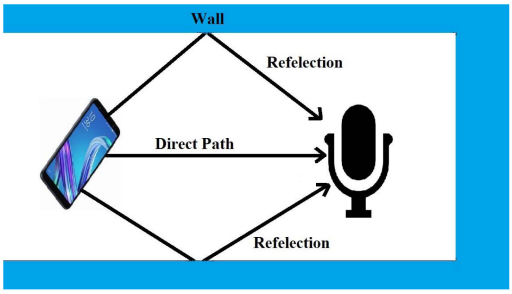

Fig. 13 *LoS condition*

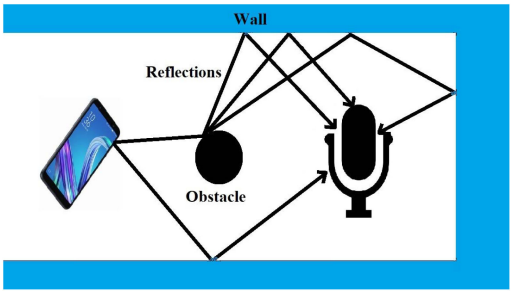

Fig. 14 *NLoS condition*

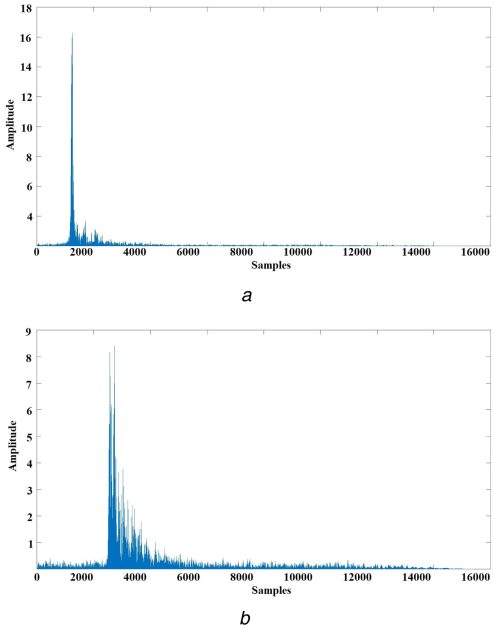

Fig. 15 *Output of the cross correlation for an experiment where*
(a) With a line-of-sight connection, ***(b)*** Without a line-of-sight connection between transmitters and receivers

RSSI-based positioning is used as a reference to determine if the position reported in the acoustic mode is valid or not. In the proposed solution, if the differences between the reported position in the acoustic mode and the BLE mode become larger than a threshold Th_{dif}, the system dumps the result.

5 Experimental validation

5.1 System prototype

The area of interest is divided into grids where each grid is populated with BLE tags. Each tag consists of SPK0415HM4H microphone, CC1310 sub-1 GHz wireless microcontroller to synchronise audio receivers and CC2640 BLE microcontroller for communication with users. Each tag is also equipped with the required resources to determine the time-of-arrival for transmitted audio signals. A smartphone app is developed using Apache-Cordova platform to test the performance of the implemented positioning system.

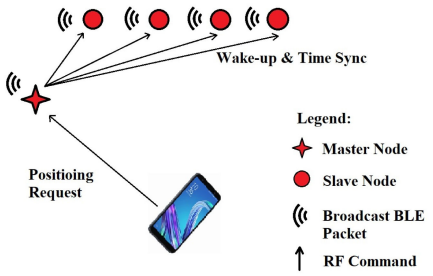

Fig. 16 *Wireless communication*

5.2 Wireless communication

As shown in Fig. 16, an audio receiver in a grid serves as a master BLE access point which receives commands directly from the smartphone. The CC2640 BLE microcontroller at all nodes broadcast BLE packets periodically, these packets can be received by a smartphone to determine its location using the RSSI-based positioning algorithm. These packets also contain audio receiver's ToA estimation results which can be extracted by the developed smartphone app. Once the master node receives the positioning request from a smartphone, it sends a command to the slave nodes in the same grid to wake them up. The slave nodes initially try to synchronise with the master node through sub-1 GHz RF. Once they are synchronised, the microphones sample the incoming audio signal to determine the ToAs and estimate the location accordingly. As the system scales, master nodes in different grids may transmit inter-node commands simultaneously, as a result a slave node in a grid may receive several commands at the same time. A grid identifier is used in the inter-node command to allow slave nodes distinguish commands from other grids and discard them. A high-accuracy time synchronisation is vital for accurate positioning. The measured time-synchronisation error is <30.52 µs corresponding to a TDoA estimation error of ∼ 1.04 cm.

The CC2640 and CC1310 modules are normally in sleep mode, their radio components are only activated for a very short time which reduces the power consumption significantly.

5.3 Optimal node placement

To minimise the TDoA errors, the tags have to be properly positioned in the area, an optimal node placement has a good TDoA error tolerance. A two-dimensional optimal sensor placement for TDoA positioning has been proposed in [28]. To cover a 3D rectangular space in this work, we placed the first four nodes at the rectangle's vertexes and positioned the 5th node at the geometric centre of the first four nodes.

We simulated a rectangular room with the length and the width of 20 m, the height of 3 m using the above-mentioned node placement strategy. We define the system's TDoA error tolerance ability as

$$\Delta\text{Error} = \frac{|E_x| + |E_y|}{\max\ (|E_{\text{TDoA}}|))} \tag{19}$$

where $|E_x|$ and $|E_x|$ are the absolute errors of the estimated x and y coordinates, respectively, and max ($|E_{\text{TDoA}}|$) is the maximum of absolute TDoA estimation error. It is clear that ΔError should be as low as possible to support a good error tolerance ability. During the simulation, each node's z coordinate changes from −25 to 25 cm with the step size of 1 cm. For each z coordinates combination, we randomly selected 1000 points in the rectangle area to compute the average of ΔError. We found that as the z coordinates vary, the average of ΔError shows a periodical change as shown in Fig. 17. Within a certain segment there are several steps which show similar $\overline{\Delta\text{Error}}$ as local minimums. This shows that there are several combinations of z coordinates supporting optimal error tolerance abilities.

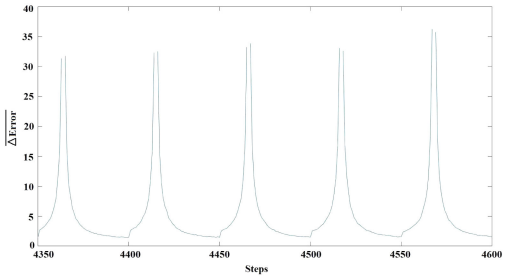

Fig. 17 *TDoA error tolerance ability with respect to z coordinates*

Fig. 18 *Area where the experimental measurements were performed*

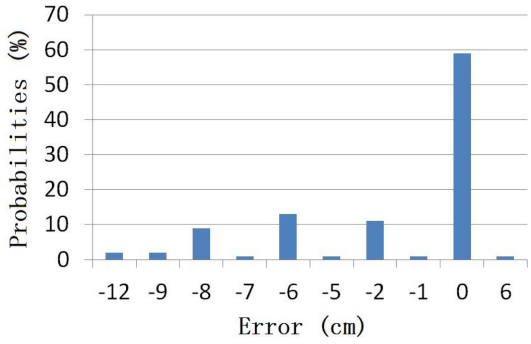

Fig. 19 *TDoA experiment results in a quiet environment*

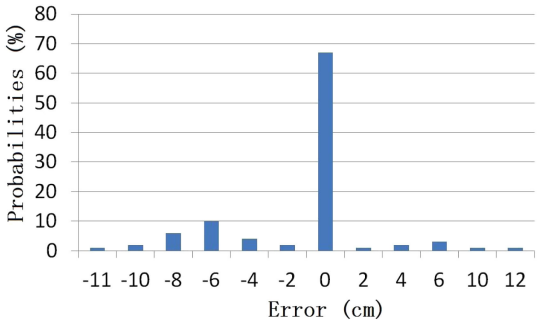

Fig. 20 *TDoA experiment results in a noisy environment*

5.4 Experimental setup

Since the key parameter in the proposed system is the TDoA estimation. The TDoA estimation experiments were conducted for three scenarios of (a) the smartphone is positioned 20 m away from the audio receivers in a quiet environment. (b) The same setup in a noisy environment with an estimated −20 dB SNR and (c) a dense multipath environment where the smartphone is placed 3 m far from the audio receivers. The location positioning experiments were conducted in a medium size room with both LoS case and NLoS paths.

The first two experiments were conducted in the hallway of the CEI building at the University of Windsor, as shown in Fig. 18. Audio receivers and smartphone are placed 20 m far from each other. For the noisy environment case, an audio player was placed next to the audio receiver tags playing a recording. The estimated SNR was ∼ −20 dB for the transmitted signal. Experiments were also conducted in the CEI building stair entrance. The stair area is small with concrete walls and can be considered an environment with dense multipath effects. In each scenario, TDoA measurements were repeated 100 times.

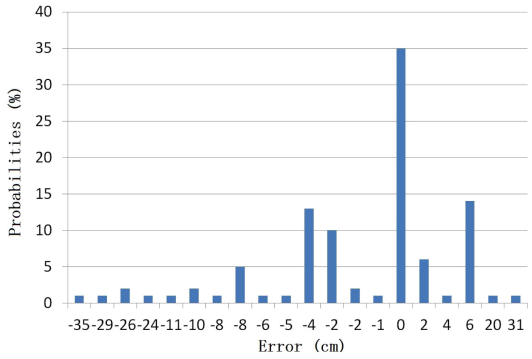

Fig. 21 *TDoA results under dense multipath effects*

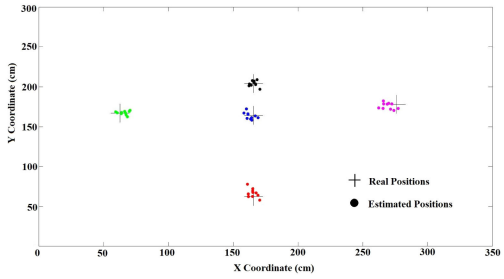

Fig. 22 *Position estimation results*

Table 2 Performance comparisons

Techniques	Accuracy	Smartphone Compatibility
proposed	<6 cm	yes
UWB [4]	centimetre	no
BLE [8]	<2.1 m	yes
ultrasound [9]	1–2 cm	no

For the location positioning experiment, four nodes were placed at the vertexes of a rectangle area with a slight difference in the heights, the fifth node was placed at the geometric centre of the other four nodes. We did experiments at different locations within the rectangle area, ten times LoS case experiments and five times NLoS condition experiments were repeated for each location. Figs. 19–22 show the measurement results.

5.5 Experiments results

For the first experiment, the mean absolute TDoA estimation error was 2.26 cm with the maximum error of 12 cm and the standard deviation of 3.46 cm. In the second experiment, the mean absolute TDoA estimation error was 2.09 cm and the maximum error was 12 cm with the standard deviation of 3.75 cm. In the third experiment, the mean absolute error was 4.53 cm, with the maximum error of 34.5 cm, and standard deviation of 8.16 cm. As for the location positioning with the LoS path, the estimated x coordinates had the mean absolute error of 3.71 cm, while the error was 2.48 cm for the y coordinate. The x and y coordinates had the maximum error of 15.40 and 15.20 cm, respectively. The average of horizontal Euclidean distances error was 5.91 cm and the maximum error was 16.11 cm.

Table 2 shows the performance comparison with the reported works in the literature. The positioning accuracies of systems in [4, 9] are comparable or even better than the proposed system. However, these systems require custom-design hardware and they are not compatible with smartphone platforms, which is a considerable disadvantage.

6 Conclusion

We have proposed a practical indoor positioning solution using RSSI for coarse positioning and near ultrasound audio signal for fine positioning. The proposed solution utilises BLE nodes for indoor positioning and utilises cellphones to determine a user's locations. The proposed solution can operate without Internet connectivity since the location is determine by user's cellphone without the need to a server. To reduce the positioning error and overcome the environmental audio noise, a chirp signal is utilised. A novel digital signal processing algorithm using a matched filter is presented to accurately determine the time-of-arrival for audio signals. The TDoAs are then used for location estimation. RSSI positioning has been used to determine whether a line-of-sight exist for the audio-based positioning. Experimental measurement results under different conditions indicate that the proposed solution can successfully determine the location of a user with <6 cm positioning error on average.

7 References

[1] Pagano, S., Peirani, S., Valle, M.: 'Indoor ranging and localisation algorithm based on received signal strength indicator using statistic parameters for wireless sensor networks', *IET Wirel. Sens. Syst.*, 2015, **5**, (5), pp. 243–249, doi: 10.1049/iet-wss.2014.0027

[2] Potortì, F., Cassarà, P., Barsocchi, P.: 'Device-free indoor localisation with small numbers of anchors', *IET Wirel. Sens. Syst.*, 2018, **8**, (4), pp. 152–161, doi: 10.1049/iet-wss.2017.0153

[3] Potortì, K.H., Yang, S.H.: 'Improved adaptive localisation approach for indoor positioning by using environmental thresholds with wireless sensor nodes', *IET Wirel. Sens. Syst.*, 2015, **5**, (3), pp. 157–165, doi: 10.1049/iet-wss.2013.0100

[4] Alarifi, A., Al-Salman, A., Alsaleh, M., *et al.*: 'Ultra wideband indoor positioning technologies: analysis and recent advances', *Sensors*, 2016, **16**, (5), p. 707

[5] Reynders, B., Pollin, S.: 'Chirp spread spectrum as a modulation technique for long range communication'. 2016 Symp. on Communications and Vehicular Technologies (SCVT), Mons, 2016, pp. 1–5, doi: 10.1109/SCVT.2016.7797659

[6] 'iBeacon – Apple Developer', https://developer.apple.com/ibeacon, accessed 06 Mar. 2019

[7] 'Indoor Positioning System for Enterprises | powered by indoo.rs', https://indoo.rs, accessed 06 Mar. 2019

[8] Zaki, F., Rashidzadeh, R.: 'An indoor location positioning algorithm for portable devices and autonomous machines'. Int. Conf. on Indoor Positioning and Indoor Navigation (IPIN), Alcala de Henares, Spain, 2016

[9] Priyantha, N., Chakraborty, A., Balakrishnan, H.: 'The cricket location-support system'. Proc. ACM Conf. on Mobile Computing and Networking, 2000

[10] Ward, A., Jones, A., Hopper, A.: 'A new location technique for the active office', *IEEE Pers. Commun.*, 1997, **4**, (5), pp. 42–47

[11] Mandal, A., Lopes, C.V., Givargis, T., *et al.*: 'Beep: 3D indoor positioning using audible sound'. Second IEEE Consumer Communications and Networking Conf., 2005. CCNC. 2005, Las Vegas, NV, 2005, pp. 348–353

[12] Höflinger, F., Zhang, R., Hoppe, J., *et al.*: 'Acoustic self-calibrating system for indoor smartphone tracking (ASSIST)'. 2012 Int. Conf. on Indoor Positioning and Indoor Navigation (IPIN), Sydney, NSW, 2012, pp. 1–9

[13] Balani, R.: 'Energy consumption analysis for bluetooth, WiFi and cellular networks', Document, Electrical Engineering, University of California at Los Angeles

[14] Sérgio, I.L., Vieira, J.M.N., Reis, J., *et al.*: 'Accurate smartphone indoor positioning using a WSN infrastructure and non-invasive audio for TDoA estimation', *Pervasive Mob. Comput.*, 2015, **20**, pp. 29–46

[15] Patrick, L., Niranjini, R., Oliver, S., *et al.*: 'ALPS: A bluetooth and ultrasound platform for mapping and localization'. 13th ACM Conf. on Embedded Networked Sensor Systems (SenSys 2015), Seoul, South Korea, 2015, pp. 73–84

[16] Murata, S., Yara, C., Kaneta, K., *et al.*: 'Accurate indoor positioning system using near-ultrasonic sound from a smartphone'. Eighth Int. Conf. on Next Generation, Oxford, UK, 2014

[17] Pérez, M.C., Gualda, D., Villadangos, J.M., *et al.*: 'Android application for indoor positioning of mobile devices using ultrasonic signals'. 2016 Int. Conf. on Indoor Positioning and Indoor Navigation (IPIN), Alcala de Henares, 2016, pp. 1–7

[18] Zhang, L., Chen, M., Wang, X., *et al.*: 'TOA estimation of chirp signal in dense multipath environment for low-cost acoustic ranging', *IEEE Trans. Instrum. Meas.*, 2019, **68**, (2), pp. 355–367, doi: 10.1109/TIM.2018.2844942

[19] Zhang, L., Huang, D., Wang, X., *et al.*: 'Acoustic NLOS identification using acoustic channel characteristics', *Sensors*, 2017, **17**, (4), p. 727(SCI/EI)

[20] 'Dev.ti.com. CC13 × 0 RF User's Guide — CC13 × 0 Proprietary RF User's Guide 1.01.00 documentation', http://dev.ti.com/tirex/content/simplelink_cc13×0_sdk_1_30_00_06/docs/proprietary-rf/html/cc13×0/index.html, accessed 17 Dec. 2018

[21] Stoica, P., Li, J.: 'Lecture notes – source localization from range-difference measurements', *IEEE Signal Process. Mag.*, 2006, **23**, (6), pp. 63–66

[22] Mensing, C., Plass, S.: 'Positioning algorithms for cellular networks using TDOA'. 2006 IEEE Int. Conf.Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proc., May 2006, vol. 4, pp. IV–IV

[23] Jin-Yu, X., Wei, W., Zhong-Liang, Z.: 'A new TDOA location technique based on Taylor series expansion in cellular networks'. Proc. of the Fourth Int. Conf. on Parallel and Distributed Computing, Applications and Technologies, Chengdu, China, 2003, pp. 378–381

[24] Hossain, S.A., Anower, M.S., Halder, A.: 'A cross-correlation based signal processing approach to determine number and distance of objects in the sea using CHIRP signal'. 2015 Int. Conf. on Electrical & Electronic Engineering (ICEEE), Rajshahi, 2015, pp. 177–180

[25] 'En.wikipedia.org. (2018). Pulse compression', https://en.wikipedia.org/wiki/Pulse_compression, accessed 22 Nov. 2018

[26] Sheugh, L., Alizadeh, S.H.: 'A note on Pearson correlation coefficient as a metric of similarity in recommender system'. 2015 AI & Robotics (IRANOPEN), Qazvin, 2015, pp. 1–6, doi: 10.1109/RIOS.2015.7270736

[27] 'What does it mean if the correlation coefficient is positive, negative, or zero?', https://www.investopedia.com/ask/answers/032515/what-does-it-mean-if-correlation-coefficient-positive-negative-or-zero.asp, accessed 22 Nov. 2018

[28] Lui, K., So, H.C.: 'A study of two-dimensional sensor placement using time-difference-of-arrival measurements', *Digit. Signal Process.*, 2009, **19**, (4), pp. 650–659, doi: 10.1016/j.dsp.2009.01.002