RAY: Improving human Interaction Using Acoustic-based Encounter Profiling and Smartphones

Amirush Javare
California State University: Sacramento
amirushramdasjavare@csus.edu

Yashasri Edukulla
California State University: Sacramento
yedukulla@csus.edu

Raghavendra Acharya
California State University: Sacramento
raghavendrasreeniva@csus.edu

Abstract

We nowadays encounter many people in different places like events or conferences or general meetups. But we can't remember the people we have met and had a talk with. So, people identifications with their place of the meeting are the need of the day. In this paper, we have tried to improvise the traditional way people interact with each other. We have developed an android application which is RAY that emits acoustic signals transmission by using sensors from the smartphones which includes speaker, microphone, and accelerometer. The process works in 2 phases first is detecting the approach when the 2 people stop in front of each other and second is voice profiling which gives the evidence that the 2 people are interacting. RAY is also able to support multiple access schemes within the limited inaudible acoustic frequency band. The main objectives of RAY are to identify the encountered persons and record the interaction context and information about the location as well. We have evaluated RAY in different scenarios and also with multiple users and obtained the accuracy of 6.81% as false positive and 9.46 % as false negative when we performed experiment in real world.

Keywords: *Acoustic signals, sensors, speaker, microphone, voice profiling, inaudible acoustic frequency, multiple access.*

1. Introduction

Individuals regularly encounter and interact with different and numerous people in a day or during a social occasion. Encounter profiling aims at identifying and logging these interactions. Our project aims at analyzing and collecting information focused on his/her interactions with other individuals. This information is important for lifelogging and memory assistance. We plan to do it by creating an android application which is the name "RAY", which will use the principles of DOPPLER EFFECT that will have the capability to collect information mainly about: (1) Who has the owner of the application spoken to today? (LIST OF INDIVIDUALS) (2) How much time did he actively interact with each individual? (DURATION) (3) Where did they encounter each other? (PLACE). We also have alternate options to do this task which are Sony's Lifelong and Google's Keep, but both of these require continuous instructions to be given from time to time by the owner, unlike the above application which does its task even without any assistance. The normal process of interaction between any two humans is they meet and then converse. So, the major steps for the application are to be capable of doing: Trajectory Analysis and Conversation confirmation. The application does trajectory analysis which will further allow smartphones to broadcast acoustic signals and by utilizing it which is Doppler effect to find the relative velocities between 2 persons. Every user has to undergo a process that involves recognition of their voices so that when another person comes into conversation the application can find out somebody else's presence using the difference in their frequencies. This is how conversation confirmation happens. This logging is very important to individuals who want to track the content of conversations they have had with a lot of people.

Example: If Ram is attending a social event, he would have had a small conversation with Sam and additionally, should have had small chats with other people. Once Ram is back, he is able to retrieve the information about those chats without any assistance of his efforts in the social event. Existing techniques that can detect human presence using handshaking with Skin Potential Level (SPL) sensors, identify human beings through trajectory pathway detection or detect human proximity with short distance communication like Bluetooth Low Energy devices. This paper produces the application that has high goals of identifying user's interaction by leveraging acoustic signal transmission with common sensors on mobile phone devices. **Moreover, none of the existing work [1][2][3][4] has taken into consideration the analysis for human walking trajectories and they all focus on single link measurements and do not address the multiple access problem. The work carried out in [5][6][7] also involves complex pattern recognition and comparison algorithms of high computational overhead and cannot be directly used for our application RAY due to long processing time and energy consumptions. Also, work carried out in [8][9] does not consider coordinating multiple access for Doppler effect measurement as RAY does. Therefore, RAY has been developed to overcome the drawbacks that were seen in the above-mentioned works.**

2. Related Work

1. Md Tanvir Islam Aumi, Sidhant Gupta, et.al.: This paper authrs have presented DopLink which is an ultrasonic-base device selection approach. It utilizes the embedded hardware from smartphones to determine if a particular device is pointed at another device which is the user waves their phones at a target in a pointing motion. The authors have tested and achieved accuracy of 95% for device selection and 97% in finding the device position [1].

2. Sidhant Gupta, Daneil Morris, Desnet Tan, et.al.:The authors have presented a new technique that leverages the speaker and microphone of smartphones to sense in air gestures around the device by generating inaudible tone, which get frequency shifted when it reflects off moving objects for example hand movements. [2].

3. Sangki Yun, Yi-Chao Chen, Lili Qiu: The authors have developed a system that can accurately track hand movements to realize a mouse. This is done by sending inaudible pulses at a few selected frequencies and uses the frequency shifts to find the speed and distance travelled. Then authors further enhanced the system by developing some techniques to calibrate the distance between the speakers and narrow down the device position using its movement trajectory and continuously track the device new position in real time [3].

4. Wenchao Huang, Yan Xiong Xiang Yang Li, et.al.: This paper proposes an acoustic direction-finding scheme which is named as Swadloon. Swadloon tracks the displacement of smartphone relative to acoustic direction with the resolution of less than 1 millimetre. The direction is then founded by combining the velocity from the displacement with the one from the inertial sensors. [4].

5. Douglas A. Reynolds, Thomas F. Quatieri, Robert Dunn: In this paper the authors have described the major elements of speaker verification which is used many NIST speaker recognition evaluations. The model revolves around likelihood ratio test for some verifications using some simple and efficient GMMs which is nothing but Gaussian Mixture model and also universal background model (UBM) is also take into consideration for presentation purpose. Moreover, they used Bayesian adaption to generate speaker models from UBM [5].

6. Hong Lu, Bodhi Priyantha, Amy K Karlson, et.al.: The authors have designed a speaker sense which is a prototype for speaker identification which takes into account heterogenous multiprocessor hardware architecture which further splits the computation that is to be carried on the lower power processor and the phone application processor. This will enable doing continuous background sensing with minimal power [6].

7. Chengwen Luo, Mun Choon Chan: This paper proposes and evaluates SocialWeaver, a sensing service running on smartphones that performs conversation clustering and builds conversation networks automatically. SocialWeaver uses a hybrid speaker classification scheme that exploits an adaptive histogram-based classifier to non-obtrusively bootstrap. The conversation clustering algorithm proposed is able to detect fine-grain conversation groups even if speakers are close together. Finally, to address energy constrain, a POMDP-based energy control scheme is incorporated [7].

8. G. Enrico Santagati, Tommaso Melodia: In this paper the authors have developed U-Wear which is a networking framework for wearable medical devices based on ultrasonic communication. It encloses a set of physical, data link and networking layer functionalities which adapts to application and system requirement to distribute information between ultrasonic wearable devices. Moreover, it also offers reconfiguration feature as well to provide a flexible platform to develop medical applications [8].

9. Rajalakshmi Nandakumar, Krishna Chintalapudi, et.al.: In this paper author have addressed different challenges of enabling NFC on smartphones. So, the authors have developed Dhwani which is a novel acoustic based NFC system that uses the microphone and speakers of mobile devices thus eliminating the need for NFC hardware. One more key feature of Dhwani is the Jam-Secure technique which uses self-jamming coupled with self-interference cancellation at the receiver to provide information theoretically secure communication channel between the devices [9].

3. Concept of Doppler Profiling

In this section we are going to describe the process of how RAY derives the doppler profiles of the acoustic signals transmitted between smartphones.

3.1. Doppler Profiling in Smart Phones

The term doppler refers to the changes in the frequency of waves when some observer is moving towards or away from the source is given by the equation [10].

$$\Delta f = (\Delta v / c) * f_0 \qquad (1)$$

Here in the above equation Δf = f – f0 which is the received frequency f subtracted from emitting frequency. C is the speed of the waves observed. If f is detected we can easily calculate the frequency offset which further gives us the velocity of the person moving. Below is the screenshot of the 2 persons who walk towards each other. It can be observed that velocity is positive when the 2 persons are walking towards each other and in figure 3.1 and comes to zero when the 2 person stops in front of each other.

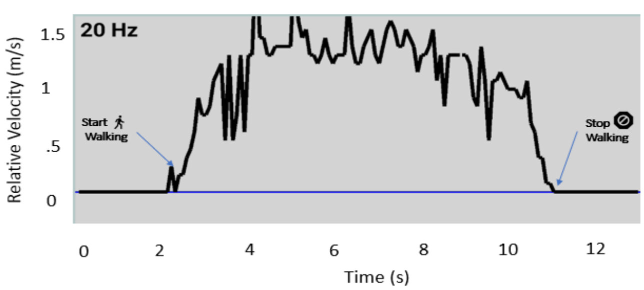


Figure 3.1.1 Example for determining start and Stop

Figure 3.1.1 depicts a simple example that we have carried out with the RAY. It can be seen when the person is standing the velocity is stationary but when the person starts to move towards another person there is a change observed in the velocity and moves at the positive side and again comes back to stationary when the walking person stops beside the other person who is the sender.

We have set the microphone sampling rate to 48Khz which is supported by most of the smartphones today and further performs 4096 FFT which is nothing but 1 FFT for 4096 samples which can be handled by smartphones easily. The corresponding velocity resolution is only 22.5 cm/s in equation (1) [20]. Therefore, to overcome this we have done under-sampling which reduces temporal fidelity as the interval between the 2 adjacent samples can be enlarged. Under-sampling has shown better results in Spartacus [12]. Oversampling is also utilized to provide high temporal fidelity which further extends the utilization into the past sampling data and results in a improved sliding window for FFT. In all these 2 techniques have helped to achieve velocity resolution better than which is obtained from equation (1).

3.2. Study of Different Trajectories

A particular trajectory is effective when a 2-person approach in front of each other have a conversation. We have categorized 5 different approaches to trajectories. First, is the direct approach when 2 persons walk directly towards each other it is shown in figure 3.2.1. Direct approaches output effective trajectories as the 2-person approach and they stop in front of each other. Figure 3.2.2 shows the Velocity indirect approach which remains steady all the time.

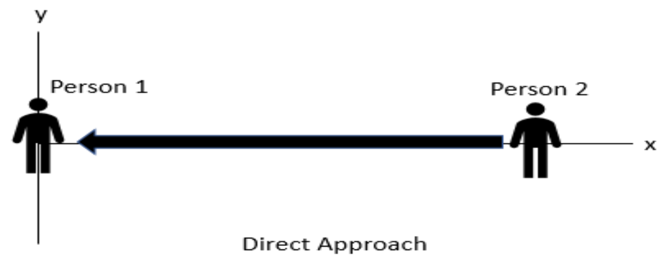


Figure 3.2.1 Direct Approach

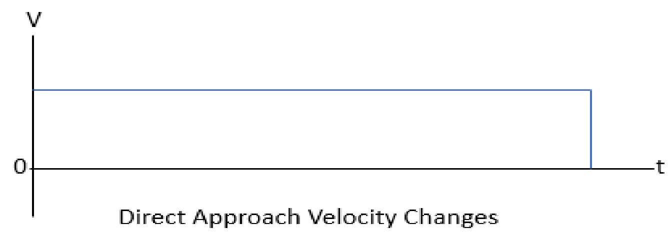


Figure 3.2.2 Direct Approach Velocity Changes

The roundabout approach is the second one in which 2 people walk towards each other in the roundabout format in some cases it might have to pass some obstacles as well. Such trajectories are sometimes effective. Velocity in this case once the obstacle is crossed or passed by the 2 people the velocity increases as they move closer to each other.

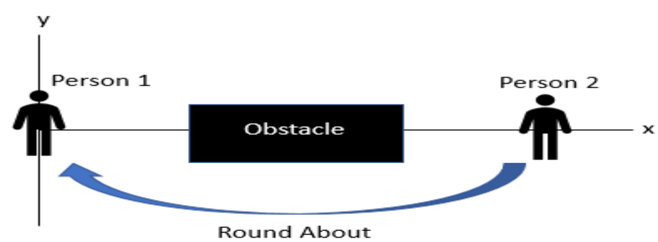


Figure 3.2.3 Round About

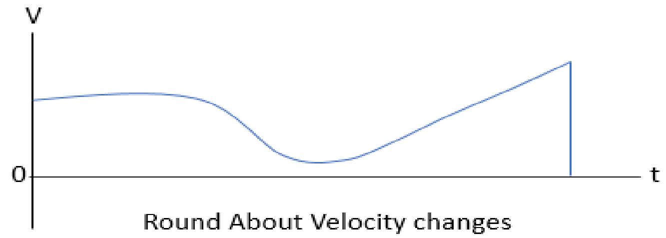


Figure 3.2.4 Round About Velocity changes

Next is the slant approach where 2 person walks slantly but they do not meet each other. The slant approach does not give us an effective trajectory as the persons don't tend to meet each other. In this case, the relative velocity decreases as the heading direction of the 2 persons keep shifting away.

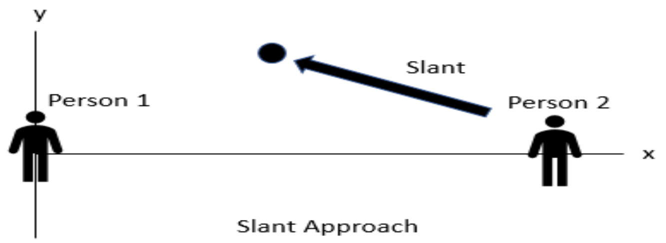


Figure 3.2.5 Slant Approach

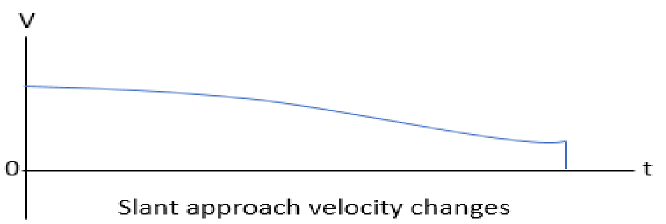


Figure 3.2.6 Slant Approach Velocity changes

Followed by Departing where 2 persons depart from each other and lastly departing and passing by scenarios. In these 2 cases, the velocity goes below 0 as they move away from each other.

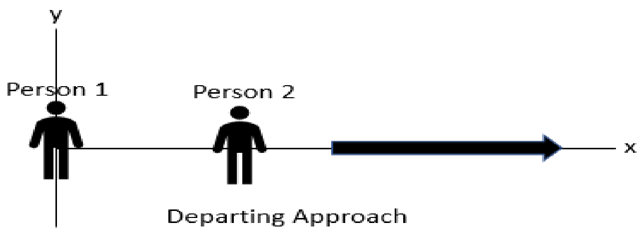


Figure 3.2.7 Departing

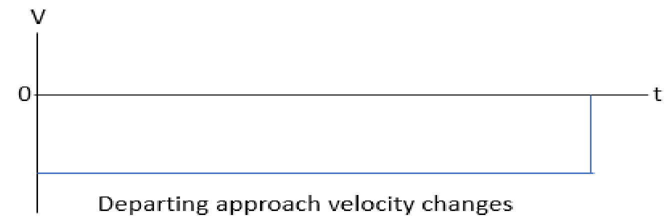


Figure 3.2.8 Departing Approach Velocity Changes

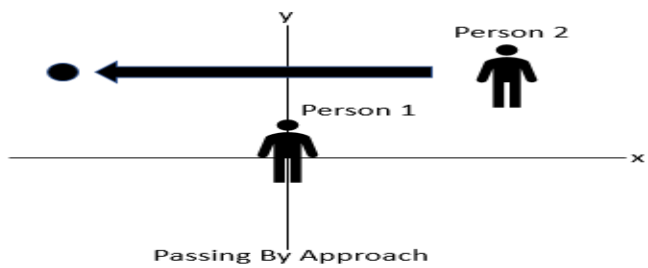


Figure 3.2.9 Passing by Approach

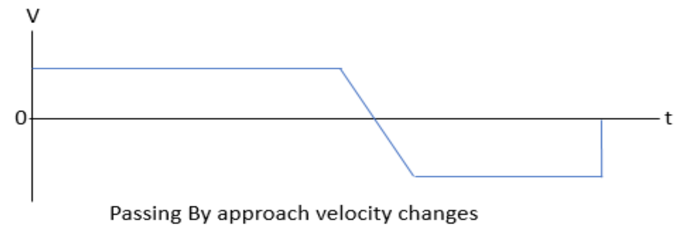


Figure 3.2.10 Passing by Approach Velocity Changes

3.3. Cleaning the Data

As there some noise and errors in the frequency band obtained which is highlighted in the below figure 3.3.1. We have tried the following 3 steps that will make some improvement in RAY.

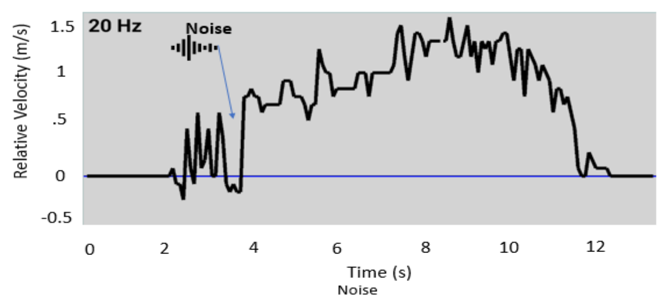


Figure 3.3.1 Noise and Errors in the Frequency Band

Step 1: Smoothing the velocity trace: When a person starts to walk at that time it is observed that peak acceleration occurs at the moment when the person's foot strikes the ground. It was observed in [13]. So, this issue leads to forming valleys in the velocity trace. To overcome the issue, we use moving maximum which exploits the fact that when the human walks steadily except when the foot strikes the ground first. Our technique replaces the relative velocity with maximum velocity within some range of adjacent samples. Moving maximum window is set to one cycle of human walking of approximately 2 - 3 steps. It's observed in many people that it takes 0.5s to make a step which is discovered in [14] the number of velocity samples for 1 walking cycle is 11.7 so our application RAY replaces the relative velocity of different samples with the equation given below [20].

$$V_i = \max\left(V_{i-5}, V_{i-4}, \ldots\ldots, V_i, \ldots\ldots, V_{i+4}, V_{i+5}\right) \qquad (2)$$

Figure 3.3.2 shows the data which we have smoothed by following the technique mentioned above.

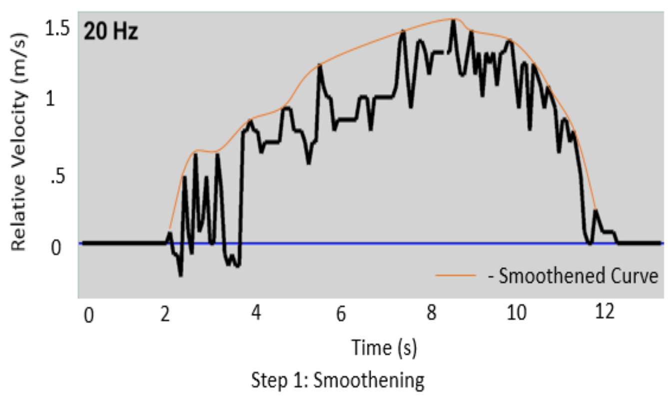


Figure 3.3.2 Smoothening

Step 2: Removing the stop phase of human walking: We have tried to remove the misleading stop phase of human walking as they are not important for trajectory classification. Therefore, the last few seconds data is removed off and the slope is founded for the remaining data. Figure 3.3.3 shows that the data from the last few seconds is removed and then the slope is calculated in the next step.

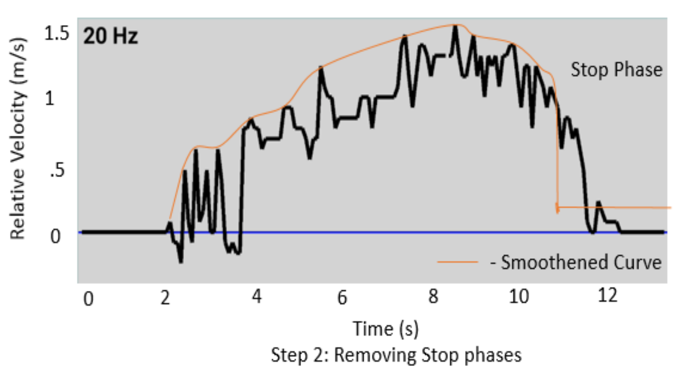


Figure 3.3.3 Removing Stop phases

Step 3: Linear fitting to find slopes: After completing the 2 steps, we applied linear fitting on the data we obtained, and then the slope is calculated of the fitting line to classify the approaching trajectories. Figure 3.3.4 shows the fitting line and slope.

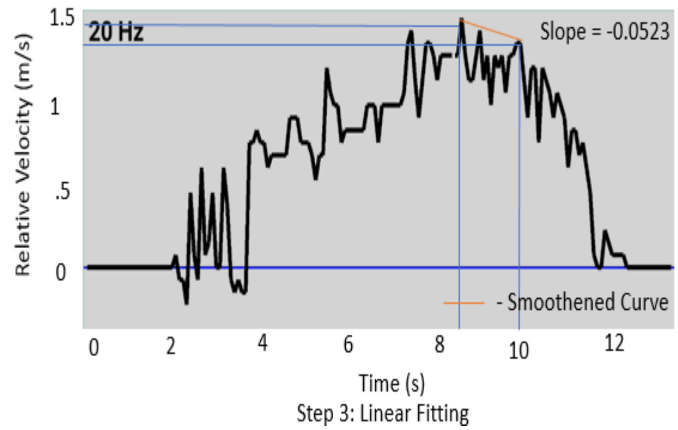


Figure 3.3.4 Linear Fitting

So, for all the trajectories we have cleaned the data that we have obtained.

For direct approach, the fitting line is obtained somewhat downwards but in our case the data this we have obtained consist of velocity going beyond 20Hz that why but it would have been horizontal approximately figure 3.3.5 shows the waves for the direct approach which is obtained after processing the 3 steps..

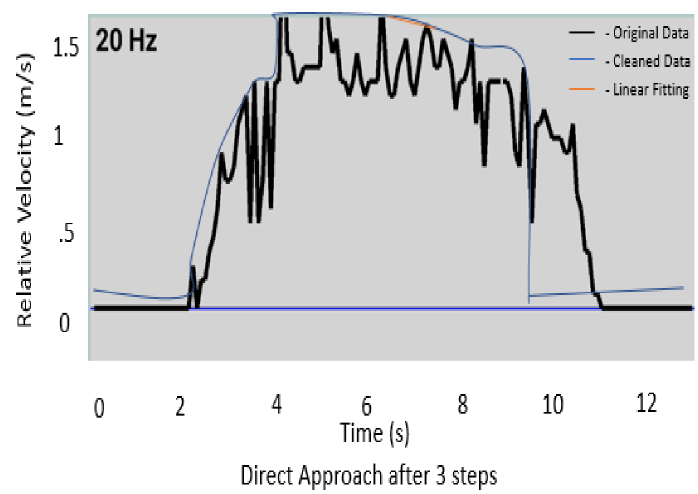


Figure 3.3.5 Direct Approach after 3 steps

Whereas in the roundabout and slant approach it observed that the linear fitting is upwards and downwards respectively. Figure 3.3.6 shows the roundabout and 3.3.6 shows slant approach.

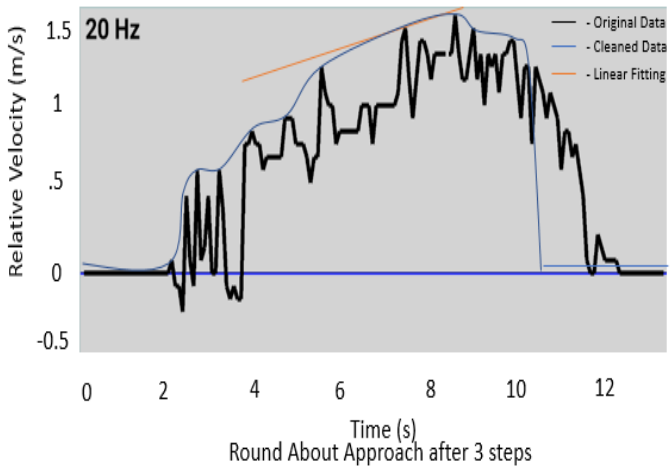


Figure 3.3.6 Round about approach after 3 steps

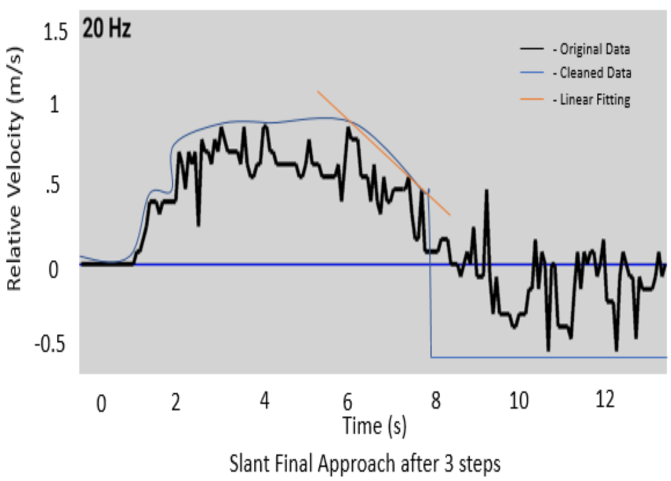


Figure 3.3.7 Slant Approach after 3 steps

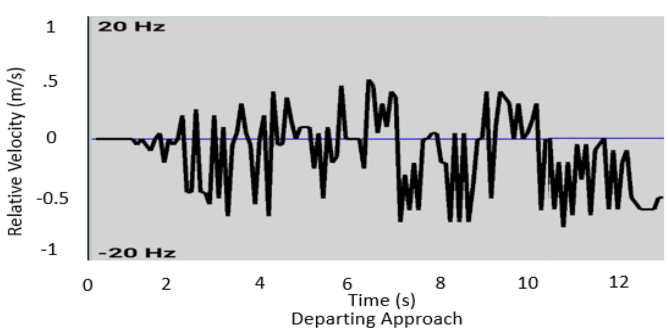


Figure 3.3.8 Departing Approach after 3 steps

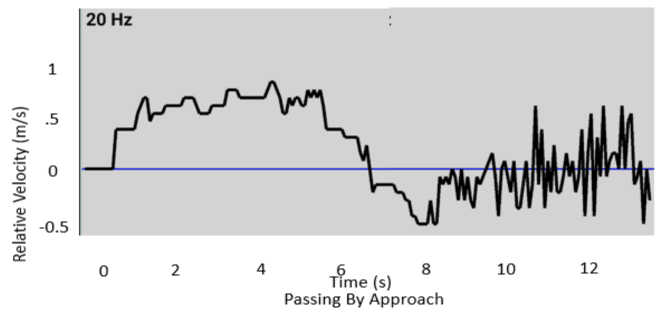


Figure 3.3.9 Passing by Approach after 3 steps

The last 2 approaches have significant differences in the graph that is in departing and passing by they are quite easy to identify as their velocity traces are have a negative spike before the person stops walking. These 2 trajectories do not result in effective trajectories so the 3 steps aren't included for the graph. Figure 3.3.8 shows velocity movement for departing approach followed by passing by the approach in 3.3.9

4. Voice Profiling and Workflow of RAY

RAY can identify the different trajectories that lead to human interactions. Sometimes it may misjudge the 2 persons, for example, let's consider Bob and Alex walks towards each other and Bob walks faster as compared to Alex. So, in the figure below we can see the changes in the velocity as it goes beyond the given limit.

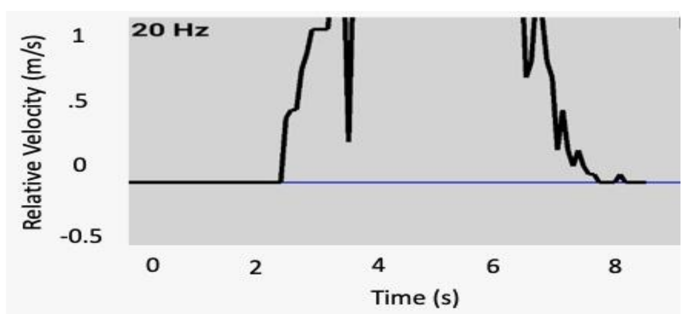


Figure 4.1 Relative Velocity Vs Time

Our application also makes use of voice profiling to make sure that the 2 people are having a conversation with others. In the study of paper [15] suggests that when 2 people are having a conversation, they are approximately 0.5m to 1.5m apart. So, our app uses Doppler profiling to trigger the voice profiling event. It only needs to detect whether there is a conversation and does not take into account the task of recognizing speakers or their voices. Therefore, we don't utilize any features of database or any pattern recognition algorithm. When there's a conversation between 2 people they speak one after the other. So, what the application does is divide one voice trace into N time slots. Which is given as Vi=1 for 1-time slots and so on. The formula for alternative ratio equation 3 is given by [20].

$$AlternativenessRatio = 1 - \sum_{i=1}^{N} \frac{(v_{ai} \,\&\, v_{bi})}{N} \quad (3)$$

Where v_{ai} and v_{bi} == 1 if and only if v_{ai} = 1 and v_{bi} = 1 [20]. The combined voice segments from the 2 speakers can be found by using another equation which is duty ratio. This equation results high when 2 persons are talking for a long time as they have some important topics to talk about. Equation 4 is given by [20].

$$DutyRatio = \frac{\sum_{i=1}^{N} v_{ai} \mid v_{bi}}{N} \quad (4)$$

4.1. Recognizing Owners Voice

Our application tries to utilize the technique that was discovered in [16] to identify the voice of the owner. As the main objective is to identify the owner's voice when he or she listens to the audio recording. So, that's why our application does not demand any voice database for this purpose.

4.2. Exchanging the Voice

When the 2 people meet each other at an event, for example, they do not know each other's identity. Moreover, the signals transmitted by the sender is not able to detect the Id of the other person as well. So, the use of technology like Bluetooth [17] or Wi-fi [18] is not suitable in this case. To overcome this thing RAY first transmits the id of the sender in the same acoustic signals that were earlier used for doppler profiling. The channel is not ready to use as the sender has stopped walking and the process of doppler profiling is completed. When the owner clicks on the start button in the application it sends out acoustic signals with the same duration of its presence of voice. So, at last, by looking at the start and endpoints of such signals the receiver can identify the voice when he hears it after some time by taking into consideration the timestamps of the recording and identify the identity.

4.3. Overall Work Flow of RAY

The workflow consists of 3 phases in the **1st phase when the person who is the receiver starts walking there can be seen changes in the velocity as it comes under the radar of the acoustic signals transmitted by the sender. When the 2 persons stop in front of each other the velocity graph moves to 0. So, now in the 2nd phase they can then have some conversation, and this conversation if the 2 people think is important, they can start recording it by pressing the start button of our RAY application. It also allows the user to know the location of the 2 people where they have met by pressing the Get location button. In the third phase when the person comes home and wants to find the name or identity of the people, he met he can easily hear the recording which is stored in his internal memory of the phone. When supposing example Bob meets 20 people in a day and doesn't know the person, he met at the beginning then in this case he can easily go through the recording that is saved in the internal storage of his phone memory. RAY consists of the Get Location button which is shown in figure 4.3.1 below. When the user clicks on the button the application is able to find the**

location by using the GPS of the phone and gives the Latitude and Longitude of the current position of the user.

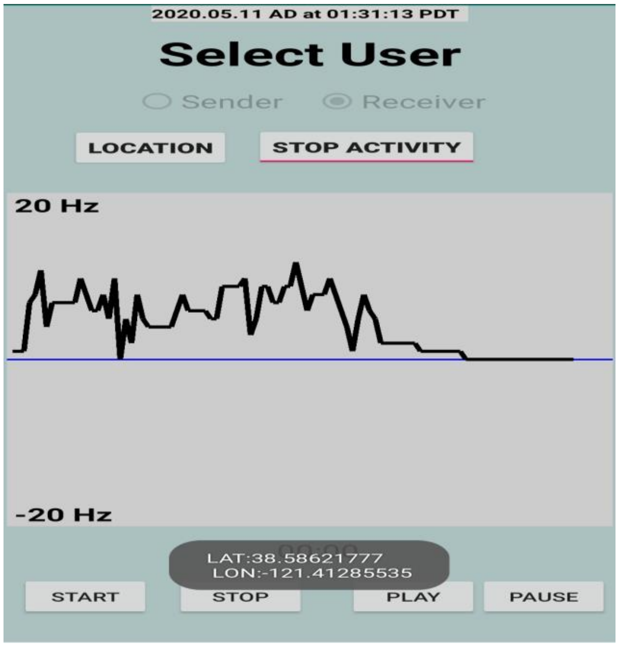


Figure 4.3.1 Get Location Button

Another thing is that if a person thinks that he needs to record the conversations which he feels important for him he ca easily press the Start Button which will enable the microphone to record the voice of the 2 people. Figure 4.3.2 shows the screenshot of the audio button and also it has an option to play the conversation which was recently recorded.

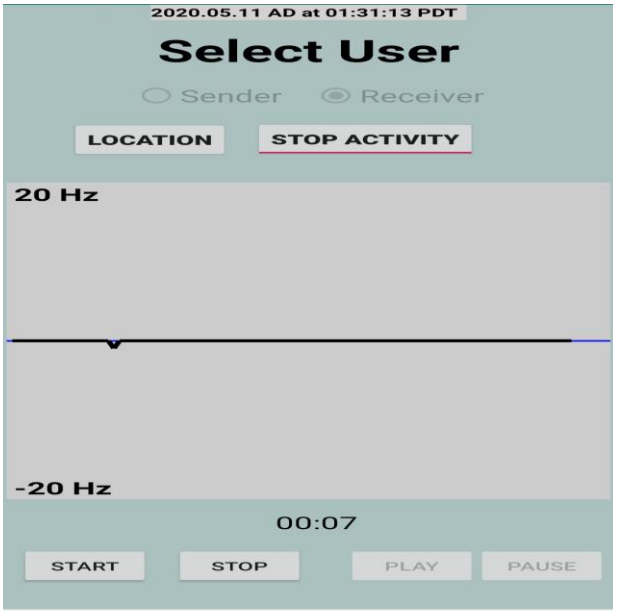


Figure 4.3.2 Audio and play button to listen to Recordings

Lastly, if a person when he goes home wants to find the identity of a particular person, he met for example in the afternoon then can do that as well by going through the recording which is available in the internal memory. The audio recording is stored in the form "Recording yyyy_mm_dd_hr_min_ss" so it becomes easy for the person to find the identity of the people by thinking, for example, Bob met Alex in the afternoon around 3 pm suppose then Bob can easily go through the recording with the timestamp of around 3 pm and hear the conversation. The figure below shows the recording that is stored in the root directory of the smartphone.

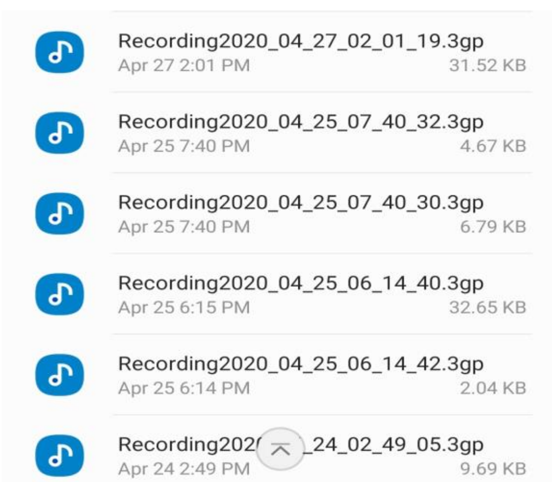


Figure 4.3.3 Saved Recordings in Internal phone Memory

5. Channel analysis

5.1. Synchronizing with More than One User

Until now, we have only talked with respect to the case of one user sending signals and one receiver on the other end. In this part of the paper, we are going to discuss the transmission of signals among multiple users. These synchronizations between multiple users can be broadly classified into four consequences which are 1- transmitter to 1- recipient, 1 – transmitter to many recipients, many transmitters to 1 – recipient, and many transmitters to many recipients. In the case of one transmitter to one recipient, the interaction is already profusely discussed above. Moving on to the next kind of interaction which is one transmitter to many recipients, which is more likely similar to the previous case discussed as the recipients do not transmit any signals. The next interaction is many transmitters to one recipient, in which multiple transmitters that generate signal will collide at the recipients end disrupting doppler profiling with each other. The last case, which is many transmitters to many recipients, which is identical to the case of many transmitters to one recipient. Because recipients do not transfer any signals, the application "RAY" needs to only coordinate transmitters to avoid collisions. The application thereby uses carrier sense multiple access to find an appropriate channel for every transmitter to send the signal. Collisions among these signals can be detected using the root mean square of the FFT signal of that particular channel. A threshold root mean square value is chosen and if the RMS detected is higher than that threshold, only then it is concluded that a collision has occurred. And as a follow up to this collision, the transmitter switches the channel after a randomly set interval. The application then performs an FFT on the inaudible frequency band, the transmitter can now obtain the root mean square for all the channels with only 1-time measurement. As an attempt to reduce the probability of a collision, it uses a unique prioritization scheme for selecting the channel.

5.2. Design of the channel

The number of channels to be used in the application "RAY" will be determined by the entire frequency bandwidth and the bandwidth of every channel. According to under-sampling theorem [12], the frequency (FL, FH) which supports follows the following relationship [11]:

$$\frac{2.f_H}{n} \leq \frac{F_S}{n} \leq \frac{2.f_L}{n-1}, \forall n : 1 \leq n \leq \left\lfloor \frac{f_H}{f_H - f_L} \right\rfloor \quad (5)$$

In the above equation:

n is the under-sampling factor; Fs is the original sampling rate and the last parameter is the flooring operation.

The frequency band supported in our system when under sampling factors are used are given in the following table 5.2.1: This table shows the possible frequency bands supported when under-sampling factor n is used. The bands which are inaudible are underlined [20].

Table 5.2.1: Possible Frequency Bands Supported when under sampling factor n is used.

n	F_s^* (kHz)	Frequency band (kHz)	Δv (cm/s)
6	8.0	(8.0, 12.0), (16.0, 20.0)	3.6
7	6.9	(13.8, 17.25), <u>(20.7, 24.15)</u>	3.0
8	6.0	(12.0, 15.0), <u>(18.0, 21.0)</u>	2.7 *
9	5.3	(15.9, 18.6), <u>(21.2, 23.9)</u>	2.5

To optimize the setting of velocity resolution and temporal fidelity (the degree to which a channel stream maintains the rhythmic integrity), the application takes the under-sampling factor to be 8.

5.3. Prioritization of every channel

When the application users exist all together the limited number of channels will cause many and frequent collisions. But as explained above, the transmitters switch to a more "free "channel. As the doppler profiling needs to measure the relative velocities after a few seconds before the user stops walking, enough data will be obtained even after switching the channel. Such random switches also cause a high probability of collisions. The application minimizes this probability using a prioritization scheme. It involves classifying the categories into three broad headings: Unconfined channels – It is free if the FFT RMS of the channel found is lower than a decided threshold. This has the highest priority. Channels with a positive sector – In doppler profiling, a positive sector and a negative sector makes up the entire channel. Considering an example of three users A, B, and C, where A is the receiver and B and C are the transmitters. Let's say, A and B are transmitting through the same channel, and for the receiver A, the frequency offset for an approaching transmitter B is positive and similarly, for a departing transmitter, C is negative. In the application, since our only concern is how to deal with users walking towards us, we do not consider the negative offsets. The noises or any disturbance on the negative sector will directly be ignored, therefore considering the mechanism followed in "ray", the transmitter can any time switch his channel as long as a positive frequency sector has no intrusions. In the case where, multiple transmitters are moving towards a single receiver and if all of them use the same channel, it is still feasible to do doppler profiling. The attenuation of signals is around 6.6dB/m. [19]. As the "ray" adopts FFT based detection to discover the frequency offset with maximum amplitude, the closest sender may effortlessly "overwhelm" the signals of distant senders. Hence, the transmitter can switch to a busy channel as long as the FFT RMS over the positive sector of that channel is very low than the threshold.

5.4. Channel Coordination

As described in the channel prioritization, a collided transmitter switches to a channel of high numbered priority. The transmitter waits for a random time interval before it performs the switch operation. This is done to avoid repetitive collisions caused by frequent switching of many transmitters. If no channel is available at that instant of time the transmitter cancels the switching process which is pending and instead re-executed the algorithm as soon as possible.

6.Evaluation

Some false positive and false negative results always exist in the application "ray" and therefore RAY is not one hundred percent accurate. The user-friendly nature of the application and adaptiveness to the channel circumstances makes the application more favourable and likable to the users. Different results of false positives and false negatives can be achieved and adjusted using a trade-off in three crucial parameters which are: threshold of Doppler Profiling, alternativeness fraction, and duty ratio of Voice Profiling. In the real-world we study; it is assumed that the application is extendable to many other users and events. It can support many immobile mobile phones (as only phones of people who move need to send signals), also additionally acoustic channel can be reused every 11m away (from base paper tests performed on Nexus Series and Huawei P7 and we have made use of Samsung S10, Pixel 3 and Samsung S9 in our practice). This Doppler Profiling method is not dependent on the positions of the phone. This is another added advantage that reduces the pressure on the user to keep his/her phone in only a particular orientation. This happens because the application uses the relative trajectories of users moving around.

The test is performed considering 3 major components i.e.; Doppler profiling, voice profiling, and multiple access. A power profiling is also considered into account as a part of evaluation analysis.

6.1. Doppler profiling

The experiment Doppler profiling evaluates the following approaches:

Direct Approach: The sender and receiver are 10m apart and the sender walks towards the receiver.

Round About Approach: The sender and receiver are 10m apart but are separated by an obstacle. The sender walks towards the receiver crossing the obstacle.

Slant Approach: The sender and receiver are 10m apart but the sender walks at θ angle from the sender which varies from 30 to 45 degrees in our trials.

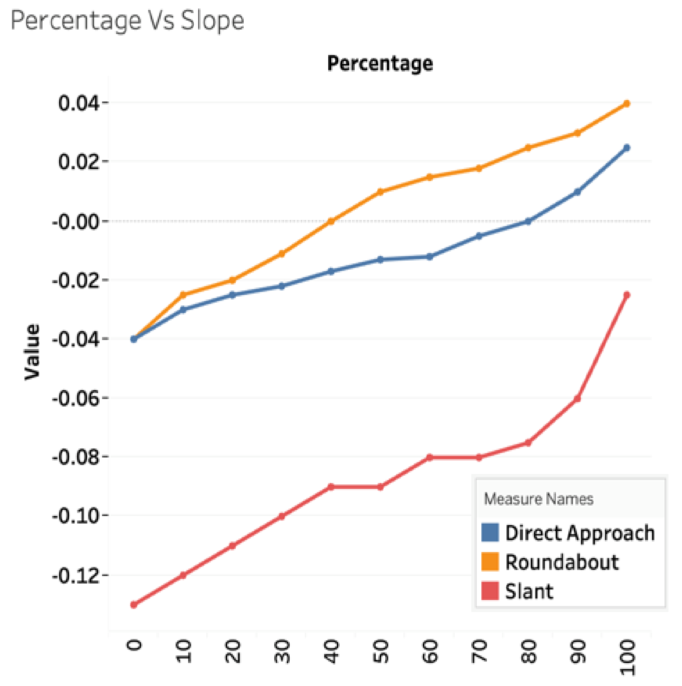


Figure 6.1 Cumulative Distribution function of the slopes for doppler profiling traces of 3 different trajectories (Actual)

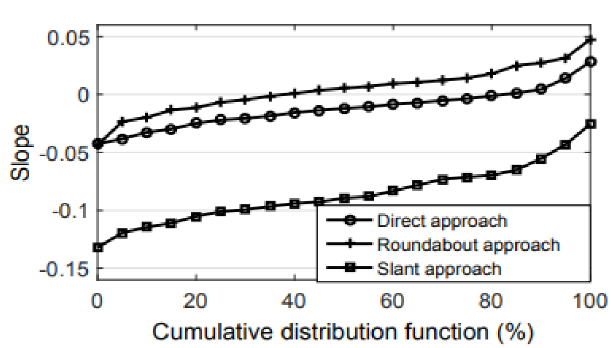


Figure 6.1.1 (Expected)

The classification of any approaching trajectories is based on the slope of the doppler profiling obtained. Figure 6.1 depicts cumulative distribution of the slopes we have obtained in the experiment that we have done. From the graph it's very easy to determine that the slant approach does not result into effective trajectories while the others 2 result in effective ones. Figure 6.1 is the actual results obtained and figure 6.1.1 is the expected one taken from base paper [20].

By looking at both the graphs its noted that the result is the same but the slope obtained in actual is somewhat different. As expected, graph is performed on large scale and also under controlled experiment environments so there might be some changes when done practically without imposing any restrictions in the environment which is in our case that is the actual graph.

Table 6.1: Slopes for slant approach trajectories with different angles and comparison with direct approach (Actual)

θ	0°	30°	40°	50°	60°
Mean	- 0.0130	-0.0581	-0.0810	-0.1020	-0.1050
Max	0.0276	-0.0271	-0.0379	-0.0721	-0.0725
Min	- 0.0432	-0.0911	-0.1101	-0.1220	-0.1290

Table 6.1.1: Slopes for slant approach trajectories with different angles and comparison with direct approach (Expected)

θ	0°	30°	40°	50°	60°
Mean	−0.0128	−0.0613	−0.0803	−0.1037	−0.1010
Max	0.0282	−0.0257	−0.0382	−0.0727	−0.0736
Min	-0.0428	−0.0967	−0.1110	−0.1276	−0.1322

Also, in table 6.1 we can see the breakdown of slopes which we have obtained in slant trajectories with different values of θ. For simplicity purpose we have just considered values from 30 to 60. Later after obtaining the values we compared those values with slopes obtained for direct trajectories for value of θ = 0. We have observed that as we go on increasing the values the slope goes on decreasing.

The results obtained in actual that is table 6.1 and compared it with the expected table 6.1.1 which we took from base paper [20]. We have observed that there is good amount of difference in the mean of the direct approach in our case its -0.0130 with the mean for slant approaches which goes on decreasing that is for 30 its -0.0581 and so on which can be seen the table 6.1. So, due to such huge difference it allows us to enable us easily distinguish these 2 approaches. Our

actual and expected is approximately the same which gives results of decrease in slopes for the slant approaches in both the cases.

Once Doppler profiling is complete, there could be a little misclassification rate, but with the next step of Voice Profiling, this misclassification can be mitigated.

6.2. Voice Profiling

Here the voice of the sender is verified and we examine if there is a one to one communication between the sender and receiver to finalize if they are engaged in a conversation. Mostly in conversations, there tends to be a pattern observed where one person talks, and then the next person would give a response to it. As the paper recommends, we performed the test with three volunteers, where they are less than a meter apart from each other. A and B volunteer would be having a one to one conversation while C would be the interferer. All the volunteers are using RAY; hence their timestamps and durations are monitored. The owner's voice is verified by RAY and the timestamp of other voices is matched with it to identify if it is a conversation. This is mathematically formulated using the alternativeness ratio and duty ratio. This experiment is performed in 3 different environments: hall, corridor, and street, where the corridor being the quietest and street being the loudest of the environments.

Accuracies of Voice profiling is defined in Table 6.2 which is actual one and table 6.2.1 is the expected results [20]. In RAY, the Alternativeness ratio confirms if a conversation pertains between A and B. The possible test cases to consider are false negative (there really exists a conversation between A and B but the application does not recognize it as a conversation) and false positive (a conversation between A and B does not exist but the application recognize it as a conversation).

We have considered 3 cases which are hall, corridor, street for evaluation. From that we observed that base paper had false positive of 15.83% with alternative ratio and when both is considered that is alternative and duty ratio, we get 6.50%. So, the average reduces by 9.3% [20]. Now in our case false positive with alternative ratio we got 15.90% and when both ratios are combined (alternative and duty ratio), we attained 6.5% then average reduces by 9.4%.

Considering False negative in the base paper it's 7.16% with alternative and 8.06 for alternative and duty ratio. Average increases by 0.90% [20]. In our case we got 7.23% in alternative and 8.30% with both thereby increasing average by 1.07%.

Table 6.2: Performance of one to one Conversation with and without duty ratio (Actual)

Location		Hall	Corridor	Street
Alternative	False Negative	9	6.9	5.8
	False Positive	15.1	13.9	18.7
Alternative + Duty Ratio	False Negative	9.5	8.5	6.5
	False Positive	5.4	4.8	9.3

Table 6.2.1: Performance of one to one Conversation with and without duty ratio (Expected)

Location		hall	corridor	street	mall
alternative	false negative	8.8%	7.0%	5.7%	9.2%
	false positive	15.4%	13.8%	18.3%	14.6%
alternative + duty ratio	false negative	9.4%	8.1%	6.7%	10.7%
	false positive	5.2%	4.7%	9.6%	8.7%

By taking duty ratio into consideration, the false positive rate is decreased and false negative rate is increased, which overall increases the efficiency of RAY. The results observed above are close to the expected result (which is the result in the base paper [20]). While the base paper has also taken into consideration, locations like shopping malls, which is a difficult test case to process as it is the loudest environment and not feasible due to time constraints.

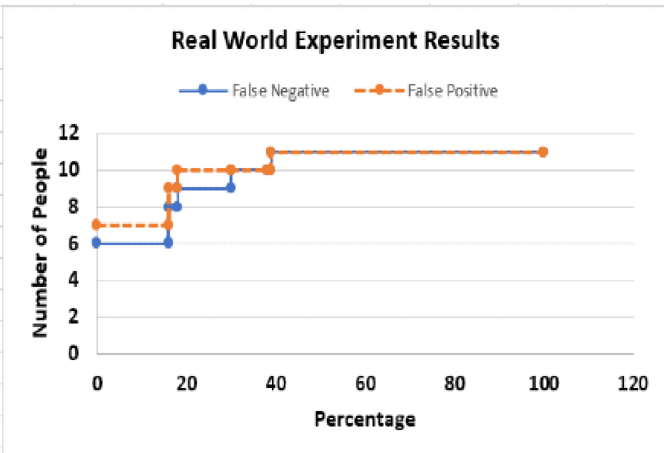
6.3. Multiple Access

The RAY's priority-based switching scheme is compared with a random switch and back off scheme. The experiment was conducted to test if these two can avoid collisions and perform Doppler profiling. The smartphones were added to the site and were configured to duplicate RAY senders. The smartphones transmitted acoustic signals on various channels. The success rate of channels is noted for a various number of RAY senders. The experiment results show that one sender had a success rate of 96% and when the number of senders increases, the success rate drops. The success rate is 75% and 55% for two and three senders per channel respectively. Because the RAY selects the best channels, the success rate is still 92%. Overall, because the RAY acknowledges low interference than random challenge switch, it is proved that it is better than the random challenge switch.

6.4. End to End Performance Results

The base paper carried out experiment on 11 people with 2 females and 9 males [20]. The participants perform the experiment as stated and recorded the list of people they have interacted with. So, for getting accuracies better or close to the base paper we also carried out the experiment with 11 people but we had taken 4 females and 7 males as our participants for evaluation.

The graph 6.3.1 shows the expected and the actual one where the first is Actual followed by Expected below it. Graph shows the number of people vs false rate which we have used it for evaluation of RAY.



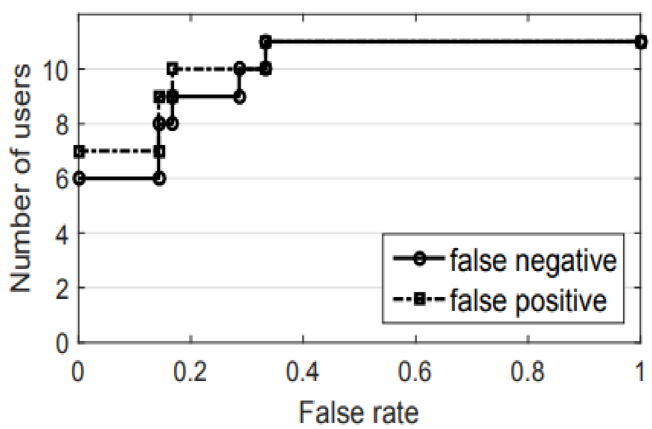


Figure 6.3.1 End to end performance of RAY (Actual and Expected)

We obtained similar results as that of base paper [20] that is from 6 out of 11 people had no false positive and false negative. From the remaining people we observed 3 people had false negative of <16 to 18% where as in base paper it just showed us that 3 people got <17%. For false positive base paper had <15% and we have got < 18% so there is some difference in accuracies as we haven't taken into consideration the scenarios where the experiment is performed in malls. But according to us we have attained better accuracies which is 9.46% for false negative where base paper achieved 9.7%. Lastly for false positive we achieved accuracy of 6.81% where on the other case base paper had 6.9%.

6.5. Power profile

Irrespective of the smartphone, the application drains out almost the same amount of power for all smartphones. The major power sources for the application are for screen time, sound recording, and other aspects of the application. Hence, the general power differences that can occur when compared with other smartphones would be the display's power consumption. The application records only if a user clicks on record, hence the screen should be turned on throughout the run time. Furthermore, RAY runs throughout the time it is turned on, even if there are no senders around. An ideal update that could potentially mitigate the problem would be making RAY idle when there are no senders present. Overall, RAY reduced 4% of battery (136mA) on the usage of approximately 30 minutes

7.Evaluation Analysis

Some false positive and false negative results always exist in the application "ray" and therefore it is not a hundred percent accurate. The user-friendly nature of the application and adaptiveness to the channel circumstances makes the application more favorable and likable to the users. Different results of false positives and false negatives can be achieved and adjusted using a trade-off in three crucial parameters which are: threshold of Doppler Profiling, alternativeness fraction, and duty ratio of Voice Profiling. In the real-world we study only with persons; it is assumed that the application is extendable to many other users and events. It can support many immobile mobile phones (as only phones of people who move need to send signals), also additionally acoustic channel can be reused every 11m away (from base paper tests performed on Nexus Series and Huawei P7 and we have made use of Samsung S10, Pixel 3 and Samsung S9 in our practice). This Doppler Profiling method is not dependent on the positions of the phone. This is another added advantage that reduces the pressure on the user to keep his/her phone in only a particular orientation. This happens because the application uses the relative trajectories of users moving around.

8.Conclusions

From the study carried out we can say that RAY is the first smartphone that is built which allows the people to identify each when they met at a particular event. RAY incorporates Doppler effects to the signals emitted by the sender /receiver to identify effective trajectories when 2 or more people approach each other. Moreover, our application also allows the user to record the voice of the people when they are interacting and having a conversation so that they can identify the people they met when the owners listens the recording and it also allows the user to know the location where they met by clicking on the get location button which is incorporated in the application. We have obtained close enough results and compared it with the ones given in the base as well for better understanding. The experiment involved many scenarios which are to be considered while evaluating the results and calculate the accuracies accordingly in different environment which is shown in results obtained Evaluation section. In all RAY we can say that it's not 100% accurate but has improvised the traditional approach of people interacting and has tried to make life of people easier by reducing efforts.

9.Acknowledgement

This work is supported by Singapore MOE AcRF Tier 1 MOE2013-T1-002-005 and NTU NAP grant M4080738.020 [20].

10.References

[1]. Md.Tanvir Aumi, Sidhant Gupta, Mayank Goel, et.al., "DopLink: Using the Doppler Effect for Multi-Device Interaction", UbiComp 13: Proceeding of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ISBN: 9781450317702, https://doi.org/10.1145/2493432.2493515.

[2]. Sidhant Gupta, Daniel Morris Desney tan, et.al., Sound Wave: Using the Doppler Effect to Sense Gestures," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems", ISBN: 9781450310154, https://doi.org/10.1145/2207676.2208331".

[3]. Sangki Yun, Yi-Chao-Chen, Lili Qiu," Turning a Mobile Device into a Mouse in the Air", Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services, ISBN: 9781450334945, https://doi.org/10.1145/2742647.2742662.

[4]. W. Huang, Yan Xiong, et al., "Shake and walk: Acoustic direction finding and fine-grained indoor localization using smartphones," IEEE INFOCOM 2014 - IEEE Conference on Computer Communications, Toronto,ON,2014,ISBN:978-1-4799-3360-0, https://ieeexplore.ieee.org/document/6847959.

[5]. D.Reynolds, Thomas F.Quatieri, et.al., "Speaker Verification Using Adapted Gaussian Mixture Models", Digital Signal Process, ISBN: 1051-2004, https://doi.org/10.1006/dspr.1999.0361.

[6]. Hong Lu, A.J. Brush, Amy Karlson, et.al., "SpeakerSense: Energy Efficient Unobtrusive Speaker Identification on Mobile Phones", Proceedings of the 9th International Conference on Pervasive Computing,ISBN:9783642217258, https://dl.acm.org/doi/10.5555/2021975.2021992.

[7]. Luo, Chengwen and Chan, Mun Choon," SocialWeaver: Collaborative Inference of Human Conversation Networks Using Smartphones", Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems, ISBN: 9781450320276, https://doi.org/10.1145/2517351.2517353.

[8]. Santagati, G. Enrico and Melodia, Tommaso, "U-Wear: Software-Defined Ultrasonic Networking for Wearable Devices", Proceedings of the 13th Annual International Conference on Mobile Systems, Application and Services,ISBN:9781450334945, https://doi.org/10.1145/2742647.2742655.

[9]. Nandakumar, Rajalakshmi and Chintalapudi, et.al., "Dhwani: Secure Peer-to-Peer Acoustic NFC", SIGCOMM Computer Communication Review, ISBN:0146-4833, https://doi.org/10.1145/2534169.2486037.

[10]. J.Rosen and L. Q. Gothard, "Encyclopaedia of Physical Science, Facts on file 2009.

[11]. Gomez C, Oller J, Paradells J. "Overview and Evaluation of Bluetooth Low Energy: An Emerging Low-Power Wireless Technology". Sensors. 2012; 12(9):11734-11753.

[12]. Sun Zheng and Bose, Raja and Zhang, et.al., "Spartacus: Spatially-Aware Interaction for Mobile Devices Through Energy-Efficient Audio Sensing", GetMobile: Mobile Computing and Communication's:2375-0529, https://doi.org/10.1145/2721914.2721918.

[13]. Nirupam Roy, He Wang, Romit Choudhary, "I am a smartphone and I can tell my User's Walking Direction", Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services,ISBN:9781450327930, https://doi.org/10.1145/2594368.2594392.

[14]. Yang M., Zheng, McClean, Harris, et.al., "Assessing the utility of Smart Mobile Phones in gait pattern analysis. Heath and Technology", 2, 81-88(2012), https://doi.org/10.1007/s12553-012-0021-8

[15]. Edward T.Hall, "The Hidden Dimension", New York: Anchor Books, 1990, ISBN: 9780385084765.

[16]. Lu, Hong, and Brush, A. J. Bernheim and Priyantha, et.al., SpeakerSense: Energy Efficient Unobtrusive Speaker Identification on Mobile Phones, Proceedings of the 9th International Conference on Pervasive Computing, ISBN: 9783642217258, https://dl.acm.org/doi/10.5555/2021975.2021992.

[17]. Chakraborty, G., Naik, K., Chakraborty, D., et al. Analysis of the Bluetooth device discovery protocol. Wireless Networks 16, 421–436 (2010), https://doi.org/10.1007/s11276-008-0142-1.

[18]. D. Camps-Mur, A. Garcia-Saavedra, and P. Serrano, "Device-to-device communications with Wi-Fi Direct: overview and experimentation," in IEEE Wireless Communications, vol. 20, no. 3, June 2013, ISSN: 1558-0687, https://ieeexplore.ieee.org/document/6549288.

[19]. L. E. Kinsler, Austin R. Frey, James V. Sanders, et.al., "Fundamentals of Acoustics", ISBN: 978-0-471-84789-2.

[20]. Zhang, Huanle and Du, Wan and Zhou, Pengfei and Li, et.al., “DopEnc: Acoustic-Based Encounter Profiling Using Smartphones”, Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking, ISBN: 9781450342261, https://doi.org/10.1145/2973750.2973775.