Accuracy of Bluetooth-Ultrasound Contact Tracing: Experimental Results from NOVID iOS Version 2.1 Using Five-Year-Old Phones

Po-Shen Loh*

June 26, 2020

Abstract

A common standard for close contact in the context of COVID-19 is 6 feet in the United States (2 meters elsewhere in the world). Many smartphone contact-tracing apps attempt to measure distance using Bluetooth received signal strength, but it remains an open question to reliably translate signal strength into an accurate distance metric. The NOVID app publicly launched for Android download with an approach combining Bluetooth and ultrasound on April 7, 2020, using sound time-of-flight calculations to reduce the number of false positive detections in app-based contact tracing. In this experiment using publicly-accessible features of the app, we systematically test the latest version of NOVID (iOS version 2.1) in a variety of realistic and adversarial at-rest settings. We find that a simple 9-foot classification threshold based on the NOVID distance measurement is very effective. Even though this experiment put the app in challenging circumstances (not an anechoic chamber), 99.6% of the 225 tested 12-foot-or-higher interactions were correctly classified as over-9-feet, while over 50% of the 187 tested under-6-foot interactions (and 94% of a representative subcategory) were correctly classified as 9-feet-or-under. This demonstrates that ultrasound can significantly contribute to app-based contact tracing.

1 Introduction

The COVID-19 pandemic has wrought havoc on ordinary life around the world. A distinctive characteristic that makes this novel disease difficult to control is its high level of contagion, coupled with a tendency for pre-symptomatic spread. Indeed, a recent paper by He, Lau, and Wu, *et al.* [2] in *Nature Medicine* estimated from their dataset that 44% of the infectiousness profile occurred presymptomatically. This naturally inspired much work on accelerating contact tracing with current technology.

In Spring 2020, a large number of smartphone contact tracing app projects emerged, spearheaded by the TraceTogether app developed by the national government of Singapore. Perhaps inspired by the TraceTogether app, which used a novel application of Bluetooth for anonymous short-range smartphone-to-smartphone communication, many projects based their core technology on Bluetooth. The core technological problem at the heart of a contact tracing app is the accurate

*Department of Mathematical Sciences, Carnegie Mellon University; and Expii, Inc. Email: po@novid.org.

detection of distance with the ability to discern at the meter-scale. Much of the world settled on a 2-meter standard (6 feet in the United States[1]).

On April 7, 2020, the NOVID app launched for public download on Android devices [3], introducing the combination of Bluetooth and ultrasound to more effectively prevent false positive classifications of faraway interactions as close-proximity. The principle was to accurately measure the time-of-flight of sound, which is theoretically possible because modern smartphones can reliably quantify time with millisecond-accuracy; when multiplied by the speed of sound, this produces sub-meter distance accuracy. To enable third-party verification of accuracy, NOVID then exposed the estimated distance proximity on its deployed app user interface for public testing, together with a button that triggers a distance scan on demand.

The experiment in this report uses this publicly accessible feature to systematically test the latest iOS version of NOVID (2.1) in a variety of real-world adversarial at-rest settings, to quantify its strengths and limitations in practical situations. The app does have Doppler correction built into its signal processing code, but this initial public report focuses on the at-rest setting because a very significant proportion of sustained human-to-human close-proximity interaction occurs when their smartphones are at rest. In addition to testing the ultrasonic distance measurement system in simple line-of-sight settings, this experiment includes a variety of adversarial settings designed to interfere with the physical limitations of any ultrasound-based approach: ultrasound has difficulty penetrating obstacles, and its reflections off objects create multipath effects that complicate time-of-flight signal processing calculations. However, the theoretical physical characteristics of ultrasound work in favor of reducing false positives, because any time-of-flight measurement of a reflected path is an overestimate of the truth. The objective of this experiment is to understand the degree to which this system can categorize interactions as under-6-feet settings with minimal false-positive rate, while still positively detecting a substantial fraction of true under-6-feet interactions.

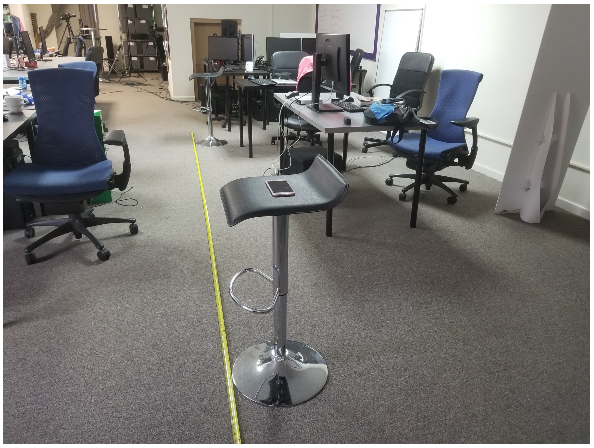
2 Materials and Methods

The latest iOS build of NOVID was used for this experiment, corresponding to the Version 2.1 branch. The distance measurement interface within the app, depicted in Figure 1(v), can be used by anyone to reproduce the results of this experiment by downloading any Version 2.1.x iOS build from the App Store.

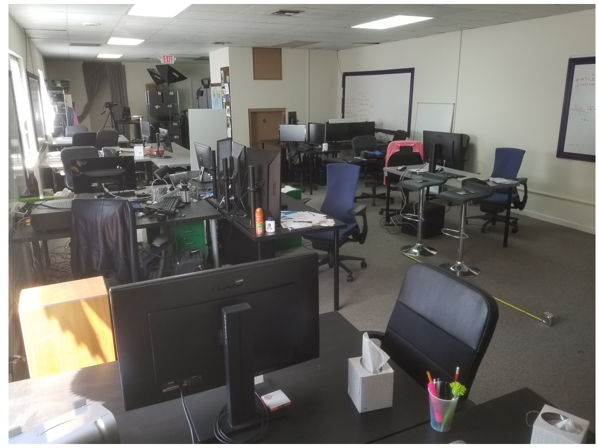
Figure 1 show a typical open-plan office environment, which was selected as the experimental setting, with standard walls, floors, and 8-foot-6-inch drop ceiling. Imperial units are used throughout this report because the USA version of the NOVID app was used for testing, and it displayed Imperial units on its user interface. Office desks, chairs, monitors, computers, cups, etc. were distributed throughout the office without obstructing the direct pathway of the experiment, as in Figure 1(ii) where the clear walkway in Figure 1(i) was used for testing the app at a range of distances. An anechoic chamber was intentionally not used so as to simulate a real-world environment with many sources of multipath interference.

The devices used were an Apple iPhone 6s and an iPhone 6s Plus, so as to test the performance on five-year-old smartphone technology. This is relevant because the app's ultrasonic communication frequency range (18.5–19.5 KHz) approaches the limit of smartphone speaker and microphone

[1]As of June 25, 2020, the relevant COVID-19 webpage maintained by the USA Centers for Disease Control [1] `https://www.cdc.gov/coronavirus/2019-ncov/php/public-health-recommendations.html` defined close contact as "$<$ 6 feet," with a footnote indicating that data to inform the definition of close contact are limited.



(i)

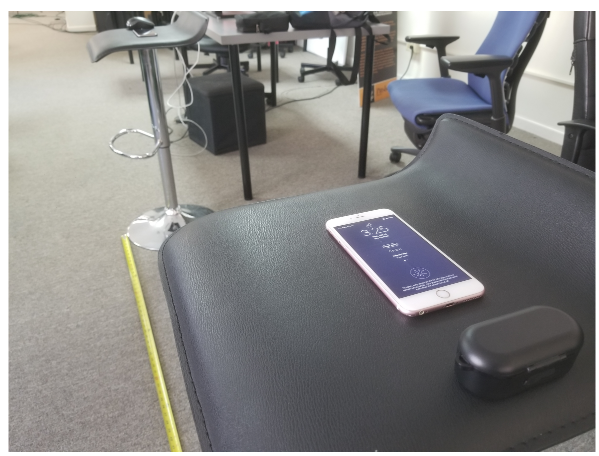


(ii)

(iii)

(iv)

(v)



(vi)

Figure 1: *Experimental setup*

capability. Throughout the experiment, the phones had at least 30% charge, so that the operating system would not engage power saving features.

The phones were placed flat on bar stools at height 30.5 inches off the ground, with each phone centered on the axis of each bar stool's rotating support pole as confirmed by visual checks from two perpendicular perspectives. The bar stools had circular bases, which facilitated center-to-center measurements through the use of a 25-foot Stanley metal measuring tape that was tangent to the perimeter of each circular base. This setup permitted the positioning of phones such that actual distance measurements were within one inch of accuracy.

The experimental procedure ran through many different setups. In each setup, three numerical measurements were sought. The nature of this version of the NOVID app is such that after each distance measurement, one of three possible results is displayed: an actual distance, or "Heard A Device But Could Not Measure The Distance" (recorded in our data as "Heard"), or "No nearby users" (recorded in our data as "Unheard"). The measurements within each setup stopped as soon as a total of three distances were received, or as soon as three non-distance results were received in a row. However, all non-distance results were recorded, so this effectively only increased the number of samples in some setups.

Throughout, the human experimenter(s) remained at least 3 feet away from the phones while they were measuring distances. In situations when one phone was placed in a mens' shorts pocket, the person stood motionless. Po-Shen Loh was the primary experimenter supervising all data collection, assisted by Vivian Loh for the portion of data collection at distances of 13 feet and higher, and for all settings where he placed one phone in his shorts pocket.

Some settings involved positioning small obstacles 2 inches away from the smartphone speakers, as in Figure 1(vi). The small obstacles used were a Logitech M185 wireless mouse and the charging box of a TOZO T10 Bluetooth earbud set. Both of these were slightly irregular shapes, with approximate dimensions $4 \times 2 \times 1$ inches. Some settings involved bags, both of which were made of durable nylon luggage-type material and depicted in Figure 1(iii). The large bag was not padded, and measured $19 \times 14 \times 5$ inches when loosely crumpled with an iPhone inside. The small bag was a padded camera bag which held its form due to the internal 1/4-inch foam-like padding, with dimensions $8 \times 8 \times 2$ inches. Some settings involved a 59-inch-tall, 20-inch-wide self-standing 1/4-inch foam core board placed at the midpoint between the two phones, obstructing their line-of-sight, as in Figure 1(iv). Some settings involved a pair of mens' shorts, made of a light khaki-type material and worn by an approximately-120-pound man. Therein, the phone was placed in different pockets, where its center was 27 inches above ground when in the front pocket, and 29 inches above ground when in the back pocket.

Background music (1990's pop hits via Spotify high-quality streaming for maximum frequency range) was continuously played through a pair of floorstanding Wharfedale Emerald 97 Mark IV speakers and a 500-watt Infinity Interlude IL120s subwoofer. The equipment was chosen so as to deliver a frequency response that would reproduce the near-ultrasonic frequencies in music intermixed with frequent spikes in the amplitude due to pop-music rhythmic beats, so as to interfere with the app operation. The floorstanding speaker high-frequency tweeters 35 inches off the ground and about 10–12 feet away generally facing toward the phones. The volume was set such that the ambient sound level at the phone position (as measured by the "Sound Meter" Android app [4] on a Samsung Galaxy S8) was in the 60–70 dB range. Conversation between the experimenters required mild shouting to achieve audible clarity.

3 Procedure

For each number of feet in $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25\}$, as measured from the center of one phone to the other, perform the distance measurement protocol in the previous section (stop after a total of 3 measurements, or after 3 consecutive non-distance measurements) in each of the following settings.

Line of sight. Place the phones on the bar stools so that they have clear line of sight, and then test the following four settings:

- **A1.** Orient the phones in the same direction, where that direction is perpendicular to the line between the phones. The external speaker is on the bottom edge of the phone, and so this orients the speakers in parallel directions which are perpendicular to the line between the phones.
- **A2.** Orient the phones so that their speakers point directly towards each other.
- **A3.** Orient the phones so that their speakers point directly away from each other, and there are no other objects within 5 feet directly in front of each phone's speaker.
- **A4.** Orient the phones so that their speakers point directly away from each other, but place small obstacles 2 inches in front of each speaker: place the wireless mouse in front of the iPhone 6s and the Bluetooth earbud charging box in front of the iPhone 6s Plus.

In bags. Orient the phones so that their speakers point in the same direction perpendicular to the line between the phones. Place the phones on the bar stools so that they have clear line of sight, except as detailed in the following two settings:

- **B1.** Leave the iPhone 6s exposed, while enclosing the iPhone 6s Plus in the large non-padded bag.
- **B2.** Enclose the iPhone 6s Plus in the large non-padded bag, and enclose the iPhone 6s in the small padded bag.

Blocked by foam core board. Orient the phones so that their speakers point in the same direction perpendicular to the line between the phones. Place the phones on the bar stools so that they have clear line of sight, except for the foam core board blocking at their midpoint and as specifically detailed in the following three settings:

- **C1.** Leave both phones exposed.
- **C2.** Leave the iPhone 6s exposed, while enclosing the iPhone 6s Plus in the large non-padded bag.
- **C3.** Leave the iPhone 6s Plus exposed, while enclosing the iPhone 6s in the small padded bag.

In shorts pocket. Place the iPhone 6s Plus flat on a bar stool, exposed and oriented such that the speaker is pointing in a direction perpendicular to the line between the phones. Place the iPhone 6s vertically in a mens' shorts pocket, with the speaker pointing upwards. The phones have a clear line of sight apart from the person wearing the shorts. That person stands next to the measuring tape, facing towards the other phone, such that their center of mass is at the specified distance from the other phone. Measure the distance in these two settings:

D1. Place the iPhone 6s in the person's front left pocket.

D2. Place the iPhone 6s in the person's back right pocket.

Note that since there are 16 distances, and 11 settings per distance, this is a total of 176 setups. After performing all measurements, perform a sanity check to confirm that in situations when ultrasound should not be detected, the app correctly reports "Unheard." In our experiment, this was accomplished by going into a bathroom 29 feet away from the iPhone 6s Plus, closing the wooden door, going to the back of the bathroom (6 feet beyond the door), and facing the door with the other phone (iPhone 6s) in a back pocket. The app was then invoked three times.

4 Results and Analysis

The following pages contain all 589 of the raw data points measured over the course of experimental measurements on June 24–25, 2020. All numbers are expressed in units of inches. All cells less than or equal to 72 inches (6 feet) are highlighted in green, and all other cells less than or equal to 108 inches (9 feet) are highlighted in yellow. The reason for this color coding is because *a priori*, a reasonable threshold for separating 6-feet-and-under interactions from 12-feet-and-over interactions would be to classify based on whether the NOVID distance measurement was less than or equal to 9 feet.

In this color-coding scheme, the green and yellow cells collectively cover all interactions at distance less than or equal to 9 feet. Consequently, it is good if most of the cells corresponding to actual distance at most 72 are green or yellow, and most of the cells corresponding to actual distance at least 144 are uncolored. The column "Un/heard in notes" totals the number of non-distance measurements recorded for each setting, if at least one distance measurement was obtained.

Actual distance	Setting	Measure 1	Measure 2	Measure 3	Un/heard in notes	Notes
24	A1	22	22	22		
24	A2	16	16	15		
24	A3	198	100	198		
24	A4	33	32	34		
24	B1	24	30	28		
24	B2	31	24	96		
24	C1	24	58	27	1	1 Heard between Measure 1 and 2
24	C2	116	155	243		
24	C3	165	53	46		
24	D1	32	22	20		
24	D2	201	276	102		
36	A1	35	35	35		
36	A2	26	26	26		
36	A3	293	238	194		
36	A4	45	45	44		
36	B1	53	49	54		
36	B2	Heard	Heard	Heard		
36	C1	36	36	35		
36	C2	153	164	166		
36	C3	159	245	149	1	1 Heard between Measure 1 and 2
36	D1	28	32	31		
36	D2	184	70	14	2	2 Heards between Measure 2 and 3
48	A1	47	46	46		
48	A2	41	40	40		
48	A3	167	380	212		
48	A4	57	58	55		
48	B1	63	84	66		
48	B2	Heard	Heard	Heard		
48	C1	78	105	97		
48	C2	166	165	155	4	1 Heard before Measure 1, then 1 Heard between Measures 1 and 2, then 2 Heard between Measures 2 and 3
48	C3	78	88	84	1	1 Heard before Measure 1
48	D1	44	44	44		
48	D2	137	83	96	2	2 Heards between Measures 1 and 2
60	A1	56	57	58		
60	A2	51	51	50		
60	A3	197	197	199		
60	A4	61	62	77		
60	B1	72	63	64		
60	B2	70	162	112	5	2 Heards before Measure 1, then 1 Heard between Measures 1 and 2, then 2 Heards between Measures 2 and 3
60	C1	114	99	103		
60	C2	127	93	133		
60	C3	240	129	145	2	1 Heard before Measure 1, then 1 Heard between Measures 1 and 2
60	D1	54	53	53	2	2 Heards before Measure 1
60	D2	234	71	83		
72	A1	70	71	72		
72	A2	75	78	77		
72	A3	342	299	189		
72	A4	89	94	89	1	1 Heard before Measure 1
72	B1	84	127	95	1	1 Heard between Measures 2 and 3
72	B2	Heard	Unheard	Heard		
72	C1	216	112	117		
72	C2	160	165	165		
72	C3	144	133	159		
72	D1	68	64	66		
72	D2	165	172	170		
84	A1	83	83	83		
84	A2	84	85	85		
84	A3	354	353	354		
84	A4	106	105	105		
84	B1	94	86	83		
84	B2	81	83	83		
84	C1	114	120	245		
84	C2	250	181	264	1	1 Unheard before Measure 1
84	C3	206	240	239		
84	D1	77	78	88		
84	D2	Heard	Heard	Heard		
96	A1	105	99	102		
96	A2	84	88	84		
96	A3	215	265	217		
96	A4	105	106	106		
96	B1	Heard	Heard	Heard		
96	B2	Heard	Heard	Unheard		
96	C1	255	131	252		
96	C2	192	186	258		
96	C3	171	255	254		
96	D1	91	87	91		
96	D2	Heard	Unheard	Heard		
108	A1	103	106	103		
108	A2	97	97	100		
108	A3	359	363	223		
108	A4	153	150	118		
108	B1	178	193	33	1	1 Heard before Measure 1

108	B2	Unheard	Unheard	Unheard		
108	C1	159	257	182		
108	C2	187	280	277	2	2 Heard before Measure 1
108	C3	258	257	256		
108	D1	97	100	103		
108	D2	Heard	Unheard	Heard		
120	A1	115	116	117		
120	A2	113	113	113		
120	A3	376	376	373		
120	A4	156	162	155		
120	B1	194	205	169		
120	B2	236			5	1 Heard and 1 Unheard before Measure 1, then 1 Heard and 2 Unheard
120	C1	196	262	140		
120	C2	192	199	195		
120	C3	187	185	184		
120	D1	108	107	153		
120	D2	393	310	295	2	1 Heard before Measure 1, then 1 Heard between Measures 2 and 3
132	A1	164	158	127		
132	A2	134	134	133		
132	A3	342	342	341		
132	A4	166	136	166		
132	B1	180	168	168	2	1 Heard before Measure 1, then 1 Heard between Measures 2 and 3
132	B2	Unheard	Heard	Unheard		
132	C1	208	167	186		
132	C2	221	283	200	1	1 Heard before Measure 1
132	C3	266	336	333		
132	D1	161	129	160		
132	D2	Heard	Heard	Heard		
144	A1	175	238	238		
144	A2	139	207	207		
144	A3	349	349	350		
144	A4	291	186	186		
144	B1	350	147	188		
144	B2	Heard	Unheard	Unheard		
144	C1	337	336	338		
144	C2	202	215	223		
144	C3	270	204	201	1	1 Heard before Measure 1
144	D1	149	150	224		
144	D2	Heard	Heard	Unheard		
156	A1	183	209	159		
156	A2	152	178	172		
156	A3	363	359	359		
156	A4	167	168	167		
156	B1	160	158	147	2	1 Heard before Measure 1, then 1 Heard between Measures 2 and 3
156	B2	Heard	Heard	Unheard		
156	C1	275	275	276		
156	C2	286			5	2 Heards before Measure 1, then 3 Heards
156	C3	74	281	272		
156	D1	148	153	149	1	1 Heard before Measure 1
156	D2	Heard	Heard	Heard		
168	A1	254	165	183		
168	A2	207	207	209		
168	A3	362	389		4	1 Heard before Measure 1, then 3 Heard after Measure 2
168	A4	204	205	209		
168	B1	Heard	Heard	Heard		
168	B2	Unheard	Unheard	Heard		
168	C1	346	346	346		
168	C2	346	354	274		
168	C3	277	353	194	1	1 Heard before Measure 1
168	D1	160	163	167		
168	D2	Heard	Heard	Heard		
180	A1	188	263	195		
180	A2	173	173	176		
180	A3	386	348	345	2	1 Heard before Measure 1, then 1 Heard between Measures 1 and 2
180	A4	207	212	208		
180	B1	Heard	Heard	Heard		
180	B2	Unheard	Heard	Unheard		
180	C1	285	292	291		
180	C2	359			3	After Measure 1, 2 Unheard and 1 Heard
180	C3	342	294	261	3	1 Heard before Measure 1, then 2 Heards between Measures 1 and 2
180	D1	268	205	202		
180	D2	Heard	Heard	Heard		
240	A1	267	275	271		
240	A2	256	259	289		
240	A3	Heard	Heard	Heard		
240	A4	270	272	270		
240	B1	244			4	1 Heard before Measure 1, then 3 Heards
240	B2	Heard	Unheard	Unheard		
240	C1	282	279	276		
240	C2	336	331		3	3 Heards after Measure 2
240	C3	317	322	314	1	1 Heard between Measures 1 and 2
240	D1	297	290	236		
240	D2	Heard	Heard	Heard		

300	A1	324	326	326		
300	A2	307	301	304		
300	A3	Heard	Heard	Heard		
300	A4	393	315	331	2	1 Heard between Measures 1 and 2, then 1 Heard between Measures 2 and 3
300	B1	Unheard	Heard	Heard		
300	B2	Unheard	Unheard	Unheard		
300	C1	370	329	330		
300	C2	Heard	Heard	Heard		
300	C3	Heard	Heard	Heard		
300	D1	317	364	345		
300	D2	Heard	Heard	Unheard		
Bathroom	Sanity	Unheard	Unheard	Unheard		

The sanity check in the last row confirms that when the phones have their ultrasound blocked by distance and a door, the "Unheard" response triggers reliably. Indeed, the earliest objective of NOVID's Bluetooth-ultrasound framework was to reject interactions from different rooms.

It turns out that NOVID's iOS Version 2.1 implementation of the combined Bluetooth-ultrasound framework achieves significantly more than just that. The laws of Physics govern the travel of ultrasonic waves, and the baseline accuracy of a particular ultrasonic implementation can be assessed in terms of its performance on Settings A1 and A2. (Even then, the task is not completely trivial in the real-world environment of this experiment, which is far from anechoic, and has much multipath interference.) Those settings do turn out to be extremely accurate in Table 1.

Setting	Mean of Error (inches)	Standard Deviation (inches)
A1	−1.0	3.1
A2	−6.3	5.6
A4	14.4	10.7
D1	−4.9	3.9

Table 1: *Accuracy of measurements ranging over true distances of 2, 3, 4, 5, 6, 7, 8, and 9 feet, in the settings where Physics would predict accurate results.*

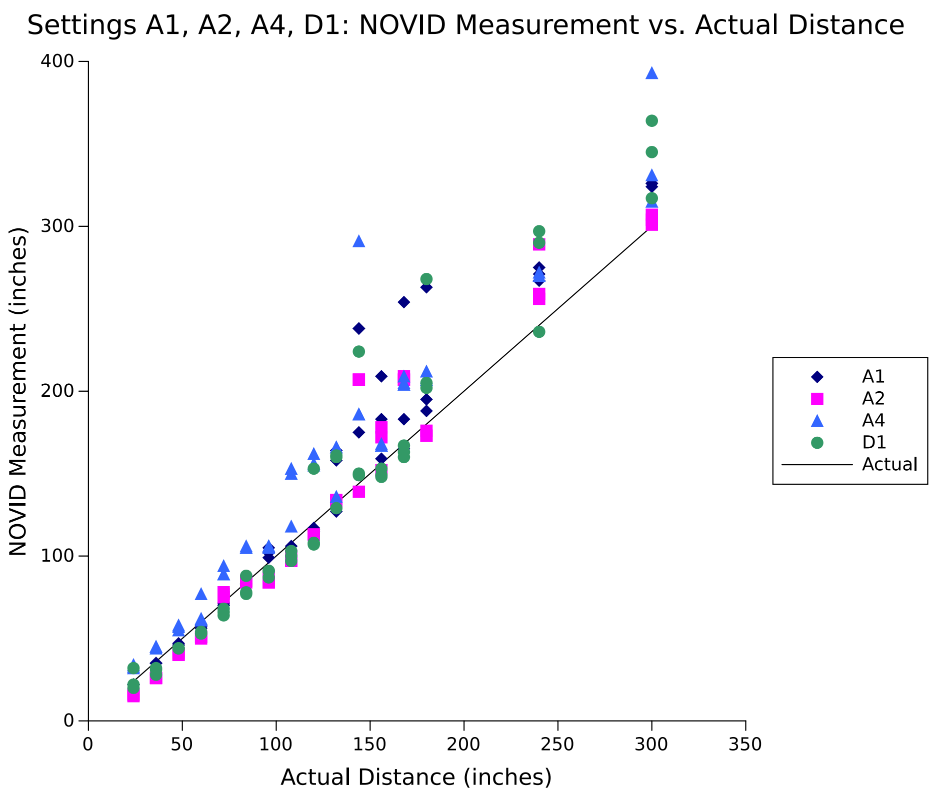


Figure 2: *All distance measurements for Settings A1, A2, A4, and D1. Note that the region of highest accuracy extends up to 108 inches (9 feet), while measurements consistently provide upper-bounds on the actual distance.*

On the basis of the very high accuracy in Settings A1 and A2, the next-most-ideal setting from a Physics perspective is A4, because each small obstacle scatters ultrasound back towards the other phone. (Physics would predict that Setting A3 gives wild overestimates, because there is nothing nearby for the ultrasound to reflect off of, and that is confirmed in the experiment.) The fact that NOVID is able to return to accuracy in Setting A4 with just a small obstacle indicates that the multipath disambiguation processing methods are robust.

Already, Settings A1, A2, and A4 indicate that when phones are placed on tables (e.g., in an office environment or at home), NOVID is sensitive enough to pick out one of the shortest multipath ultrasound trajectories and accurately estimate its length. Taking reflections into account, this motivates imposing a proximity detection threshold of 9 feet (108 inches) for the purpose of capturing interactions which in reality are within 6 feet. It is here that the (geometric) triangle inequality shines: the shortest path between two points is the straight line, and so ultrasound is particularly robust to avoiding false positives. Indeed, if one used the 9-foot-threshold to differentiate between 6-feet-and-under and 12-feet-and-over interactions, then there was only one false positive.

True Distance (feet)	Number of Samples	Number Correctly Categorized	Percentage
≤ 6	187	103	55.1%
≥ 12	225	224	**99.6%**

Table 2: *Accuracy of categorizing by thresholding the reported distance against 9 feet.*

The fact that a substantial fraction of truly-under-6-foot interactions were successfully categorized by this threshold, even in a test suite that included many adversarial settings, shows that this very high level of robustness to false positives is not simply achieved by turning off all detection. Rather, many useful positive settings are correctly handled.

The particular percentages should be taken with a grain of salt, because they are as proportions of test settings in this experiment. However, a real-world correction may work in ultrasound's favor. In real-world settings, the correctly characterized percentage of 6-feet-and-under interactions could be substantially higher than the 55.1% above, particularly if Setting A4 is representative of the disproportionate norm. Indeed, in Setting A4, 15 out of 16 measurements of truly-under-6-feet interactions were correctly detected as under-9-feet, corresponding to a success rate of 94%. Setting A4 is particularly interesting because it represents accurate detection of the shortest multipath reflection with phones that are out-of-pocket. Further discussion will be in the Conclusion.

We then turn to study NOVID's ability to detect in obstructed situations. Table 1 shows that Setting D1 still has very good performance. From a Physics perspective, since the phone is entirely enclosed in a front pocket, the ultrasound will emanate from the entire pocket, and the effective detection indicates that NOVID's algorithms are sufficiently sensitive to handle some attenuation. In fact, the extremely accurate detections in Setting C1 at close ranges (2 or 3 feet) indicate that the signal processing methods are able to discern substantially attenuated signals through a 1/4-inch foam core board, even in an environment which has many multipath reflections. Even when one phone is in a bag (Setting B1), detection is reasonable up to 7 feet (Table 3).

Setting	Mean of Error (inches)	Standard Deviation
B1	13.7	13.9

Table 3: *Accuracy of measurements corresponding to true distances of 2, 3, 4, 5, 6, and 7 feet.*

We now turn to the settings in which ultrasound was not effective. They are consistent with what one would anticipate based on the Physics of ultrasound. Setting D2 was the worst by far, and is likely because from a back pocket, there is no single-reflection transmission pathway off the ceiling, floor, or a side wall. The other setting which generally did not produce distances was B2 (both phones in bags), presumably because the signals were so attenuated that they were too diffuse to accurately calculate time-of-flight for. However, even then, among the B2 instances when the devices were within 6 feet, the ultrasonic ranging protocol only produced "Unheard" one time, out of a total of 20 trials, thereby detecting the presence of an ultrasound signal with 95% success. In light of the sanity check which confirmed that devices in different rooms reliably produce "Unheard," this indicates that if additional smartphone sensors are used in tandem (such as the proximity sensor to detect whether the phone is inside a bag), the "Heard" vs. "Unheard" signal could still potentially have value.

The remaining Settings C2 and C3 (one phone in a bag, but with a foam core board blocking line of sight) effectively turn out to be similar to Setting B1, because the ultrasound just reflects off of another object in the room. As expected, the measured lengths are then overestimates of the true length. However, it is interesting that the C2 measurements at each distance appear to be relatively self-consistent (as are the C3 measurements amongst themselves). Table 4 highlights this, showing that in many situations the standard deviation of the received measurements is small. Although the sample size is too small to draw significant conclusions, it would be interesting to further investigate whether NOVID's signal processing algorithm has locked onto a particular multipath reflection avoiding the obstacle.

Actual Distance (inches)	Measure 1	Measure 2	Measure 3	Standard Deviation
24	116	155	243	65
36	153	164	166	7
48	166	165	155	6
60	127	93	133	22
72	160	165	165	3

Table 4: *Self-consistency of measurements in obstructed Setting C2, suggesting accurate detection of multipath reflection length.*

5 Conclusion

This systematic experiment reveals the significant potential of ultrasound in controlling the false-positive rate for contact tracing. The details of the signal processing implementation are important, in order to handle real-world situations with multipath interference, such as the various challenging environments in this test suite. In particular, this experiment indicates that bags and barriers do attenuate ultrasonic signals.

This experiment also shows a way out. In the context of COVID-19 contact tracing applications, it is important to identify interactions that represent close physical proximity (within 6 feet). This is why no experiment was done with many scattered obstacles in the line of sight between two phones at distance 15 feet. Instead this experiment already indicates that NOVID's ultrasound method is generally effective at ruling out false positives even in obstructed situations, because it consistently upper-bounds the true distance when there is no direct line of sight due to the triangle inequality. At close physical ranges, the 94% categorization accuracy in Setting A4 suggests that if people set their phones on a table next to them, in the vicinity of other common items (mouse, earbud charger, wallet, cup, purse, etc.), then ultrasound will find a way to reflect outward, possibly scattering off other objects and the ceiling to reach the other phone. This experiment did not set up such a situation due to the complexity of defining it with sufficient specificity for reproducibility, and so that will be postponed for subsequent experimentation. An example of a relevant test setting would consist of devices 6 feet apart, on tables, with a variety of objects nearby, some of which serve as obstacles but some of which (such as computer monitors or people's chins) facilitate pathways of reflection. In the interim, anecdotal observations indicate that when phones are out-of-pocket at close distance in real settings, there are often enough nearby objects to facilitate a short reflected ultrasonic pathway.

Fortunately, it is already a common behavior pattern for some people to set their phones on a nearby surface when they intend to stay in the same place for some time (e.g., office and home environments, or cafes and restaurants). These situations correlate with extended interactions with significant opportunity to transmit viral load. This, combined with the experimental validation of some level of robustness to obstruction in pockets and bags, lead us to conclude that ultrasound has high practical potential for sensing proximity for contact-tracing purposes.

References

[1] Centers for Disease Control and Prevention. Public Health Guidance for Community-Related Exposure | CDC. Accessed on June 25, 2020. https://www.cdc.gov/coronavirus/2019-ncov/php/public-health-recommendations.html

[2] He, X., Lau, E.H.Y., Wu, P. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med* **26**, 672–675 (2020). https://doi.org/10.1038/s41591-020-0869-5

[3] NOVID Website. Snapshot from April 7, 2020. https://web.archive.org/web/20200407200920/https://www.novid.org/

[4] Splend Apps. Sound Meter Android App Version 1.67. Accessed on June 25, 2020. https://play.google.com/store/apps/details?id=com.splendapps.decibel