Abstract—We report on the results of a Covid-19 contact tracing measurement study carried out on a commuter tram in Dublin, Ireland. Our measurements indicate that in the tram there is little correlation between received signal strength and distance between handsets. We applied the detection rules used by the Italian, Swiss and German apps to our measurement data and also characterised the impact on performance of changes in the parameters used in these detection rules. We find that the Swiss and German detection rules trigger no exposure notifications on our data, while the Italian detection rule generates a true positive rate of 50% and a false positive rate of 50%. Our analysis indicates that the performance of such detection rules is similar to that of triggering notifications by randomly selecting from the participants in our experiments, regardless of proximity.

I. INTRODUCTION

We report on the results of a Covid-19 contact tracing measurement study carried out on a commuter tram in Dublin, Ireland. The tram is of a standard French design widely used in Europe. Measurements were collected between 108 pairs of handset locations and we have made those publicly available [1]. This adds to the body of data on contact tracing app performance on public transport initiated in [2].

Contact tracing apps based on the Google/Apple Exposure Notification (GAEN) API [3] are currently being rolled out across Europe, with apps already deployed in Italy, Switzerland and Germany. These apps use Bluetooth received signal strength to estimate proximity and will likely be used as an adjunct to existing manual contact tracing and test systems. Existing manual systems can usually readily identify the people with whom an infected person share accommodation and with work colleagues with whom the infected person is in regular contact. More difficult is to identify people travelling on public transport with whom an infected person has been in contact, since the identities of these people are usually not known to the infected person and are generally not otherwise recorded. Public transport is therefore potentially an important use case where effective contact tracing apps may be of significant assistance in infection control.

In summary, our measurements indicate that in the tram there is little correlation between received signal strength and distance between handsets. Similar ranges of signal strength are observed both between handsets which are less than 2m apart and handsets which are greater than 2m apart (including when handsets are up to 5m apart). This is likely due to reflections from the metal walls, floor and ceiling within the tram, metal being known to be a strong reflector of radio signals [4], [5], and is coherent with the behaviour observed on a commuter bus [2].

We applied the detection rules used by the Italian, Swiss and German apps to our measurement data and also characterised the impact on performance of changes in the parameters used in these detection rules. We find that the Swiss and German detection rules trigger no exposure notifications, despite around half of the pairs of handsets in our data being less than 2m apart. The Italian detection rule has a true positive rate (i.e. correct detections of handsets less than 2m apart) of around 50%. However, it also has a false positive rate of around 50% i.e. it incorrectly triggers exposure notifications for around 50% of the handsets which are greater than 2m apart. This performance is similar to that of triggering notifications by randomly selecting from the participants in our experiments, regardless of proximity.

We observe that changing the people holding a pair of handsets, with the location of the handsets otherwise remaining unchanged, can cause variations of ±10dB in the attenuation level reported by the GAEN API. This is pertinent because this level of “noise” is large enough to have a substantial impact on proximity detection.

II. METHODOLOGY

A. Experimental Protocol

Our experimental measurements were collected on a standard light-rail tram carriage used to carry commuters in Dublin, Ireland, see Figure 1(a). We recruited seven participants and gave each of them Google Pixel 2 handsets. We asked them to sit in the relative positions shown in Figure 1(b). This positioning aims to mimic passengers respecting the relaxed social distancing rules likely during easing of lockdown and with the distances between participants including a range of values < 2m and a range of values > 2m, see Figure 2. Each experiment is 15 minutes duration giving around 3 scans by the GAEN API when scans are made every 4 mins. A Wifi hotspot was set up on the tram and the participants were asked to hold the handset in their hand and use it for normal commuter activities such as browsing the internet.

After the first experiment was carried out participants were then asked to switch seats (they chose seats themselves) and a
second 15 minute experiment run. After the second experiment participants were again asked to change seats for the third 15 minute experiment and, in addition, two participants were asked to place their handsets in their left trouser pocket (in an orientation of their choice).

Each handset had the GAEN API and a modified version of the Google exemplar Exposure Notification app [6] installed, and was registered to a gmail user included on the Google GAEN whitelist so as to allow use of the GAEN API by the Exposure Notification app. Each handset also had a GAENAdvertiser app developed by the authors installed. This app implements the transmitter side of the GAEN API and allowed us to control the TEK used and also to start/stop the broadcasting of Bluetooth LE beacons.

At the start of each 15 minute experiment participants were asked to configure the GAENAdvertiser app with a new TEK and then to instruct the app to start broadcasting GAEN beacons. At the end of the experiment the GAENAdvertiser stopped broadcasting beacons. In this way a unique TEK is associated with each handset in each experiment, and these can be used to query GAEN API to obtain separate exposure information reports for each handset in each experiment.

Following all three experiments the handsets were collected, the TEKs used by each handset extracted and the GAEN API on each then queried for exposure information relating to the TEKs of the other handsets. In total, therefore, from these experiments we collected GAEN API reports on Bluetooth LE beacon transmissions between 108 pairs of handset locations. This measurement data is publicly available [1].

To provide baseline data on the radio propagation environment we also used the standard Android Bluetooth LE scanner API to collect measurements of RSSI as the distance was varied between two Google Pixel 2 handsets placed at a height of approximately 0.5m (about the same height as the tram seating) in the centre aisle of the tram carriage.

B. Ethical Approval

The experimental protocol was reviewed and approved by the Ethics Committee of the School of Computer Science and Statistics, Trinity College Dublin. The ethics application reference number is 20200503.

C. Hardware & Software Used

We used Google Pixel 2 handsets running GAEN API version 202512001\(^1\), which includes a major update by Google issued on 13th June 2020.

We used a version of the Google exemplar Exposure Notification app modified to allow us to query the GAEN API over USB using a python script (the source code for the modified app is available on github [6]).

In addition we also wrote our own GAENAdvertiser app that implements the Bluetooth LE transmitter side of the GAEN API [3]. GAENAdvertiser allows us to control the TEK, and in particular reset it to a new value at the start of each experiment. In effect, resetting the TEK makes the handset appear as a new device from the point of view of the GAEN API, and so this allows us to easily collect clean data (the GAEN API otherwise only resets the TEK on a handset once per day). We carried out extensive tests running GAENAdvertiser and the GAEN API on the same device to confirm that under a wide range of conditions the responses of the GAEN API on a second receiver handset were the same for beacons from GAENAdvertiser and the GAEN API, see [7] for further details.

GAENAdvertiser is open source and can be obtained by contacting the authors (we have not made it publicly available, however, since it can be used to facilitate a known replay attack against the GAEN API [8]).

D. Querying the GAEN API

The GAEN API takes low and high attenuation threshold values as input, together with the TEK and time interval of interest, and responds with three attenuation duration values, giving the duration (in minutes) that the attenuation level is below the low threshold, the duration the attenuation level is between the low and high thresholds and the duration above the high threshold. For each TEK and time interval we made repeated queries to the GAEN API holding the low threshold constant at 48dB and varying the high threshold from 49dB to 100dB (in 1dB steps up to 80dB, then in 5dB steps since noise tends to be higher at higher attenuation levels). By differencing this sequence of reports we can infer the attenuation duration at each individual attenuation level from 48dB through to 100dB.

We present the attenuation duration data obtained in this way using a coloured heatmap. We split the range of attenuation

---

\(^1\)As reported in the Settings-COVID 19 Notifications handset display.
for transmitting beacons, see [7].

Beacons transmitted on one of the three radio channels used by Bluetooth LE values are reported by Bluetooth scans. Beacons are generally not decoded successfully and so no RSSI attenuation of around 75dB i.e. above this attenuation level is around -100dB in a Pixel 2, giving a maximum measurable RSSI below which beacons can no longer be reliably decoded (i.e. with duration zero) are left blank. Where appropriate we also include a solid line in plots that indicates the average attenuation level at each transmit power level (the average is calculated by weighting each attenuation level by the duration at that level and then summing over all attenuation levels).

The GAEN documentation does not precisely state how the attenuation level is calculated, nor does it give details as to how the attenuation duration is calculated. The analysis in [7], and subsequent personal communication from Google, establishes that the attenuation level is calculated as \( P_{TX} - P_{RX} \), where \( P_{TX} \) is the transmit power level sent in the beacon metadata and \( P_{RX} \) is given by a filtered RSSI measurements plus a calibration offset.

For the Google Pixel 2 handsets and GAEN API version 202512001 used in our experiments \( P_{TX} \) is -31dB and the calibration offset is -6dB. Google supplied us with the calibration and offset values used for all handset models in GAEN version 202512001 and we have posted these in our online study archive [1]. Note that we observed that the noise floor (the RSSI below which beacons can no longer be reliably decoded) is around -100dB in a Pixel 2, giving a maximum measurable attenuation of around 75dB i.e. above this attenuation level beacons are generally not decoded successfully and so no RSSI values are reported by Bluetooth scans.

3For Google Pixel 2 handsets (and others) the RSSI is recorded only from beacons transmitted on one of the three radio channels used by Bluetooth LE for transmitting beacons, see [7].

### III. Results

#### A. Attenuation vs Distance

Figure 3(a) plots the attenuation measured between two handsets placed at seat height in the aisle of the tram as the distance between them is varied. These measurements were taken using the standard Android Bluetooth LE scanner API (rather than the GAEN API). This scanner API reports an RSSI value for each received beacon. Following [7] updated to reflect GAEN calibration changes pushed by Google on 13th June 2020, for the Google Pixel 2 handsets used in our experiments we map from RSSI to attenuation level using the formula -31-(RSSI-6) dB.

It can be seen that the attenuation initially increases as the distance is increased from 0.5m to 1.5m, as might be expected. But thereafter the attenuation level stays roughly constant with increasing distance out to 2.5m. There is then a sharp rise in the attenuation at 3m. This corresponds to the end of a group of seats and the start of a flexible joint between two carriages. As the distance is increased further it can be seen that the attenuation starts to fall. The attenuation is around 52dB at 1.5m and around 60dB at 4m.

These baseline measurements indicate that the radio attenuation within the tram does not simply increase with the distance between handsets. This is similar to the behaviour observed in previous GAEN measurements taken on a bus [2], and is of course relevant to the use of attenuation level as a proxy for distance.

#### B. Attenuation Between Passengers

The full attenuation duration data reported by GAEN API is given in the Supplementary Material and is publicly available online [1]. In this section we analyse two aspects of this data: (i) the relationship, if any, between attenuation level and distance between handsets and (ii) the magnitude of the variations in the attenuation level induced by differences in the way participants hold their handsets.

#### C. Trend With Distance

Figure 4 plots the mean attenuation level vs the distance between handsets.
similar ranges of attenuation levels observed at all distances, except perhaps for distances below 1m where the attenuation level is more tightly clustered.

The GAEN API records the duration at each attenuation, and so effectively the full distribution of attenuation levels rather than just the mean. Figure 5 plots the sum-duration that the measured attenuation level is below 55dB, 63dB, 68dB and 73dB. For each pair of handsets these values are the rescaled empirical CDF of the attenuation level evaluated at the specified values. Recall that a typical definition of a proximity event is spending 15 minutes or more at a distance of 2m or less apart. We have therefore indicated the 2m distance with a vertical line in Figure 5, and attenuation durations greater than 15 minutes by the shaded areas.

For reliable detection of proximity events what one might like is that for an appropriate choice of threshold value the attenuation levels lie within the shaded area when the distance is less than 2m and outside the shaded area when the distance is greater than 2m. Unfortunately we do not see such behaviour in Figure 5. Instead, consistent with Figure 4 we see no consistent trend between attenuation duration and distance below/above 2m.

### D. Magnitude of Inter-Test Variations

Between each of the three experiments the participants switch seats. The seat positions themselves remain the same, only the person sitting in the seat changes, allowing us to see the impact of differences in the way that each participant uses their handset. For beacons transmitted from each seat position Figure 6 shows the mean attenuation level observed at the other seat positions (see the Supplementary Material for the full attenuation duration data). The attenuation level observed in test 1 is plotted vs the attenuation level observed in test 2. It can be seen that the points are clustered around the 45° line, but variations of ±10dB between the two tests are common.

### E. Exposure Notification True/False Positive Detection Rate

The GAEN API is intended for use by health authority Covid-19 contact tracing apps [3]. When a person is found to be infected with Covid-19 the TEKs from their handset are uploaded to a central server. The health authority app on another person’s handset can then download these TEKs, and use them to compare against the set of beacons received by the handset. If there is a match, the attenuation duration values reported by the GAEN API can then be used to estimate the risk of infection and trigger an exposure notification is this risk is sufficiently high.

A typical requirement is for a person to have spent at least 15 minutes within 2m of the infected person in order to trigger an exposure notification. The mapping from GAEN attenuation durations to exposure notification is therefore largely based on use of attenuation level as a proxy for proximity between handsets.

1) Swiss & German Exposure Notification Rules: Switzerland deployed a Covid-19 contact tracing app based on the GAEN API on 26 May 2020 [9]. The documentation for this app states that it queries the GAEN API with low and high attenuation thresholds of $t_1 = 50$dB and $t_2 = 55$dB and then bases exposure notifications on the quantity $ES = B_1 + 0.5B_2$, where $B_1$ is the attenuation duration below 50dB reported by the GAEN API and $B_2$ is the attenuation duration between 50dB and $t_2$ [10]. An exposure notification is triggered if $ES$ is greater than 15 mins.

Germany deployed a Covid-19 contact tracing app based on the GAEN API on 15 June 2020 [11]. The app is open source. By inspecting the documentation and code, and querying
the server API to obtain the app configuration settings\(^3\), we determined that the app follows an approach similar to the Swiss app for triggering an exposure notification, but uses values \(t_1 = 55\text{dB}\) and \(t_2 = 63\text{dB}\).

We applied the Swiss and German exposure notification rules to the tram dataset. Figure 7(a) plots the true and false positive rates for \(t_1 = 50\text{dB}\) and as \(t_2\) is varied from 55dB upwards and the ES threshold varied from 10 minutes to 15 mins. The mean rates are shown with one standard deviation indicated by the error bars. The mean and standard deviation are obtained by a standard bootstrapping approach\(^4\). Figure 7(b) plots the true and false positive rates when \(t_1 = 55\text{dB}\).

It can be seen from Figure 7(a) that selecting \(t_1 = 50\text{dB}\) and \(t_2 = 55\text{dB}\) (the values used in the Swiss app) yields no positive detections, despite approximately 50% of the handset pairs in the tram dataset being within a 2m distance of one another. Increasing \(t_2\) to 62dB and above yields a small increase in detection rate, with true and false detection rates roughly equal (we comment further on the implications of this below).

It can be seen from Figure 7(b) that selecting \(t_1 = 55\text{dB}\) and \(t_2 = 63\text{dB}\) (the values used in the German app) there are no detections when the threshold for ES is 15 minutes. Reducing this to 10 minutes, the true and false positive detection rates both rise to 9%. Increasing \(t_2\) does not increase these detection rates.

2) **Italian Exposure Notification Rule**: Italy deployed a Covid-19 contact tracing app based on the GAEN API on 1 June 2020 [12], [13]. The app is open source. By inspecting the documentation and code, and querying the server API to obtain

\(^3\)This means that the app configuration can be dynamically updated. We downloaded the detection settings from the server on 21 June 2020 and they are included in the study data repository [1].

\(^4\)The dataset was resampled with replacement \(n = 1000\) times, the exposure notification percentage calculated for each sample and then the mean and standard deviation of these \(n\) estimates calculated. We selected \(n\) by calculating the mean and standard deviation vs \(n\) and selecting a value large enough that these were convergent.

---

Fig. 7: Exposure notification true and false positive rates when the threshold strategy used in the Swiss and German contact tracing apps is applied to the GAEN tram dataset. Data is shown vs attenuation level and duration thresholds, solid lines indicate true positive rates and dashed lines the corresponding false negative rates.

Fig. 8: Exposure notification true and false positive rates when a simple threshold strategy is applied to the GAEN tram dataset. (a) True and false positive rates vs attenuation level and duration thresholds, solid lines indicate true positive rates and dashed lines the corresponding false negative rates. (b) ROC plot corresponding to mean rates in (a), dashed line indicates 45° line.

---

We downloaded the detection settings from the Italian app server on 21 June 2020 and they are included in the study data repository [1].
IV. DISCUSSION

A limitation of this study is that it is confined to handsets using the Android operating system. The GAEN API is also implemented on Apple iOS devices, but Apple have severely limited the ability of testers to make measurements (each handset is limited to querying the GAEN API a maximum of 15 times a day, and Apple has no whitelisting process to relax this constraint. Our measurement approach uses 34 queries to extract fine-grained attenuation data per pair of handset locations.

We equipped participants with the same model of handset in order to remove this as a source of variability in the data and instead focus on variability caused by the radio environment and the way that people hold their handsets. Google and Apple are currently undertaking a measurement campaign to select calibration values within the GAEN API with the aim of compensating for differences between handset models. We therefore expect that our measurements should also be applicable to a range of handsets, although this remains to be confirmed.

V. CONCLUSION

We report on the results of a Covid-19 contact tracing measurement study carried out on a commuter tram in Dublin, Ireland. Our measurements indicate that in the tram there is little correlation between received signal strength of distance between handsets. We applied the detection rules used by the Italian, Swiss and German apps to our measurement data and also characterised the impact on performance of changes in the parameters used in these detection rules. We find that the Swiss and German detection rules trigger no exposure notifications on our data, while the Italian detection rule generates a true positive rate of 50% and a false positive rate of 50%. Our analysis indicates that the performance of such detection rules is similar to that of triggering notifications by randomly selecting from the participants in our experiments, regardless of proximity.

ACKNOWLEDGEMENTS

The authors would like to extend their thanks to the Irish Health & Safety Executive (HSE) for arranging with Google for us to have whitelisted access to the GAEN API, and to Transdev Dublin Light Rail for kindly providing access to one of their trams. Trinity College Dublin, (the authors’ employer) funded the “Testing Apps for Contact Tracing” (TACT) project that has allowed us the time and handsets required here. We emphasise that any views expressed in this report are the authors own, and may not be shared by the HSE, Transdev or Trinity College Dublin.

REFERENCES


Figures 9 - 11 plot the exposure information between each pair of handsets reported by the GAEN API for each of the three experiments. To assist with interpreting the plots the reports in each plot are ordered by increasing distance between the pairs of participants (see Figure 1(a)). No data is shown when no beacons were received between a pair of handsets, e.g. between participants 2 and 3 in Figures 9(b) and 9(c). It can be seen that occasionally there is an increasing trend in attenuation, for example see Figures 9(c) and 11(c), but this is infrequent. Occasionally there is a decreasing trend in attenuation, for example see Figures 9(e) and 11(d). Overall, however no consistent trend is evident in the change in attenuation level with increasing distance.

In Figure 11 participants 1 and 3 place their handsets in their left trouser pocket rather than their hand. Intuitively, one might expect this change to increase the attenuation level since the participants body is now more likely to affect transmission and reception of radio signals. However, comparing Figures 11(a) and 11(c) with Figures 9 and 10 it can be seen that this change does not cause any consistent change in the observed attenuation level. For example, comparing Figures 11(a) and 10(a) the attenuation level between participants 1 and 5 decreases from test 2 to test 3, while the attenuation level between participants 1 and 3 increases.

Fig. 9: Attenuation durations reported by GAEN API on completion of the first test. Pairs indicated on x-axis in each plot are ordered by increasing distance.
Fig. 10: Attenuation durations reported by GAEN API on completion of the second test (with the same participants as in the first test, but with their seating positions swapped about). Pairs indicated on x-axis in each plot are ordered by increasing distance.

Fig. 11: Attenuation durations reported by GAEN API on completion of the third test (with the same participants as in the first test, but with their seating positions swapped about and participants 1 and 3 with handsets in their trouser pocket rather than their hand). Pairs indicated on x-axis in each plot are ordered by increasing distance.