Assessing the Impact of Multi-Channel BLE Beacons on Fingerprint-based Positioning

Jovan Powar*, Chao Gao*, and Robert Harle*

Abstract—Bluetooth Low Energy (BLE) beacons are attractive for indoor location due to their ease of deployment, wide support on consumer devices and low cost. Co-ordinate location can be estimated using radio fingerprinting techniques applied to Received Signal Strength (RSS) values from BLE beacons. Earlier work on BLE fingerprinting has noted that the separate BLE advertising channels can exhibit different RSS values, although modern consumer devices do not provide channel information and therefore report a composite RSS with artificially inflated variance that can lead to reduced positioning accuracy.

In this work we quantify how often this channel dispersion occurs in typical environments, what the extent of the dispersion is, and what impact it has on positioning. Furthermore we explore how to obtain channel information when using today’s consumer devices.

We find that the dispersion is both common and significant and leads to significantly reduced positioning accuracy; in our experiments only 10% of composite variances were within 1 dBm of an individual channel’s variance, and 40% of composite variances were at least 4 dBm greater. We also show that the behaviour of the channels is sufficiently distinct that a fingerprinting scheme that uses a signal map for each achieves significantly increased positioning accuracy (up to 3 m).

I. INTRODUCTION

Much of the research into indoor positioning systems aims to produce a system that is robust, accurate and scalable. Unfortunately robust, accurate systems have tended to rely on extensive infrastructure support that hinders scalability. Similarly, scalable systems typically suffer from robustness issues since they are based on opportunistic signals over which there is, by definition, no control. Fingerprinting systems fall into the latter category, whereby a spatially-variant quantity is measured and compared to a previously-captured map of that same quantity in that space.

The ubiquity of WiFi has seen it dominate fingerprint-based systems, mostly using Received Signal Strength (RSS) as the spatially-variant quantity. However, the opportunistic use of WiFi has a number of serious shortcomings. Firstly, the access points are deployed for optimal communications rather than positioning. This means access point ranges overlap only as much as is necessary and the access points themselves are rarely deployed with good positioning geometry in mind. Secondly, continuous WiFi scanning is both power hungry and disruptive on the mobile device (an active scan limits other communications). Thirdly, devices are not optimised to scan for positioning. They often report the observed access points at the end of the scan, which can take anywhere between 0.5 s and 5 s (the latter associated with scanning the larger 5 GHz band). The effect of this is that a moving device can only output fingerprints at a low rate and those readings are associated with a large spatial range rather than at a single location—the fingerprints are ‘smeared’ across space.

Bluetooth Low Energy (BLE) is emerging as an answer to these issues. BLE features an advertising mode whereby a message is regularly broadcast to surrounding devices. A BLE beacon simply broadcasts its identity in a locality, traditionally providing proximity-based location. However, a set of BLE beacons can be used for fingerprinting just as with WiFi. The advantages of BLE include wide consumer device support; lower power continuous background scanning; simple, low power beacons that can be battery powered for years on a single battery and deployed anywhere; support for high density deployments with minimal interference; and instantaneous beacon reporting at the handset (no fingerprint smearing).

Although BLE uses the same 2.4 GHz ISM radio band as WiFi, it does so in a very different way. It divides the allocated spectrum into a series of narrow radio channels, three of which are used for advertising. Consumer BLE devices do not report the channel a particular beacon is heard on, and so BLE beacon fingerprinting to date has not tended to use the channel information. However, these channels are widely spaced in the frequency domain and we show that they often experience very different propagation characteristics in indoor environments. In this paper we advocate the use of per-channel signal maps rather than the composite channel map that it used presently. We argue that the latter often leads to artificially high variances in the survey map and we show this leads to sub-optimal positioning.

In this paper we make the following contributions:

- we investigate how to obtain BLE advertisement channel information in a device-agnostic and standards-compliant manner;
- we quantitatively assess how often and by how much the RSS values between advertising channels separate in typical indoor environments; and
- we quantitatively assess the impact on positioning accuracy.

The remainder of the paper is structured as follows. Section II discusses related work, while Section III describes the BLE advertising process in detail. Section IV shows how to adapt BLE beacons to allow the extraction of channel information at a consumer listener. Section V describes our evaluation of the value of using channel information, analysing the raw

*Jovan Powar, Chao Gao, and Robert Harle are with the Department of Computer Science and Technology, University of Cambridge, 15 JJ Thomson Avenue, Cambridge CB3 0FD, UK jsp50@cam.ac.uk, cg500@cam.ac.uk and rkh23@cam.ac.uk

978-1-5090-6299-7/17/$31.00 © 2017 IEEE
RSS data and the resultant positioning. Finally, Section VI concludes.

II. RELATED WORK

Indoor positioning is a mature research field—comprehensive overviews can be found in [14], [13], [11]. In this work we focus on radio positioning using empirical fingerprinting techniques such as those developed in [2], [11], [16], [12]. These involve positioning by pattern matching current observations to a previously-obtained survey of a space. Although applied to different radio technologies, these techniques have been developed primarily with WiFi in mind.

Positioning with Bluetooth prior to version 4.0 (so-called Bluetooth Classic) has used various techniques from proximity [9], [4] to trilateration [6], [15] to fingerprinting [5], [15]. The limiting factor of these techniques has been the slow scan rate of 10 s or more. Such latencies are not present in BLE [3].

We have previously looked at fingerprinting with BLE [7], where we noted the channel separation and the issues it might give. However, we did not look at how common the separation problem is in common environments, nor how pronounced. These issues were also noted by Zhuang et al. [17], who developed a statistical model combining per-channel fingerprinting with a regression model and extended Kalman filtering. They showed that channel information improved positioning when the channels diverge. However, they did not investigate how often or to what extent the individual signal differs from the composite signal. Furthermore they relied on Apple’s now-obsolete iOS 8 to obtain channel information.

III. BLE ADVERTISING

A. Specification

BLE supports both message broadcast (one to an unspecified many) and direct communication between device pairs. All transmissions are in the 2.4 GHz ISM radio band, which BLE splits into forty 2 MHz channels. In the current specification (4.x) three of these channels are dedicated broadcast channels for BLE advertisements. These are nominally labelled 37 (centre frequency 2402 MHz), 38 (2426 MHz) and 39 (2480 MHz). Since advertisements are broadcast there is no opportunity to use interference mitigation schemes such as the frequency hopping BLE uses once in a connection with another device. Similarly, no channel sensing occurs prior to a transmission. Instead, the advertising channel frequencies are chosen to avoid the dominant WiFi channels (1, 6 and 11, see Figure 1) and the system gains redundancy by broadcasting each advertisement on each of the three advertisement channels in quick succession.

A BLE beacon will send out advertisements on a regular basis. To prevent multiple in-range beacons accidentally synchronising and repeatedly interfering, each beacon period is jittered randomly. Between advertisements the hardware enters a deep sleep mode to conserve energy.

B. Device Listening Schemes

A device scanning for BLE device advertisements does so by cycling over the three advertising channels, pausing on each and reporting any packets it observes. The length of the pause on each channel is device-specific. In the case of consumer smartphones, manufacturers have adopted a range of listening strategies. At one extreme we have observed a consumer device that listened on each channel for 5 s before switching to the next. However, a more common strategy is to cycle through the channels tens of times per second: the iPhone 6 and iPad Pro, for example, spend only a few milliseconds on each channel in their current incarnations.

To complicate matters, some devices apply a duty cycle to their scanning, whereby they turn off listening altogether for limited periods in order to conserve energy. Current versions of iOS, for example, listen continuously (i.e. a 100% duty cycle) for 60 s, before throttling back to a much lower duty cycle. Crucially, the BLE specification does not provide a mechanism for a BLE controller to report the channel on which it received a given message. Non-standard extensions may make this possible (for example, iOS 8 supplied this information) but we are not aware of any current smartphones that give access to the channel information. Collecting a sequence of advertisements at one location is then a random sampling of the three channels. Properties of the sequence such as RSS values can be viewed as samples from a composite signal.

C. Implications For Positioning

The three BLE advertisement channels have well-spaced centre frequencies and may therefore have very different propagation characteristics, which translates to drawing RSS measurements from distinct distributions. Composite measurements will exhibit a higher variance whenever distributions have poor overlap since the measurement will be a sampling from a random choice of distribution. As an example, Figure 2 shows an example RSS capture for a specific beacon taken from the survey we describe later in this paper. We observe that the three advertising channels are exhibiting reasonably constant outputs, while the composite (in purple) appears far noisier as it jumps between them randomly.

To use such a composite signal for positioning with radio maps requires care. More readings are required to account for the increased noise and to make it more likely that a fingerprint is formed from packets broadcast on each channel, capturing a full profile of the signal. Even then, the higher noise will limit the spatial locality of the signal and adversely impact positioning accuracy.

Conversely, if full channel information is available then separate radio maps can be created for each channel. We hypothesise that each map would have the same or (more likely) smaller variance than the composite map at any given position. We expect this to improve positioning accuracy since the map for a single channel would be of a higher quality than the composite. In fact, we expect that the single channel
maps might be sufficiently distinct to improve positioning even further by using all three in one scheme.

IV. Obtaining Channel Information

The BLE specification does not include an API by which a listening BLE chip can report the channel it received an advertisement on. Previous work using BLE channel information such as Zhuang et al. [17] depended on using Apple’s iOS 8 operating system, which provided this information. This version is now outdated and Apple have removed the channel information from current versions. Other devices such as Android handsets have never made channel information available. Our listening device-agnostic approach is to adapt the beacon to broadcast on only one channel.

In an ideal world each beacon would add the channel information to the advertisement message at the point of transmission. Although technically feasible, BLE performs the advertisement process in hardware and manufacturers do not provide access to the low level firmware to make this modification to off-the-shelf devices. It is, however, likely that a manufacturer seeking to produce a beacon optimised for positioning would do this.

In the absence of this capability it is possible to create similar behaviour using channel maps in BLE. This allows blacklisting of specific transmission channels, ostensibly to avoid channels suffering strong interference during frequency hopping communications. In practice, this capability is used primarily for data connections, although the specification does have provision for blacklisting advertising channels. As such a standard transmitter can be set to broadcast on a single advertisement channel by manually setting a bit field such that only one of the channels is enabled.

The Bluetooth 4.x standard defines a command \texttt{HCI\_LE\_Set\_Advertising\_Parameters} that provides this functionality. One of the arguments to this function, \texttt{Advertising\_Channel\_Map}, is a bit field that indicates which of the channels 37, 38, and 39 are to be used for transmitting advertising packets. This channel map can be set at the time of device configuration to enable/disable channels before advertising begins (it cannot be changed during an ongoing advertisement cycle). The specification’s description of the map is reproduced in Table I.

In this work we create beacons from Raspberry Pis fitted with a BLE USB dongle. The advertising frequency was set at 31.25 Hz\(^1\). The multi-channel scheme identification in [7] involved using three beacons stacked on top of each other, with the composite signal formed by simply concatenating the three channels. However, this may introduce artificial differences: different hardware, different antennas and slightly different locations could all imply channel differences that would not be observed in a single piece of hardware. Instead, we used a single BLE chip and antenna and cycled through four modes: one that transmitted only on 37, one only on 38, one only on 39 and one with no restrictions (i.e. all channels in use as per the default). Each mode was associated with a unique payload to allow the receiver to identify the channel being used. The beacons continuously cycled through the four modes, pausing on each for a preset time, \(T\).

\(^1\)This high advertising rate helped us to collecting statistically significant measurements at each survey point quickly. In practice fewer samples would be needed and/or the data collection time at each point could be longer.

<table>
<thead>
<tr>
<th>Value</th>
<th>Parameter Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>00000000b</td>
<td>Reserved for future use</td>
</tr>
<tr>
<td>xxxxxxx1b</td>
<td>Channel 37 shall be used</td>
</tr>
<tr>
<td>xxxxx1xb</td>
<td>Channel 38 shall be used</td>
</tr>
<tr>
<td>xxxx1xxb</td>
<td>Channel 39 shall be used</td>
</tr>
<tr>
<td>00000111b</td>
<td>Default (all channels enabled)</td>
</tr>
</tbody>
</table>

TABLE I: Advertising\_Channel\_Map specification in Bluetooth 4.0 (Volume 2, Part E, Section 7.8.5)
### V. Evaluating the Value of Channel Information

#### A. Experimental Setup

We investigated the differences between single-channel and composite advertising by performing surveys of advertising RSS in three rooms. Room 1 was a computer room at the Computer Laboratory in Cambridge, approximately 18 m by 12 m, containing four large circular desks housing 10 desktop computers, and a thick divider cordonning off an area approximately 6 m by 12 m. Room 2 was the living area in a small flat, and part of the corridor leading up to it. The room contained a glass dining table, a large sofa, armchair, coffee table, and a television standing atop a small table. The room measured approximately 6 m by 4 m. Room 3 was the open plan living space in the ground floor of a small house. It measured approximately 7 m by 3 m, and contained two sofas, a stone and wood dining table, a number of chairs, and a snake enclosure.

Four beacons were placed in each location, widely spaced apart. The placement of beacons was chosen to emulate how such a system might realistically be deployed in the environment (an office or a home). We expected that by investigating a naturalistic deployment—on walls, desks, and shelves—we would be able to properly characterise the differential behaviour of each advertising channel.

The listening device was an iPhone 6 with iOS 10 running a custom app that logged BLE advertisements to file. It was placed atop a non-metallic stand of height 1.3 m. As mentioned in Section IV, we used a mode cycle time of \(T = 10\) s. To ensure all channels were observed, the app recorded 43 s of data at each site. In total, we surveyed at 93 points in Room 1, 35 in Room 2, and 31 in Room 3. Ground truth location was obtained in Room 1 by alignment with the 0.6 m square floor tiles; and using a Google Tango device in Room 2 and Room 3.

#### B. Single-channel vs. Composite Raw Data

We computed the RSS mean and standard deviation at each survey position for each test site, separating the results by channel or composite and plotting the values as empirical cumulative distribution functions. The CDFs of the mean RSS values (Figures 4(a), 4(c) and 4(e)) do not clearly distinguish the composite channel from the others. This is to be expected: the composite channel should exhibit a mean that lies between the mean range of its constituent channels. The CDFs of the standard deviations (Figures 4(b), 4(d) and 4(f)), however, clearly distinguish between the separate channels and the composite. The lines indicate higher variances are significantly more likely for the composite channel. Taking the larger Room 1 as an example, the three separate channels exhibit standard deviations of around 2 dBm at the 90th percentile. In contrast, 2 dBm is at the 20th percentile for the composite channel (around 6 dBm at the 90th percentile).

These graphs capture the overall distribution of variances, showing that there must be survey points where the variance of the composite differs from the separate signals. They do not, however, quantify how often there is a difference. To this end, we computed the differences between the composite mean \(\bar{C}\) and the individual channel means \(\bar{c}_i\) for channel \(i\) for each survey position:

\[
|\bar{C} - \bar{c}_i|,
\]

Similarly, the difference in standard deviations was computed as:

\[
\sigma_c - \sigma_i,
\]

where \(\sigma_c\) is the composite signal deviation and \(\sigma_i\) is the variance for channel \(i\) at that survey position. We retain the sign information for the deviation differences to emphasise that the composite variance is increased.

Figure 5(a) shows the CDFs for the difference in means in the larger Room 1 (results for the other rooms were similar but omitted to save space). They illustrate that around 80% of the composite mean values differed from the individual channel means by over 1 dBm, while 90% had a difference of up to 6 dBm. Figure 5(b) shows the CDF for the difference in deviations. It shows that around 10% of the composite deviations were comparable (within 1 dBm) to the individual
Fig. 4: Summary of raw RSS survey data
channel deviations at the same point. Thus almost every survey point was exhibiting a significantly higher variance for the composite than the individual channels.

Figure 5(b) also has a small number of survey sites where the difference was negative. Further investigation revealed these cases were associated with periods of large outliers in the RSS during one period of data collection—see Figure 6 for an example. Such outliers are consistent with a short-lived change to the environment, such as a person moving around (which did occur during the surveys). Because we measured the channel RSS values in series rather than parallel, a transient effect such as this would only affect the channel that was currently being measured, giving it an unexpectedly high variance. We include these results for completeness, but stress that they are irregular outliers—negative differences were observed for only 39 of 1,196 survey readings (3%) and all displayed short-lived disruption on one channel.

Overall the three rooms exhibited remarkably similar behaviour in terms of the dispersal of the channel RSS values. Taken together these results validate our hypothesis that not only can the three BLE advertising channels disperse, but this is a common occurrence in real world environments. Moreover the dispersal is of a non-negligible magnitude.

C. Regression-based Comparison

Regression techniques are commonly used to generate more comprehensive RSS signal maps from raw survey data. Gaussian Process (GP) regression has emerged as the preferred regression scheme for this domain. GP regression is a non-parametric technique that models the signal at each location using a Normal distribution. As such, locations are associated with a mean level and a variance, and the latter can be used to further emphasise the value in separating channels. Figure 7 gives an example GP output for a specific beacon in Room 1. We observe that the variances for the composite channel (Figure 7(i)) are notably higher than the variances for the individual channels (Figure 7(j), 7(k) and 7(l)).

We have previously developed a metric, named RSS$_{90}$, to quantitatively compare two GP maps. We give an overview here—full details are in [10]. Within two GP maps covering the same area, the RSS$_{90}$ value for a given point represents the value within which the two maps would agree 90% of the time. If the maps match at that point, the RSS$_{90}$ value will be small. Larger values indicate disagreement between the maps at that point.

We evaluated the RSS$_{90}$ value at each survey point for each pair of channels. Figure 8 visualises the results for two pairs. We found that the maps exhibited RSS$_{90}$ values in the
<table>
<thead>
<tr>
<th></th>
<th>Composite Signal</th>
<th>Channel 37</th>
<th>Channel 38</th>
<th>Channel 39</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey Measurements</td>
<td><img src="image1" alt="Image" /></td>
<td><img src="image2" alt="Image" /></td>
<td><img src="image3" alt="Image" /></td>
<td><img src="image4" alt="Image" /></td>
</tr>
<tr>
<td>GP $\mu$</td>
<td><img src="image5" alt="Image" /></td>
<td><img src="image6" alt="Image" /></td>
<td><img src="image7" alt="Image" /></td>
<td><img src="image8" alt="Image" /></td>
</tr>
<tr>
<td>GP $\sigma^2$</td>
<td><img src="image9" alt="Image" /></td>
<td><img src="image10" alt="Image" /></td>
<td><img src="image11" alt="Image" /></td>
<td><img src="image12" alt="Image" /></td>
</tr>
</tbody>
</table>

Fig. 7: Sample GP maps for Beacon 3 in Room 1. Please note that on the bottom row, the red areas (where the variance values are much higher than else where) correspond to the areas where no signal samples were made (because of table areas etc.) during the data collection.

range 15–20 dBm, which indicates a statistical difference as expected.

D. Positioning Comparison

We evaluated the positioning using the signal strength-based location estimation method described in [8]. A standard Bayesian localisation algorithm was used to incorporate both the mean and variance estimate at each GP map location. To estimate a location, $x$, conditioned on observation of a set of signal strength measurements (fingerprints), $z$, we used:

$$p(x|z) \propto p(x)p(z|x)$$  \hspace{1cm} (3)

where $p(x)$ is the prior probability of $x$ and we set $p(x)$ to a uniform distribution over the whole space to simulate one-shot positioning; and $p(z|x)$ is the likelihood of observing a set of signal strength measurements at location $x$. $p(z|x)$ is given directly by the GP regression (where channel information was used, we formed a separate GP map for each channel). We divided the environment into a grid of square cells of length 0.5 m. For each positioning request we computed $p(x|z)$ at each cell and selected the cell with the maximum as the location estimate. Signal strength measurements lower than $-90$ dBm were discarded.

<table>
<thead>
<tr>
<th></th>
<th>RSS$_{90}$ map</th>
<th>RSS$_{90}$ CDF</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image13" alt="Image" /></td>
<td><img src="image14" alt="Image" /></td>
<td><img src="image15" alt="Image" /></td>
</tr>
<tr>
<td><img src="image16" alt="Image" /></td>
<td><img src="image17" alt="Image" /></td>
<td><img src="image18" alt="Image" /></td>
</tr>
</tbody>
</table>

Fig. 8: RSS$_{90}$ maps and CDFs. The results for other channels are similar so are omitted here for brevity.
In further work, we intend to investigate the use of beacons that use multiple beacons, as an alternative to the channel switching scheme we have outlined. We also hope to further characterise dispersion and its implications for positioning using larger testbeds and by investigating the effect of varying beacon parameters, such as transmission power.

VI. CONCLUSIONS

Our results have shown that the different RSS behaviour of the three BLE advertising channels almost always has a significant effect on the composite signal, as well as significant implications for positioning.

We have shown that channel dispersion occurs frequently and to great extent in typical environments: 80% of composite mean values differed from the individual channel values by at least 1 dBm in our experiments. Furthermore the variance of a beacon’s RSS value is artificially increased by taking the composite channel: in our experiments only around 10% of composite variances were within 1 dBm of an individual channel’s variance, and 40% of composite variances were at least 4 dBm greater. This raw data analysis indicates that a single channel signal is highly preferable to a composite signal, and this is borne out by our positioning results. Beyond simple dispersion, we have also shown that the three channels behave distinctly in typical environments. Thus by constructing three signal fingerprints, we arrive at a greatly improved positioning scheme.

We hope this work will motivate addition of the channel information in future BLE standards. In the interim, however, we have outlined a standards-compliant scheme for beacons to cycle through the channels via channel maps.

Figure 9 gives the CDF of the positioning results for Room 1, which are also summarised in Table II. We found that the use of the channel information and its lower-variance maps produced a significant improvement in accuracy—around 3 m—at all percentiles.

VI. CONCLUSIONS

Our results have shown that the different RSS behaviour of the three BLE advertising channels almost always has a significant effect on the composite signal, as well as significant implications for positioning.

We have shown that channel dispersion occurs frequently and to great extent in typical environments: 80% of composite mean values differed from the individual channel values by at least 1 dBm in our experiments. Furthermore the variance of a beacon’s RSS value is artificially increased by taking the composite channel: in our experiments only around 10% of composite variances were within 1 dBm of an individual channel’s variance, and 40% of composite variances were at least 4 dBm greater. This raw data analysis indicates that a single channel signal is highly preferable to a composite signal, and this is borne out by our positioning results. Beyond simple dispersion, we have also shown that the three channels behave distinctly in typical environments. Thus by constructing three signal fingerprints, we arrive at a greatly improved positioning scheme.

We hope this work will motivate addition of the channel information in future BLE standards. In the interim, however, we have outlined a standards-compliant scheme for beacons to cycle through the channels via channel maps.

In further work, we intend to investigate the use of beacons that use multiple beacons, as an alternative to the channel switching scheme we have outlined. We also hope to further characterise dispersion and its implications for positioning using larger testbeds and by investigating the effect of varying beacon parameters, such as transmission power.

TABLE II: Statistics of CDFs in Figure 9 (all in metres).

|                | 50%   | 75%   | 95%    | mean
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No channel info</td>
<td>8.1</td>
<td>10.9</td>
<td>15.1</td>
<td>8.0</td>
</tr>
<tr>
<td>With channel info</td>
<td>3.8</td>
<td>7.8</td>
<td>13.2</td>
<td>5.2</td>
</tr>
</tbody>
</table>

Figure 9: CDF of positioning results

REFERENCES