

Bluetooth scanning

why it is bad to have a device in discovery mode

J.C. van Nieuwenhuizen

November 26, 2006

1 Introduction

A lot of research has been put in getting data from a specific targeted bluetooth device. In this article we explore an approach in which we use one stationery scanning point which scans for bluetooth devices in the neighbourhood. Further more we show some examples of what can be done with the acquired data.

2 Scanning

The first step that had to be taken to do our analysis, was to design the software and database we were going to use to collect the data. We had the following requirements for our scanning tool:

1. Being able to scan for hardware addresses of bluetooth devices in the vicinity
2. Being able to get the name from a bluetooth device
3. Being able to store the found data in a database

2.1 Hardware

The hardware we choose to use was a Trust BT-1300TP bluetooth usb adapter. This adapter should be able to communicate over a distance of 100 meters, we should however note that this range was limited by the walls surrounding our scanning point.

2.2 Software

The software we created to scan was programmed in Perl, but makes use of the build in FreeBSD bluetooth program *hccontrol*. This Perl program logged to a simple PostgreSQL database. Each device we see is logged when

it is discovered in the current scan cycle and was not seen in the previous scan cycle or the last time the device was logged as present was more than 5 minutes ago. This choice was made to avoid flooding our log when only one device is present.

3 Results

In this section we explore the results of our scanning. The first few days of our scanning we discovered between 600 and 800 new devices each day. Later on this dropped, but remained between 400 and 600 new devices per day. We suspect that if we had extended our scanning period with a few more weeks this number would have dropped further.

3.1 Devices we see often

A simple query can be asked against the database to find the devices we saw far more often than other devices. I.e XX:XX:XX:XX:26:eb is a device we saw constantly and is likely an appliance in our neighbourhood.

A more interesting device was however XX:XX:XX:XX:10:78 this device was only seen between 8:00 and 17:00 but then very often. On Thursday when the shops are also open in the evening we saw this device also after 17:00 but never after 21:30. Which might indicate that this device belongs to someone working in the shops in the vicinity.

3.2 Devices we see together

Another interesting point that we tried to extract from our database was whether there are devices we saw together on more than one day at the same time. This would indicate that someone has 1 or more enabled bluetooth devices with him/her. But it can also indicate that 2 persons know each other. As the database grows over time the change of making a correct distinction between those two will probably increase.

A simple example from our database to illustrate the kind of data that can be extracted this way:

Device	Date - Time
XX:XX:XX:XX:65:3b	2006-07-29 11:08:07.512983
XX:XX:XX:XX:65:3b	2006-07-29 11:37:24.924039
XX:XX:XX:XX:65:3b	2006-07-29 11:44:45.770688
XX:XX:XX:XX:65:3b	2006-07-29 11:48:48.062865
XX:XX:XX:XX:65:3b	2006-08-05 12:04:17.446093
XX:XX:XX:XX:65:3b	2006-08-05 12:32:46.915755

Device	Date - Time
XX:XX:XX:XX:1d:45	2006-07-29 11:37:21.618499
XX:XX:XX:XX:1d:45	2006-07-29 11:44:30.599211
XX:XX:XX:XX:1d:45	2006-07-29 11:48:54.937834
XX:XX:XX:XX:1d:45	2006-08-05 12:32:39.821909

Device	Date - Time
XX:XX:XX:XX:19:b5	2006-07-29 11:08:14.572135
XX:XX:XX:XX:19:b5	2006-07-29 11:44:24.433205
XX:XX:XX:XX:19:b5	2006-07-29 11:48:38.777555
XX:XX:XX:XX:19:b5	2006-08-05 12:04:01.838533
XX:XX:XX:XX:19:b5	2006-08-05 12:32:39.842027

The occurrence of the devices at the same time on more than one day at least suggests a connection between these devices. An important point to notice is that these devices are not in our list of often seen devices, which further reduces the chance that seeing them at the same time is a coincidence.

3.3 Twice a day

In our analysis some devices showed up 2 times a day with an interval of 10-50 minutes between sightings. After a few days this pattern repeated for the same set of devices. Because our scanning point is located across a supermarket it is likely that those devices belong to customers of this supermarket.

3.4 Devices and their name

Because people who take the effort to change the name of their device will probably also put personal information in we decided to run a query against the found names of devices. The names we found can roughly be divided in to the following four categories: name, name and surname, default device name as set by the manufacturer and other.

We took this analysis a step further and queried for devices which have changed their name over time. One of the results here was someone changing the device name from his name to his name followed by loves her name. The most were however changes from the default manufacturer name to the persons own name.

4 Improvements and further research

One of the problems we had during scanning was that while we were trying to determine the name of a found bluetooth device we could not scan for other devices. Especially we have to wait a whole timeout cycle when the device for which we are trying to determine the name has moved out of

range. A solution for this problem would be to have one or more dongles dedicated to scanning for new devices and one or more dongles dedicated to name resolution.

Another point of improvement is probably targeting our database design more towards the data we want to extract. The current model is a very flat and simple design which makes some analysis very expensive. In example finding devices which are often seen at the same time at the same location is relatively slow using the current design.

5 Conclusion

The primary idea of the scanning was to find trends like how busy is the centre of the city based on day of the week and/or weather conditions outside. We however found out that more information can be extracted from the scan data. I.e. Names can be found, a guess at why the person carrying the device is in in the neighborhood of our scanning device can be made and relationships can discovered. Some of this data is however open to interpretation, and no hard conclusion can be made for all devices. To protect your self against this form of data gathering it is however necessary to disable the discovery mode of your bluetooth enabled device.