

# Social Density Estimation – Experimental Setup (Draft)

Guntur Dharma Putra

October 4, 2016

## 1 Introduction

This document presents an experimental setup for social density estimation using WiFi. The correlation between the number of unique device and available Access Point in a particular area is investigated. The unique devices are sensed using WiFi probe-request, while number of available access points is taken from normal scan in WiFi managed mode. This document also describes the experimental setup for MAC address randomization experiment. Furthermore, the availability of voice activity will also be taken into account.

## 2 Experimental Setup

The experimental setup is divided into two sections: WiFi sensing and MAC address randomization.

### 2.1 WiFi Sensing

The objective of WiFi sensing is to capture WiFi probe-request packets and scan and count available Access Point. The device will capture probe-request packets for 10 minutes, and then scanning available WiFi access point. This cycle is repeated accordingly, depending on the total scanning duration. As a validation, Voice Activity Detection (VAD) technique is also implemented. The scanning process is automated using a bash script.

#### Sensing setup

The experimenter captures the probe-request packets, scans for available Access Point, and creates an audio recording of the surrounding. The experimenter and the scanning instrument will remain still at a certain location during this process. The instrument for this setup is only a laptop with built-in microphone and network card.

### **Sensing Area**

The area is classified to:

- Outdoor
  - High density of crowd
  - Medium density of crowd
  - Low density of crowd
- Indoor
  - Large-sized hall, e.g., university library
  - Small-sized hall, e.g., house

#### **Note:**

University complex will also be considered as it has medium to high crowd density but only (possibly) one available SSID (**eduroam**). For each sensing, the experimenter records the GPS coordinates as well. Other places are also possible.

### **Duration of Sensing**

The duration of sensing is classified to:

- Short (15 to 20 minutes)
- Medium (30 to 40 minutes)
- Long (60 to 90 minutes, or more)

## **2.2 MAC Address randomization**

This research is trying to find a way to overcome MAC address randomization in Android. In order to do so, the experimenter will use laptop and scan for any MAC address changes in a location where no other WiFi probe-requests can be captured, i.e., remote areas. The experimenter first tries to redo the experiment carried out in [1, 2, 3].

## **3 Data Analysis**

The experimenter must carry out data analysis once the log seems to be adequate.

### 3.1 Data Preprocessing

The log will be filtered out from unidentified MAC addressed based on Organizationally Unique Identifier (OUI)<sup>1</sup>. The unidentified OUI, which might be a randomized MAC address, will be filtered out. Furthermore, data preprocessing also involves duplicates removal.

**Note:**

Some people might just pass the location, but they also broadcast probe-request, which may affect the result. This point will also be taken into consideration.

### 3.2 Voice Activity Detection

In this analysis, we use library from [4, 5], which is available online. The VAD result will be matched with the WiFi scanning result. In this scenario, the experimenter uses two available libraries: a) robust VAD<sup>2</sup> and b) Crowd++<sup>3</sup>.

## 4 Result

I did a quick experiment to scan for WiFi probe request and available AP both in indoor (IKEA Groningen) and outdoor (Grotemarkt Groningen) on last Saturday. The result, in my opinion, is promising for further experiment.

### 4.1 IKEA Groningen

IKEA provides an affordable breakfast menu until 10:30AM, which attracts many customers. This makes IKEA a great location to do social density experiment. We can see the result in Figure 1 and 2, in which blue and green line denote number of access point and unique devices respectively. We can see that around 400 MAC addresses are removed.

---

<sup>1</sup>[https://en.wikipedia.org/wiki/Organizationally\\_unique\\_identifier](https://en.wikipedia.org/wiki/Organizationally_unique_identifier)

<sup>2</sup><https://github.com/mvansegbroeck/vad>

<sup>3</sup><https://github.com/lendlice/crowdpp>

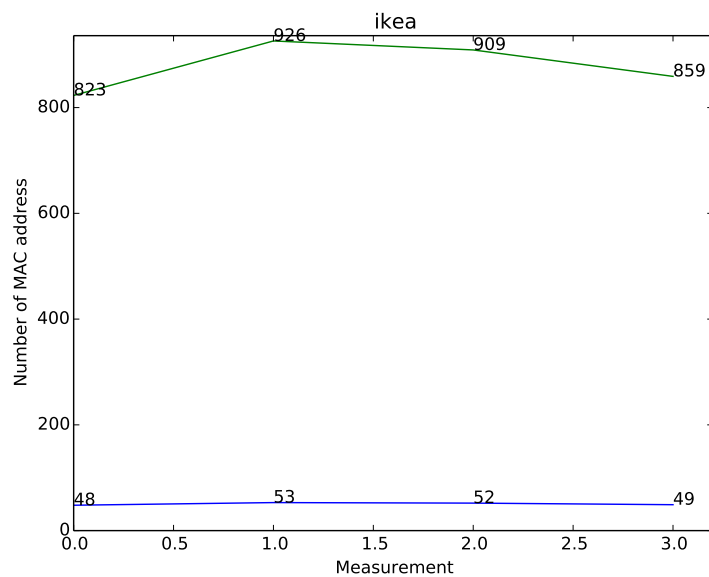


Figure 1: IKEA experiment result before MAC address removal.

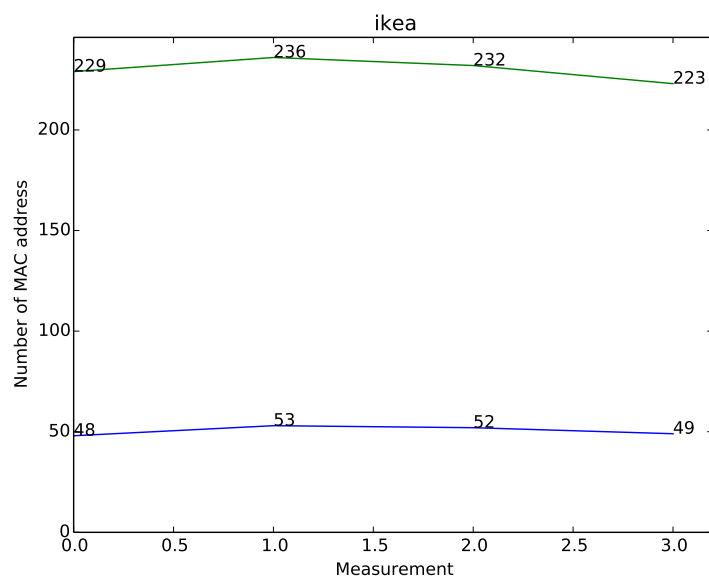


Figure 2: IKEA experiment result after MAC address removal.

## 4.2 Grotemarkt Groningen

Grotemarkt Groningen is a favorite place for many Groningers in the weekend, which makes this place relatively crowded during weekends. As we can see in Figure 3 and 4, the number of devices drop significantly. This is possibly due to this location, which is very dynamic, which has high probability of people going in and out in this area.

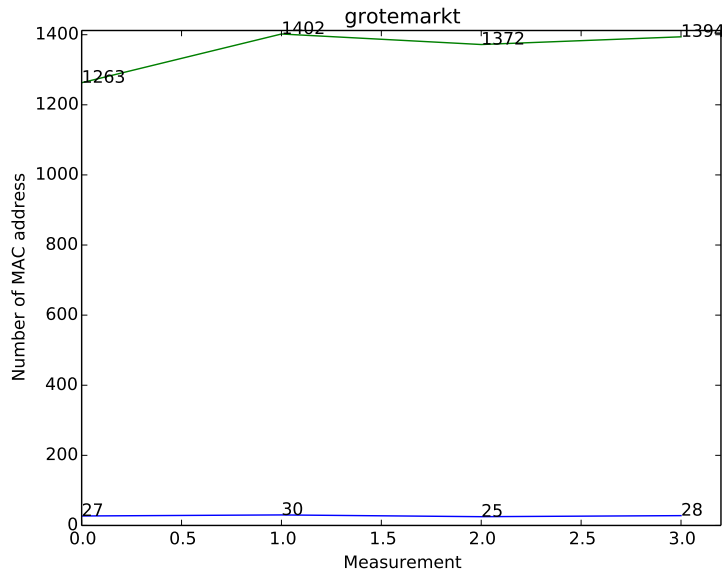


Figure 3: Grotemarkt Groningen experiment result before MAC address removal.

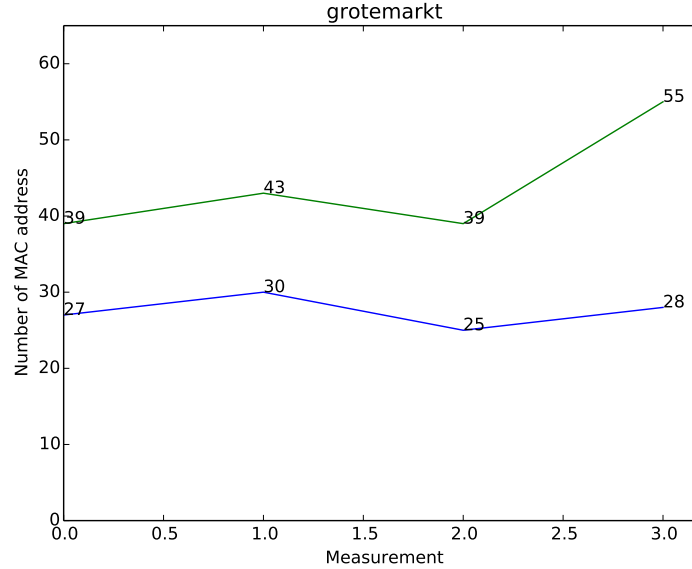


Figure 4: Grotemarkt Groningen experiment result after MAC address removal.

**Note:**

If we have adequate data, we can present the result in a scatter graph, which could give us clearer correlation overview.

## References

- [1] E Vattapparamban, B S Çiftler, I Güvenç, K Akkaya, and A Kadri. Indoor occupancy tracking in smart buildings using passive sniffing of probe requests. In *2016 IEEE International Conference on Communications Workshops (ICC)*, pages 38–44, 2016.
- [2] Julien Freudiger. Short: How Talkative is your Mobile Device? An Experimental Study of Wi-Fi Probe Requests. In *WiSec '15 Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, pages 1–6, New York, New York, USA, 2015. ACM Press.
- [3] Bhupinder Misra. iOS8 MAC Randomization Analyzed! <http://blog.mojonetworks.com/ios8-mac-randomization-analyzed/>, 2014. [Online; accessed 28-September-2016].
- [4] Maarten Van Segbroeck, Andreas Tsiartas, and Shrikanth S. Narayanan. A robust frontend for VAD: Exploiting contextual, discriminative and spectral

cues of human voice. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, number August, pages 704–708, 2013.

- [5] Chenren Xu, Sugang Li, Gang Liu, and Yanyong Zhang. Crowd ++: Unsupervised Speaker Count with Smartphones. In *UbiComp*, pages 43–52, New York, New York, USA, 2013. ACM Press.