s[[1]](#footnote-1)

Elephant8266 – A wireless nut sniffer

G. Pilotto, University of Padova , Computer Science Dept.

*Abstract*—Wireless technology now a days are widely used from many people since it able them to share contents and connect in a easy way to everyday-use technologies. With a few step they can use many ways to communicate with other devices, share file, image, videos and post on social network. The prevalence of the portable devices make the wireless technologies like Wi-Fi Protocol and Bluetooth protocol spreading among the world population. At home, at work, a the supermarkets, bar , restaurant

,hospitals, all private an public places have a Wi-Fi network. The use of the massive use of this technologies every day, make the owner in some cases traceable and expose to risks of privacy violation. In this paper we record Wi-Fi and Bluetooth traffic in a library for 5 hour and then we do an offline data analysis in order to get more informations about the oweners.

A survey is provided to state the knowledge base and the behavior of different users using Wi-Fi and Bluetooth.

*IndexTerms*—Wireless,Wi-Fi, Bluetooth, Security.

# INTRODUCTION

T

HE use of wireless technology spread out through population becoming an interesting topic also in the home domotic secotor thanks to the IoT devices (Internet of Thing) connected to the home Wi-Fi network. Portable device like smartwatch and Fitbit are interfaced with a device (usually a smartphone), and they grab to the owner many different type of data like : time, physical performance, notifications and so on.

Peripherals device enhance the use of the Bluetooth technology[1] , the latest version decrease the working power using Bluetooth low energy (Bluetooth BLE) since due to the

number of application it is involved.   
Due to the number of information and how fast them change, people need to be updated more and more in a reasonable time.

Internet connections allow them to be constantly updated.

The way that the information are promoted and the type of information that we can retrieve online, are very big and some times they can be complex and massively connected together.  
In Facebook, for example, is possible to retrieve, picture, text, videos, gif,360° photo, stories and so on, all in one platform.

In the case of social media user usually are subscribed to more than one social media, visiting them simultaneously.

Wi-Fi or IEEE 802.11is always the preferred choices [2] of the users ,when is provided of course, since most of the time is the faster than a normal 4G phone connections ( this could not be the case, when there are too many people using the Wi-Fi network, and the Wi-Fi network in not well implemented in order to sustain all the users, in this paper we assume WiFi network to be more fast and than 4G).

The worldwide hotspots distribution in 2012 was 750.000, estimate both in homes and business and public spaces [3].

Usually more than one AP that has been connected with one device, leaving it signature inside the device.

Many people using this wireless technology are not aware about this fact , underestimate the leakage of information flowing out from their devices.

Some users left the devices wireless service on even when they are not connected increasing the probability of privacy attacks.

In this paper we would like to exploit data gathered using both Wi-Fi and Bluetooth technologies, listen just at the traffic that devices and APs are exchanging.

Once we have collected enough datas for a profile, the author try to achieve more informations using other online tool.

If Wi-FI MAC ad. And Bluetooth MAC ad. match, cause they are hosted in the same wireless card, we have a link between the device name and the probes request list of a devices.

# Method

We want to record the information leaked out from the devices of the user, in order to track them and gather information without any connection to the network where the devices are connected.

In the active service discovery mode the station (device) is actively search for knows AP’s by probing each channel

broadcasting probe request message for know APs and listening to the probe response message.

The probe request contain the SSID of the target AP. the probe is repeated in each channel until the station successfully associated to an AP.

In the passive service discovery beacon are broadcasted from APs, on different operating channels and stations passively listen to the beacon thorough the set of all the possible channel.

Active mode periodically broadcasting in clear the SSID of the AP in which a device has been previously connected to, including also the MAC address of the Wi-Fi card, that is a globally unique identifier.

Using MAC address of the WLAN network card we can retrieve the vendor of a device that has broadcasted the signal, understanding which type of device the owner used.

Received Signal Strength Indicator (RSSI), is also available through the Wi-Fi, increasing the information for spot out the device.

In order to transform that measure in a meters we should have more than one device and set up a triangulation device system.

Bluetooth broadcast device’s name ,the MAC address, and the Bluetooth RSSI when a device is inquired by another Bluetooth device.

From probe request list, in some cases, it is possible to match SSIDs with physical address location using an online too..

Once we have a name or and address of a person, it is possible to exploit social networks in order to restrict our search owner list, if of course they have a profile active and registered on it.

II.I BACKGROUNDS

## Wi-Fi

A Wi-Fi network describe a client as a station which connects to an access point (AP). A typical AP has a range of 35 meters indoor and 100 meters outdoors.

The Wi-Fi defines three high level wireless frame-types: data, management and control.

We are interested in the management level ,since probe-request and beacon are management frame subtypes.

Stations (client devices) associate to an AP by sending probe-request, which includes the station’s MAC address and the SSIDs of the desired network.

The AP responds with a probe response after which security options and transfer rates are negotiated.

Finally the authentication process occurs.

## Bluetooth

Bluetooth operates at 2.4 GHz range and it is designed for transmitting data over short distance between mobile devices and for creating Personal Area Network.

A Bluetooth network (piconet) consist of one master devices and up to seven slave devices.

The master can send and request data to the slaves, the slaves can only talk to the master.

A slave can ask to a master to become master, with a request.

The master can communicate with one slave at a time , but can rapidly change which it is communicating with in a round-robin fashion.

II.II HARDWARE

The hardware used comes up in two separate boards:

1. *Arduino Yun and Bluetooth dongle:*



Fig. 1. *Arduino Yun and Bluetooth dongle*

Arduino comes up with this new board host in a Wifi

network card already mounted in the board and a OpenWRT [8] distribution, a really famous open source firmware that modify a simple router in a very powerful tool, that permit to filter packets, handle I/O ports of the router, and embed custom scripts, is really common between people that develop IOT project.

Arduino provide also a USB interface, that is different from the microUSB used for programming the core chip.

Through this USB is possible to connect whatever is attachable to a USB and once the right way to access it is found, it is to use.

On the usb port,the author has attached a USB Bluetooth dongle, supporting both Classic and BLE Bluetooth, accessible through the linux distribution (Linino), we can actively scan the nearest (10m) Bluetooth devices, gathering the name and the MAC address of the Bluetooth network card.

The cost of the hardware is around 50$.

1. *ESP8266 + HC05 Bluetooth:*

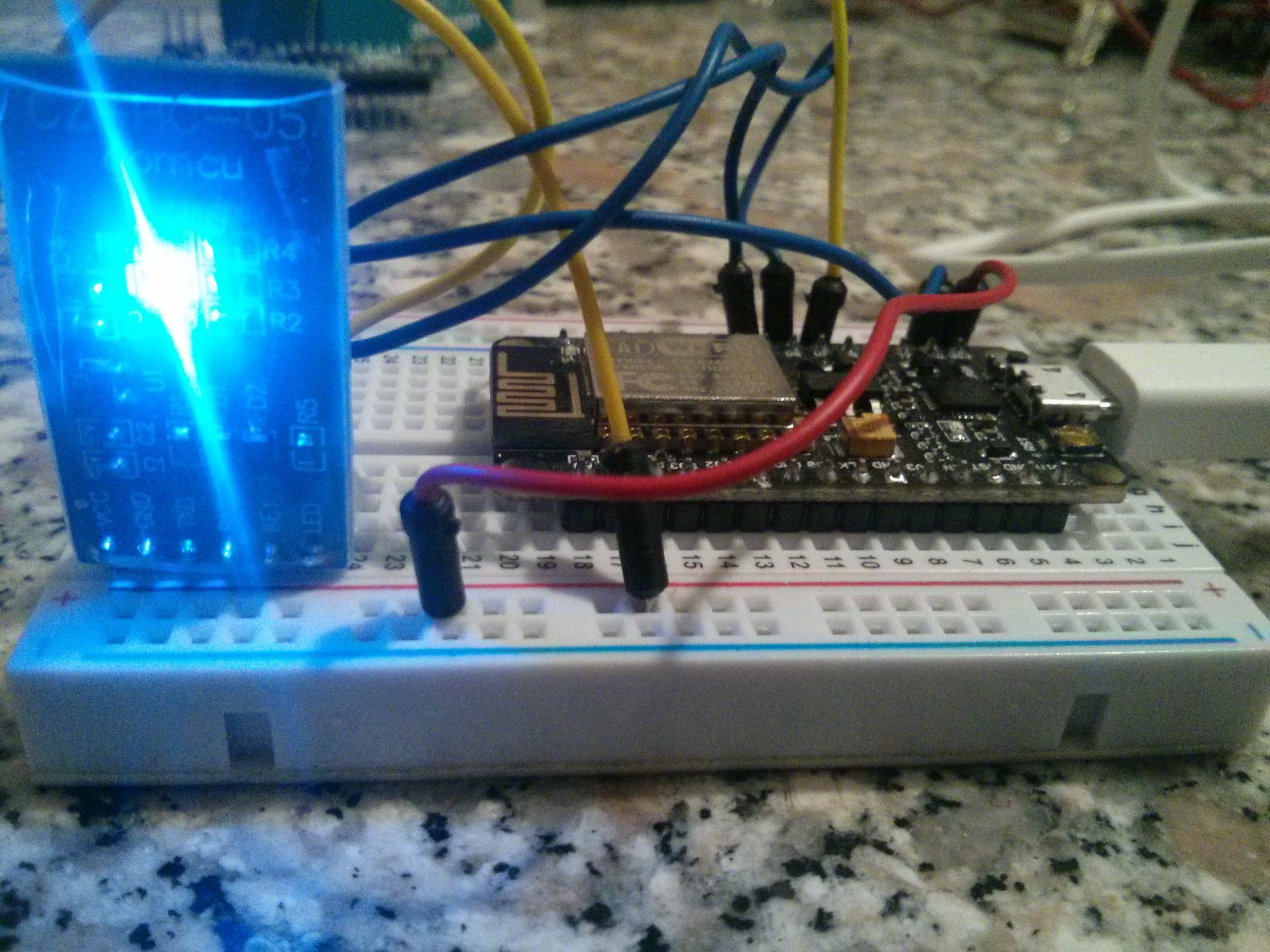


Fig. 2 ESP8266 + HC05

The lower part are the devices sniffed from the wifi, giving the BSSID of the the MAC address the power of the singnal (if the driver provide it) packets, and the probes

This solution is more customizable solutions and it comes up with two different circuits.

The first esp866 is a well-known chip used to connect to Wi-Fi and use his pinout to control sensor and actuators.

Made by Espressif System s.r.l , given out with his SDK, people of the ESP community create a firmware that can handle Arduino code and library, and this fact make the things a lot easier for who is not a C/C++ master.

Thanks to the Arduino code inside we can exploit a software serial in order to attach the Bluetooth HC05.

This Bluetooth become is very popular because itcan provide a COM port easily,

And without implement too much code, actually you have just to connect them in rx,tx pin of the serial to get a cable-serial communication converted to a wireless communication, with a max range of 10m and a cost of 3$.

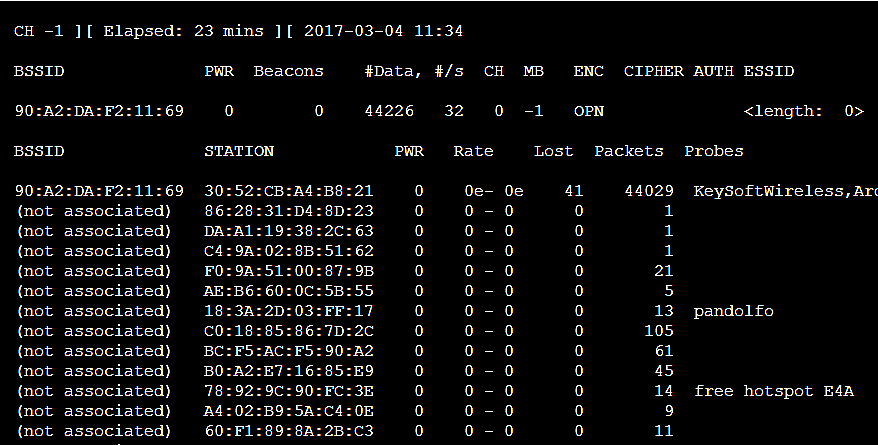
The ESP costs from 10$ to 20$.

II.III SOFTWARE

The hardware needs a couple of implementations software in order to receive data that does not belong to them:

* **Monitor Mode**: The hardware in order to scan devices and read the SSID, BSSID, RSSI signals needs to put the adapter in monitor mode, in order to gather packets that does not belong to them.
* **Channel Switching**: Probes and beacon from APs and devices are prompted along different channels, our hardware needs to pass all channel and ensuring the complete discovery of all devices.

The software used in the Arduino Yun is an aircrack package, made on purpose for OpenWRT distribution named aircrack-ng, and it provide the possibility to put our wireless network adapter in monitor mode and then launching the dump through the adapter it starts to show a table of AP and devices listing the address and the respective connection (Fig4).

Fig. 4. Table of information leakage gathered from Arduino Yun + Bluetooth dongle.

The upper part regards the APs, reveleated from the wireless adapter.

The lower part are the devices sniffed from the wifi, giving the BSSID of the the MAC address the power of the singnal (if the driver provide it) packets, and the probes.

BSSID tell us which connection a devices is associated to, and looking at the MAC address we can distinguish the each different devices.

In probes list we can find a list of connections that a device beacon out, that are possible connection visited by devices.

For the Bluetooth scanning part we need to open another two different sessions, the first will inquiry the Bluetooth asking for the name, the MAC and the class.

The second will listen at the ingoing/outgoing traffic in order to get the RSSI signal strength and the real name of the device, associated to the Bluetooth MAC address (bdaddr in the fig.5).

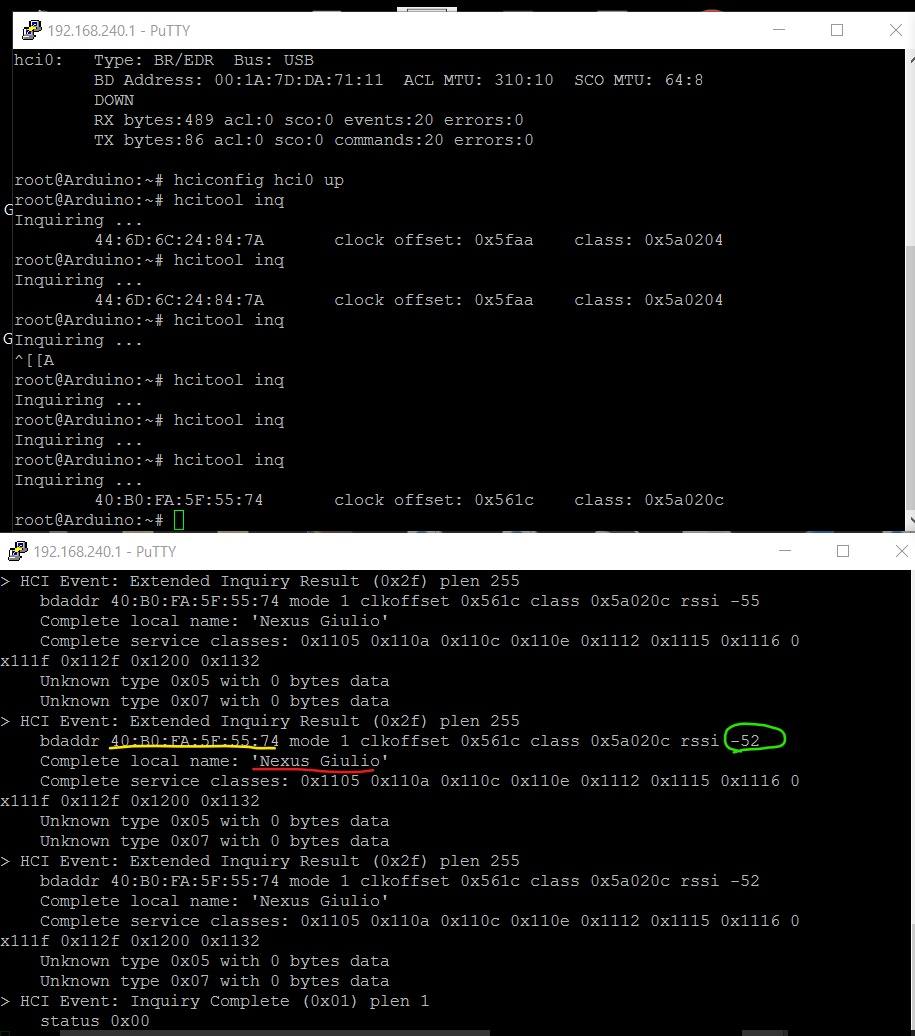


Fig. 5. In the upper session Arduino yun inquired the Bluetooth devices around and retrace the MAC address and the class and the class offset, using command hcitool inq.

The lower session is the command hcidump –a showing all the activity performed by inquiring command revealing the MAC address (in yellow), the RSSI (in green) and the Name of the Bluetooth(in red).

The software for the ESP8266 + HC05 is actually all running in the core processor of ESP8266, where we have to write a special C code in order to implement Wi-Fi scanning and Bluetooth inquiry.

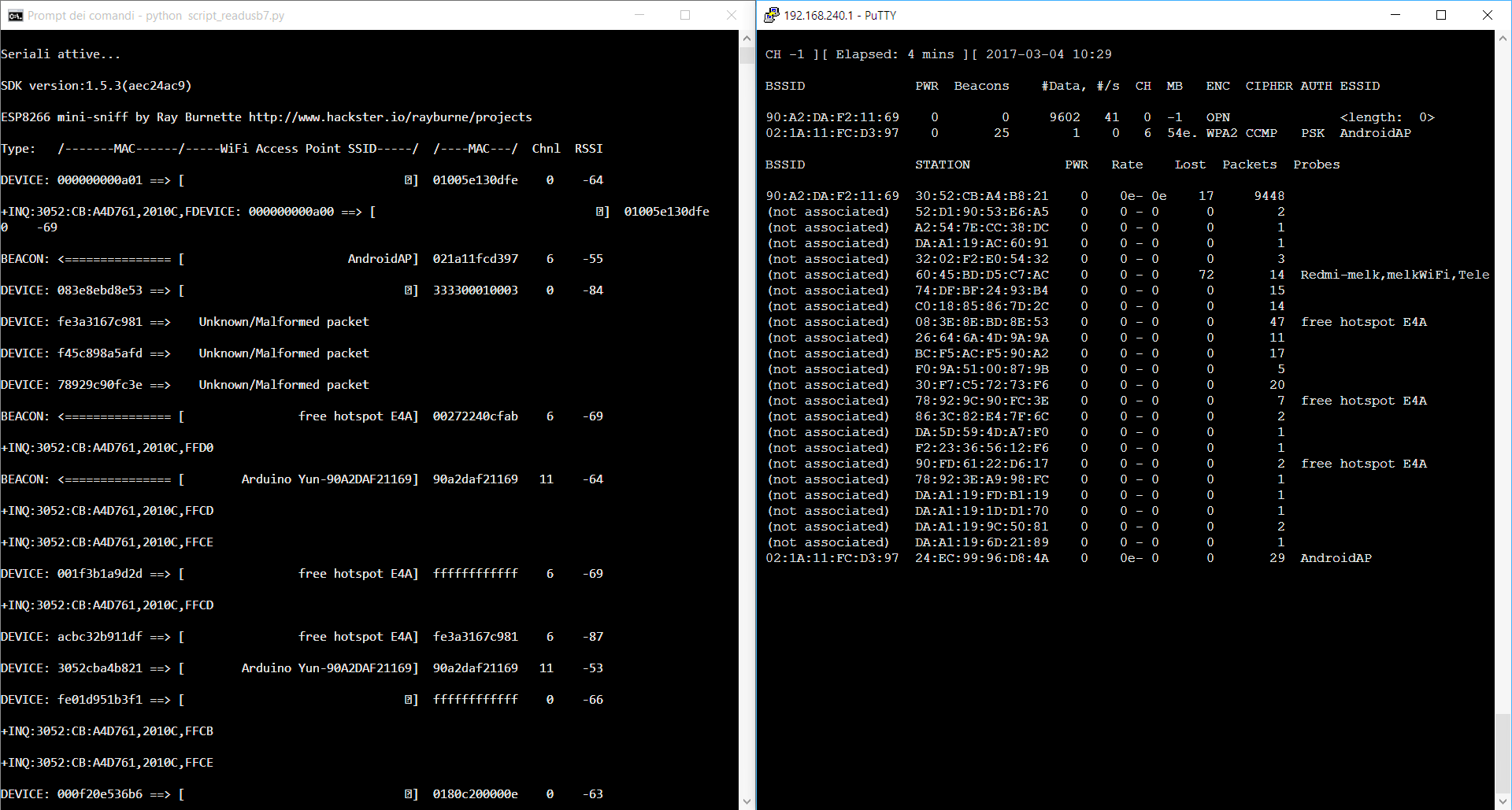
Luckily for the author the community already provide a code for scanning the Wi-Fi network, differencing between beacon coming from AP and devices connected and talking to the network.

The probes list is retrieved colleting the broadcast request sent from devices.

The WiFi RSSI is also collected , the RSSI is a signal indicator of the received power: generally speaking, the higher the numeric number the weaker the signal.

The ESP8266 will talk through a software serial to the HC05 that has been programmed through AT mode in order to inquiry the Bluetooth around, returning the MAC address the RSSI and the class of devices.(Fig. 6).

All the code is available through the GitHub of the author: <https://github.com/giulio93/Elephant8266>.

Fig. 6. The two different output coming out from the two different solutions,

On the left the ESP8266+HC05 and on the right the Arduino Yun + Bluetooth Dongle.

Packets without destination are classified with Unknown/Malformed packet.

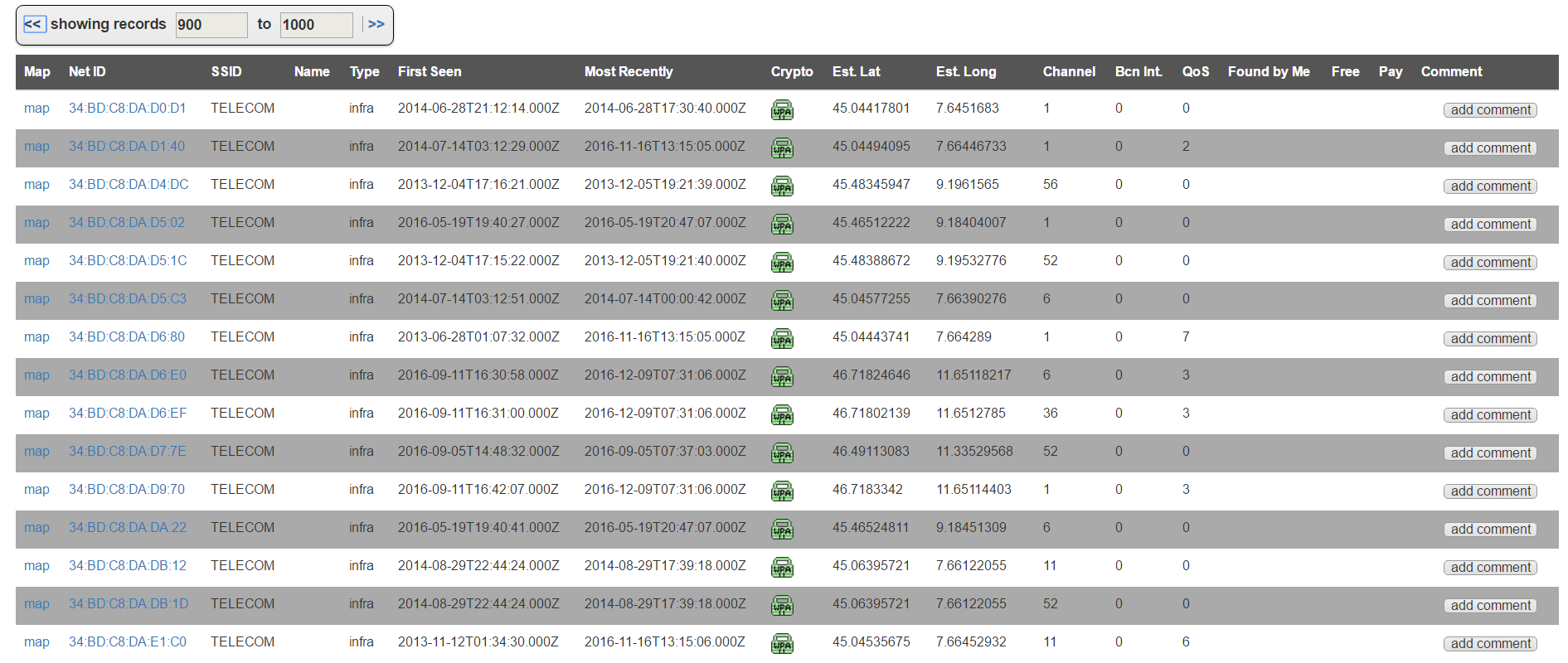
II.IIII PROFILING THE OWNER

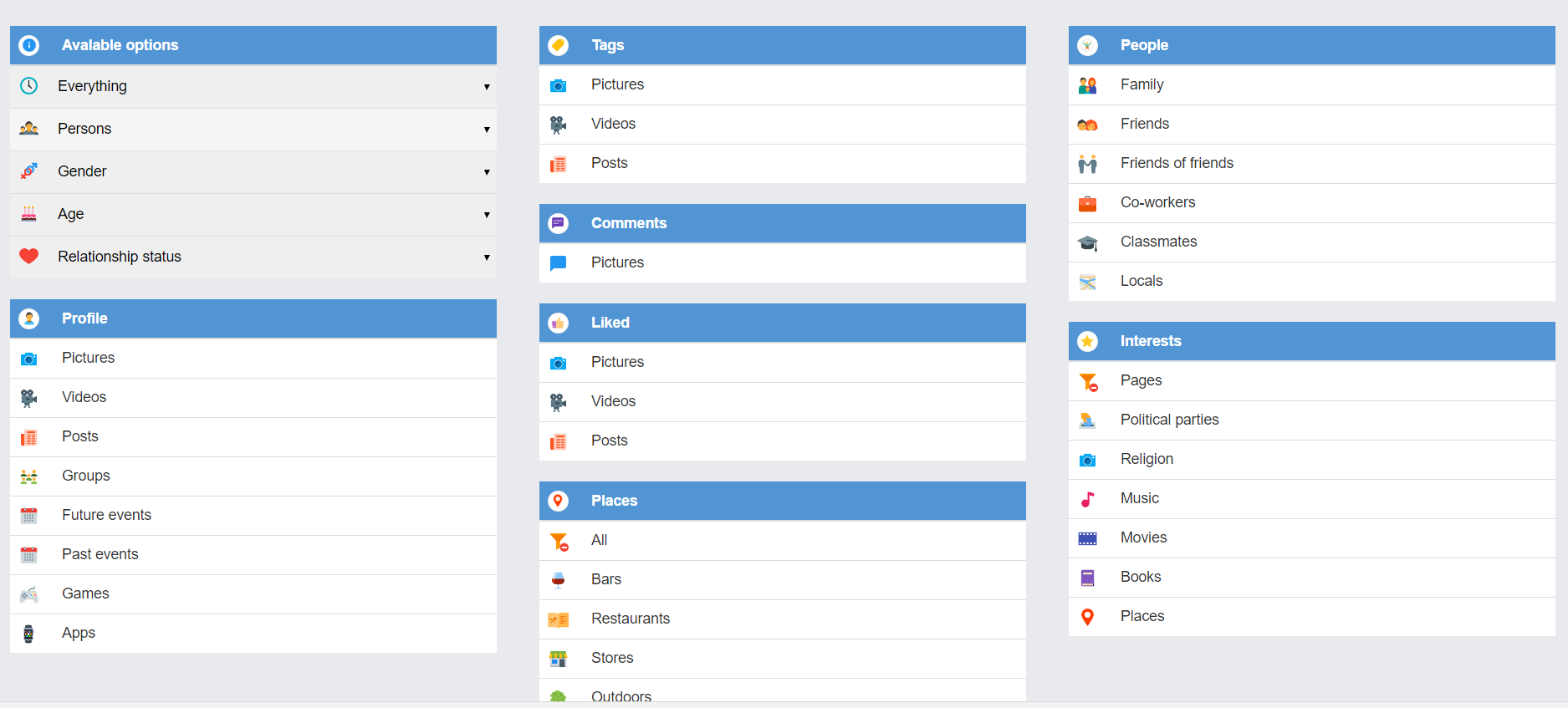
The information collected at this point are based at the low level, unique MAC addresses, list of Wi-Fi connections, Name of the devices, strength of the signal.

Using this low level information, the author want to have more meaningful material in order to make high level inference about the users.

This can be done using a couple of online tools like:

* **WiGLE.net**[4]: is a physical map of all access point sniffed and uploaded from volunteers, it can help us to find more information about probes. Feeding the tool with the access point MAC and the SSID, WiGLE retrieve the geolocation position , MAC address, SSID, First seen and most recently seen, the type of Cryptography algorithm using, and the channel where it was working when it was spotted out. We can lookup to the interested SSID both directly on the map or using a research filter where we set the position range other configuration parameters in order to give us a filter list of access point (Fig.7). There are many python script online that can help us to manage and automatize the process in order to parse a big cluster of data and find what we are looking for. With a free account search are limited.
* **Facebook + StalkScan.com[5]:** Facebook has recorder 1.76 monthly active users in December 31,2016 [6], compared to the world population, 7.5 billion in 2017, Facebook earn almost 1/5 of the total world populations. Facebook is a powerful tool that make inference about people and provide profile search filter in order to find in a efficient way: people, places and many other social events.Information like: name, type of deivce) geo-located addresses (i.e. work place or university attended) is possible to build a list of possible profile matching our low level data. In order to infer a better profile list StalkScan provide us all the information that the privacy of the profile allow to see(Fig. 8), for better saying, it just give us the right question to prompt in the facebook environment and show us information difficult to find using the normal Facebook interface.

Fig. 7. Output of WiGLE.net looking for all hotspot containing the word Telecom, with a range 45.6716942 to 45.6816942 Lat. and 11.928238800000031 to 11.93880000031, around Castelfranco Veneto, where the registrations for this paper were made.

Fig. 8. Output of StalkScan.com looking at the data retrieved from the author Facebook profile, all the section create a new facebook tab where the interested objects are shown.

# RESULTS

The recorded a session of 5 hours no stop from 14:00 to 19:00, in the library of Castelfranco Veneto,Treviso, Italy.

The row data area available in the source code repository.

The author process the data offline after the registration period, and the graph drew from analyzed data are available in the Appendix B.

A questionnaire was also provide to different colleagues, students of different universities, high school, and workers.

The author tried understood the behavior of the users, using wireless technologies in particular Wifi and Bluetooth, and see if there is some match with the results find from the recorded session.

The questions and the results are available in the Appendix A.

Data collected from the ESP8266 solution discovered 79 devices and 8 APs, the Arduino Yun solutions find out: 202 devices and 3 APs.

The is a lot more than what the library can handle, in terms of users , but the author suppose it could be traffic coming from outside the library and library personnel’s traffic.

Taking in account the graphic B1, different AP are shown in the X axes the AP’s powers collected with the ESP8266 ( without the negative symbol) are plotted.

More power mean less signal, so inferred that low signal can be related to more distance, or some obstacle in the middle like walls.

The last two bends get a better signal and they are actually connections to the Arduino Yun solutions and the author hotspot used to surf on net.

In the Fig.B2 a distance fashion-like plot give a better idea of distribution of the AP in terms of distance.

The author always refer at the distance not as a real distance but as a measure of the signal quality.

Devices distribution is showed in Fig.B3, as a reference we take the computer (dotted red line) of the author, that is just next to the ESP8266 solution,

so all the other devices recorded are more distant or at least at the same distance, as the graph show.

The Arduino Yun solution did not provide any information about the power due to the hardware Wi-Fi card characteristics.

Bluetooth scanning give us just a couple of results, one is the phone of the author and the second is a computer that was around there (Fig. 9), that show up just for a few time, but was impossible to link with any retrived Wi-Fi MAC address.

root@Arduino:~# hcitool scan

Scanning ...

40:B0:FA:5F:55:74 Nexus Giulio

30:52:CB:A4:D7:61 DESKTOP-T9OCP9N

Fig 9- Arduino scanning a bluetooth.

Using the MAC address collected from the devices in Fig.B4 the author plot all the vendors retrieved from the initial octet of the MAC address (XX:XX:XX), Google Inc. , (Android devices) is dominating other vendors , in the second line there is Apple Inc. and then all the other follows, this make sense if the readers think about what type of devices are more purchased in the last years.

Arduino Yun solutions provide many probes for a reasonable amount of owner: in the Fig.B5 the author plot the number of time a probe combinations (i.e. “Telecom-123, math.unipd, Home-Dlink”) is requested by different devices.

The outcome are 30 differentiable combinations, the blu big slice of the pie cake is the library “free hotspot E4A” , the second orange slice is the eduroam network, eduroam (education roaming) is the secure, world-wide roaming access service developed for the international research and education community.

And this result make sense too since, most of the people that going to the library are supposed to be students, especially in the afternoon.

Let’s inspect the smaller pieces, remember that we are looking for data that can distinguish and clamp our research field in order to get more information about the user, so less people sharing the same probes vector, mean singularity and more large are probes vectors more information we can have about the owner.

Focusing on those informations Fig. B5 is plotted, and then it easy to get the MAC address with more Probes.

In those profile the author found like:

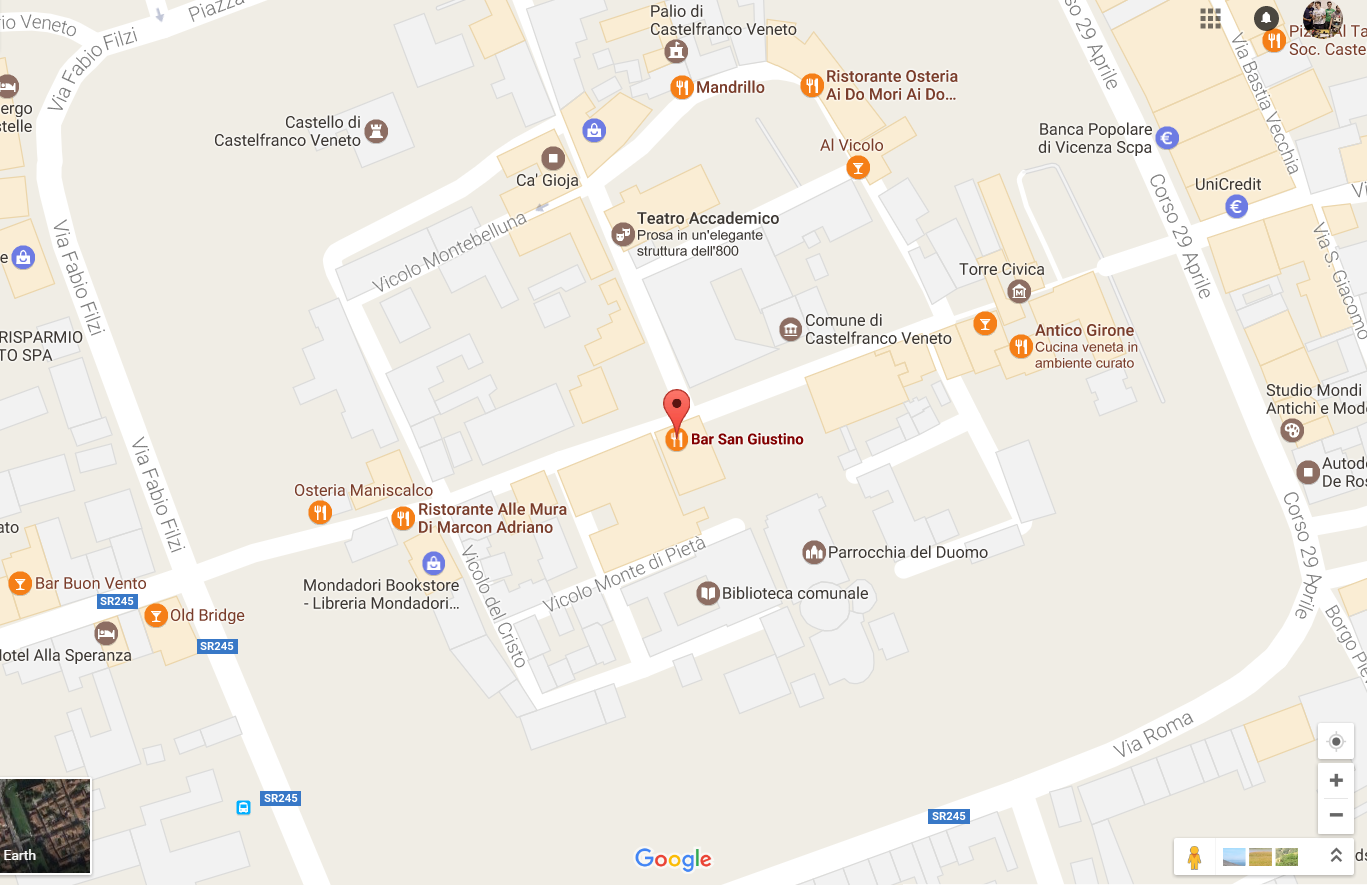
bar name SSID (like: “EATANDGO-BAR”, “BISTRO SAN GIUSTINO”) (Fig 10).  


Fig. 10

SSID discoverable using Wigle.net, like “CastelfrancoV.toWifi” (Fig. 11).

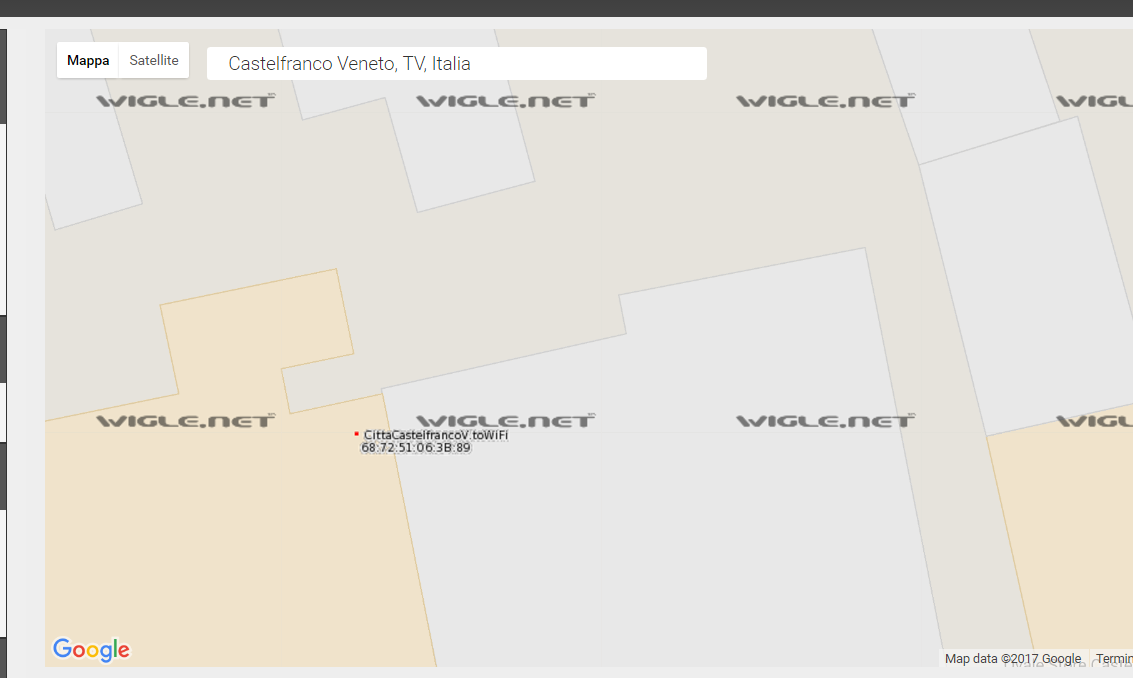


Fig 11.

SSID “OROBEACH” about beach resort in Jesolo: <http://www.jesolospiagge.it/en/oro-beach.html>.

There was also a couple of Name inside the SSID list “Guglielmo” and “Iphone di Piero”.

“Guglielmo” SSID get another pair of SSID name, that’s unfortunately searching in WiGLE.net does not give any result, cause of the too old mapping in WiGLE.net.

The other one “Iphone di Piero” is much more interesting, even if there is no mapping in WiGELE we could make some inference about it, since it appear also the SSID “eduroam” meaning that he could be a student.

And he has to be involved in Castelfranco Veneto library in some way, if it is not there just for a bad time schedule stop.

So searching on Facebook about Piero, Castelfranco Veneto, students, the author get a list of possible people to infer.  
Using stalk scan the author find a list of 3 people fitting the parameters, that could be the owner that are looking for.

# CONCLUSION

The aim of this paper is try to find a way to get informations about the owner looking at the WiFi Bluetooth leakage.

The ESP8266 solution did not give us a good feedback about probe request it notify a few broadcast search by devices for the network “ free hotspot E4A “ that is actually the Wi-Fi free network system of the library.

From the Bluetooth prospective the ESP8266 solution collected some record of a device around the table of the author ,giving the MAC address, the class and the RSSI, more studies are needed about HC05 since it is not clear if the antenna has a really short range, or during the collect session there was not any Bluetooth devices.

The cause of the poor devices collected, could be releated to the low performance of the WiFi Antenna in the esp, and also the difference in the firmware that run through Arduino library talking to the core module, since many threads are going on that is not an easy procedure to handle.

The Arduino Yun performed quite well in the Wi-Fi recorder session giving us quite good informative material, aircarck-ng work well thanks to various and large support given by the community unlikely fot the ESP-code, made few people and modify by the author.

Blutooth in Arduino needs more studies, about the capabilities of inquiry devices in terms of range (meters) and pattern searching. Maube

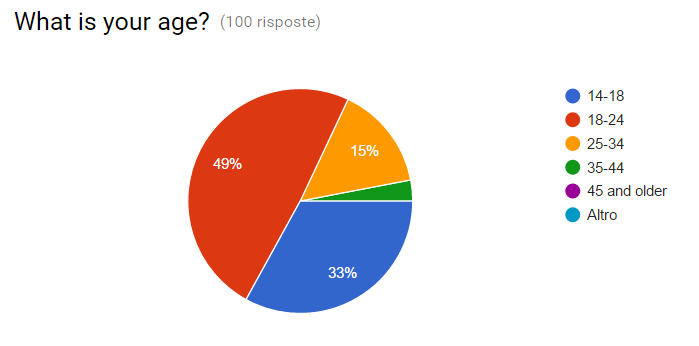
References

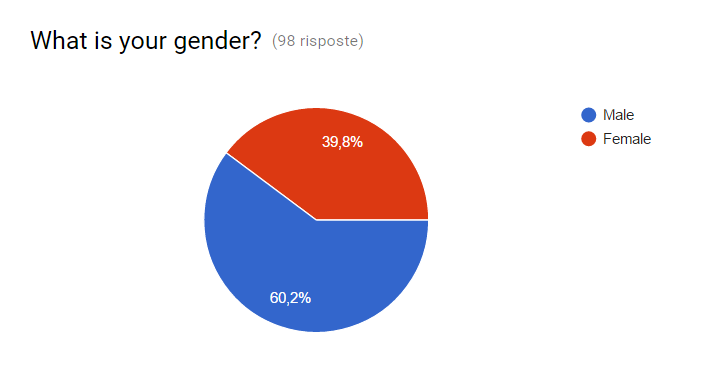
*Basic format for books:*

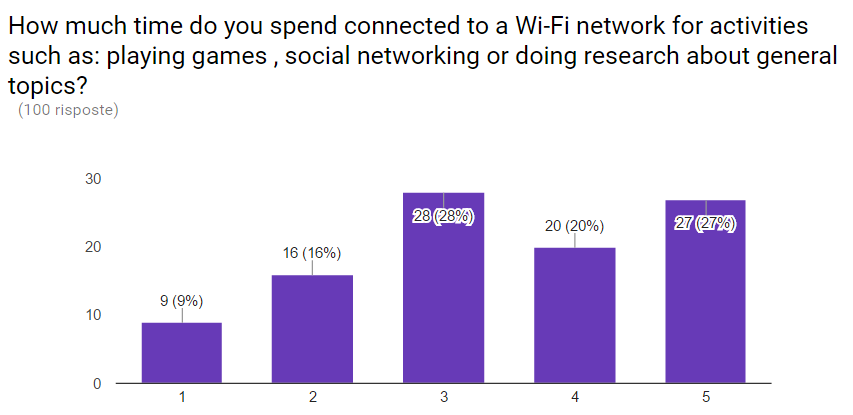
1. M. Kazantdis “Bluetooth: An Enable for Personal Area Networking,” Ericson Research ,IEEE Network ,September 2001
2. B.Scholbrock “Characterizing and Modeling User Activity on Smartphones: Summary “ Dept. of Electrical Engineering and Computer Science Northwestern University, Evanston, IL, USA
3. M. Cunche, “I know who you will meet this evening! Linking wireless devices using Wi-Fi probe requests,” *HAL,* 2 November 2012.
4. WIGLE.NET. Available: <https://www.wigle.net/>
5. STALKSCAN.COM. Available: <https://stalkscan.com/>
6. Facebook, Company informations. Available: <http://newsroom.fb.com/company-info/>
7. Worldometers . Available:<http://www.worldometers.info/world-population/>
8. Claudio E. Palazzi , “An OpenWRT solution for future wireless homes” , Universita degli Studi di Padova, Padova, Italy, 23 September 2010
9. Musical toothbrush with adjustable neck and mirror, by L.M.R. Brooks. (1992, May 19). *Patent D 326 189*

APPENDIX A – SURVEY RESULTS (100 INTERVIEWS)

(<https://docs.google.com/forms/d/1QfvNkMDf2Nzz64BTZkqTtgdU2ADnJA6jWcI1ByjgjVU/viewanalytics#start=openform>)

Fig. A1.

Fig. A2.

Fig. A3.

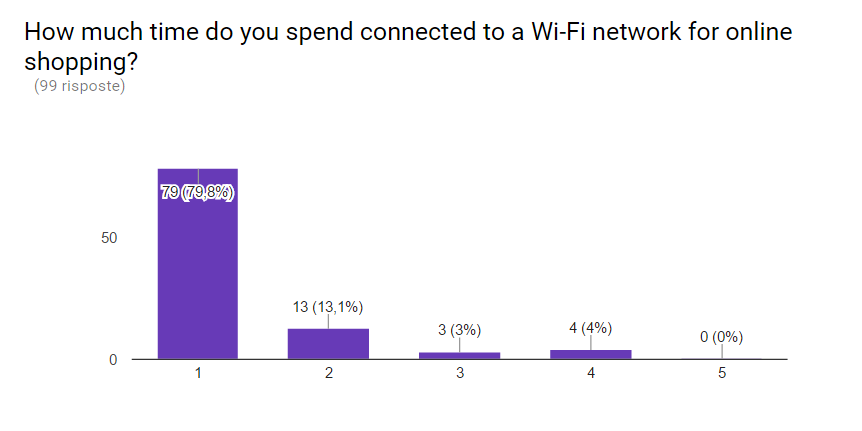


Fig. A4.

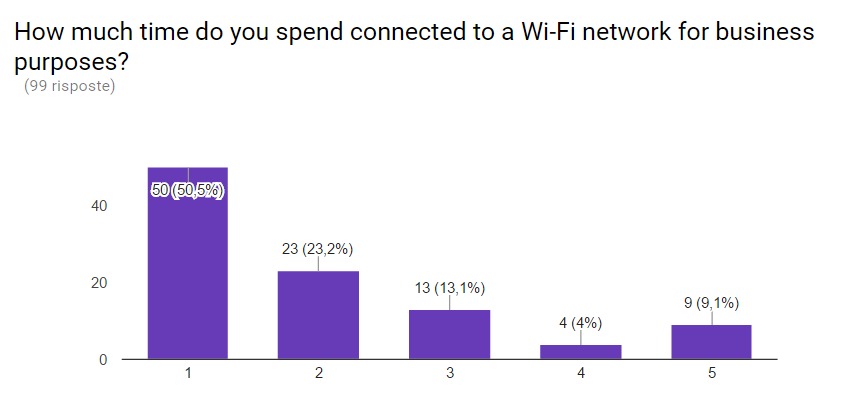


Fig. A5.



Fig. A6.

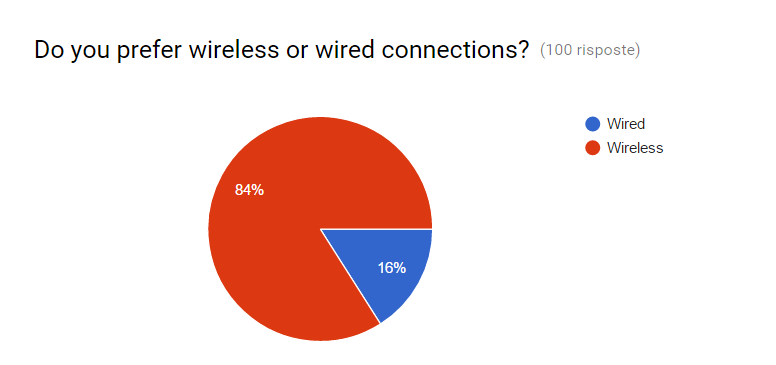


Fig. A7.

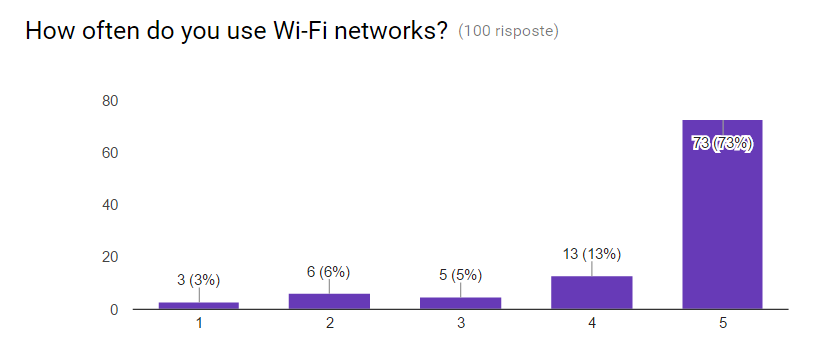


Fig. A8.

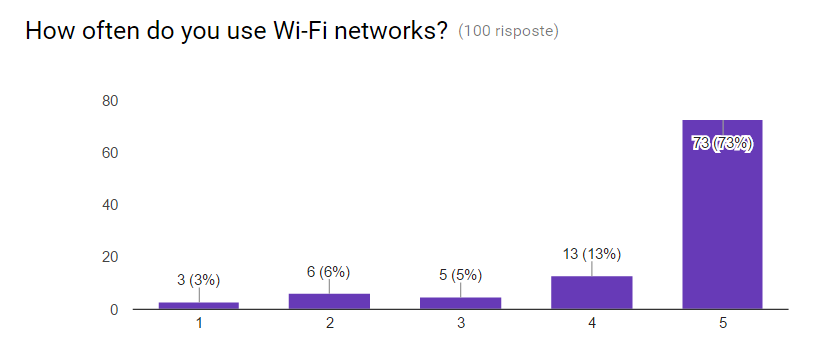


Fig. A9.

APPENDIX A – SURVEY RESULTS (100 INTERVIEWS)

(<https://docs.google.com/forms/d/1QfvNkMDf2Nzz64BTZkqTtgdU2ADnJA6jWcI1ByjgjVU/viewanalytics#start=openform>)

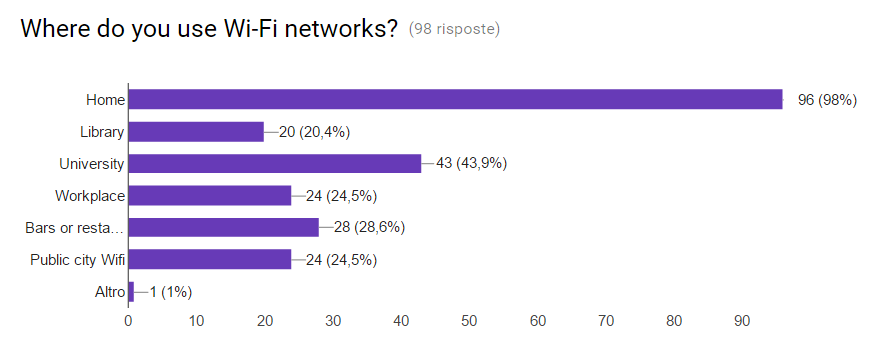
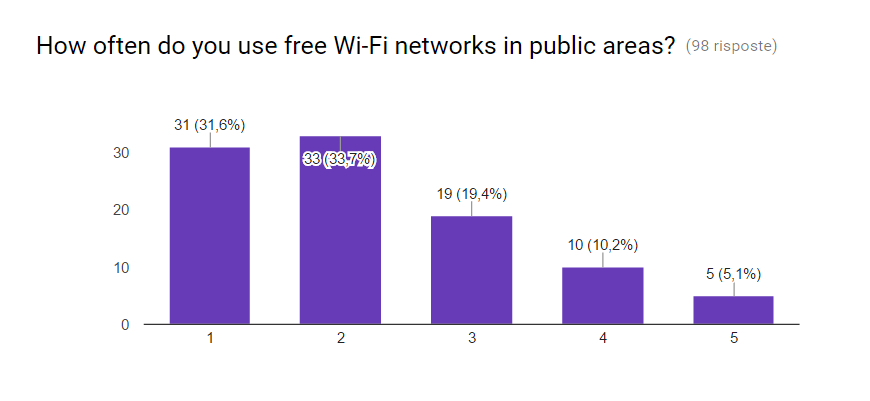


Fig. A10.

Fig. A11.

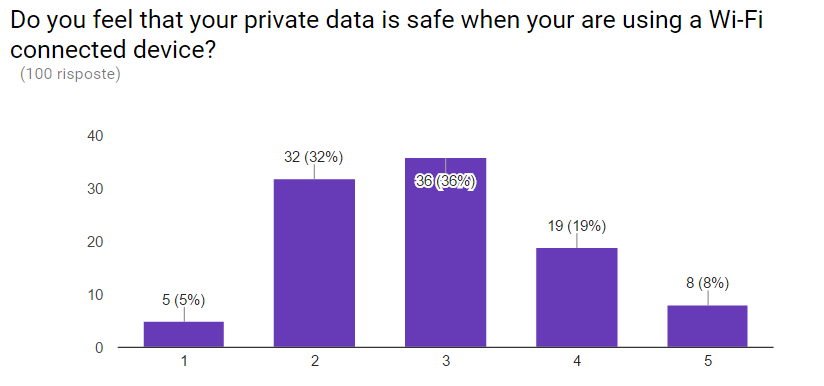


Fig. A12

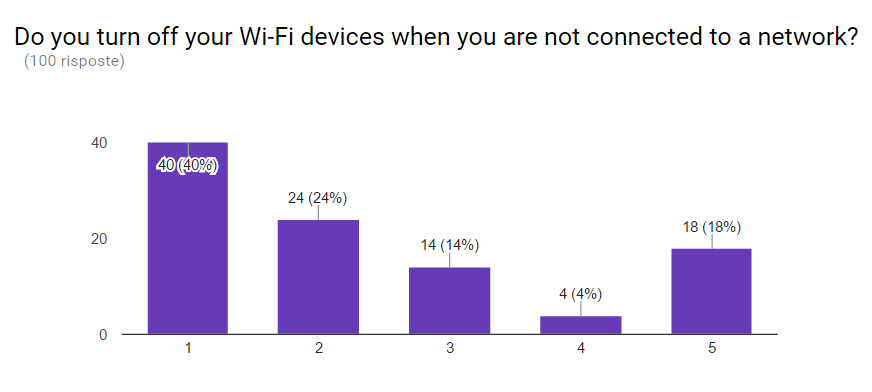


Fig. A13

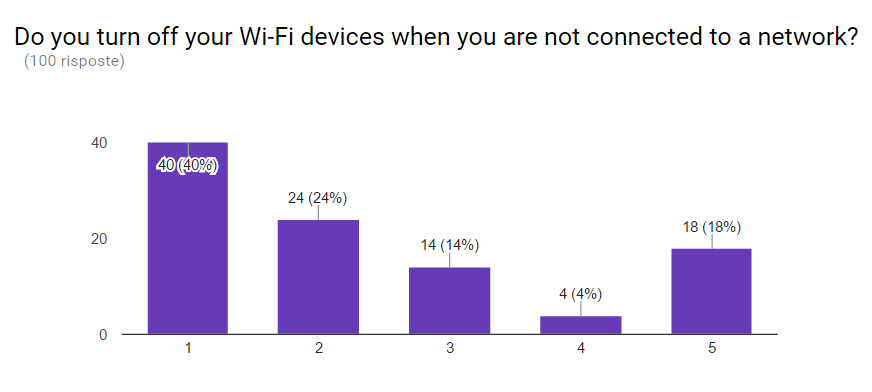


Fig. A14

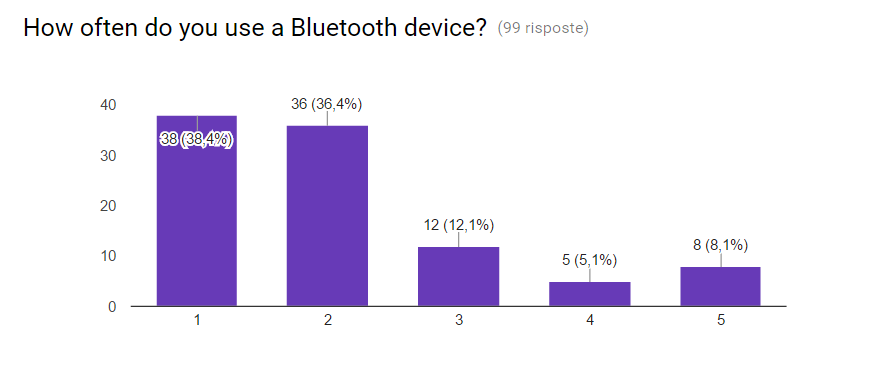


Fig. A15

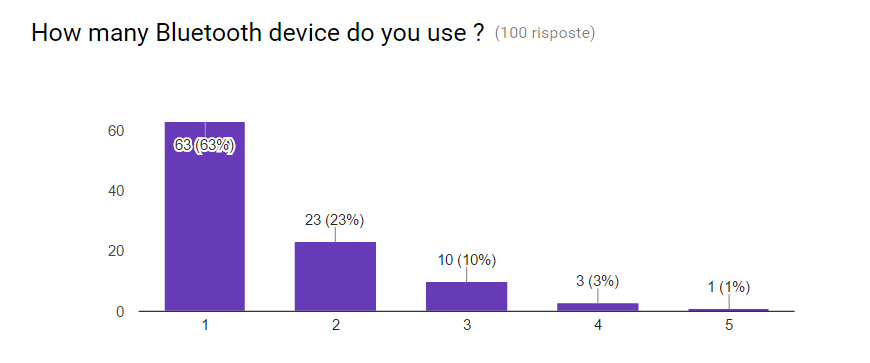


Fig. A16

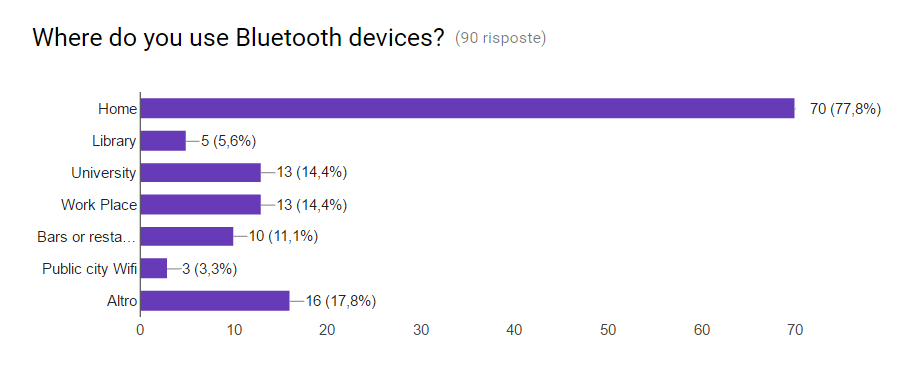


Fig. A17.

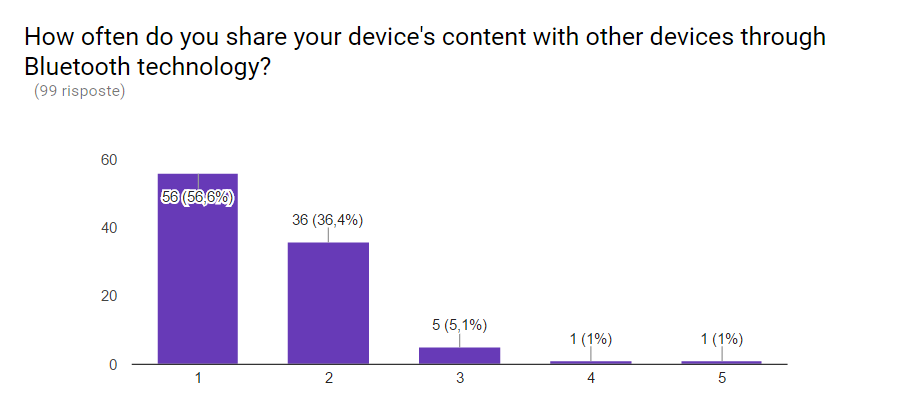


Fig. A18.

APPENDIX A – SURVEY RESULTS (100 INTERVIEWS)

(<https://docs.google.com/forms/d/1QfvNkMDf2Nzz64BTZkqTtgdU2ADnJA6jWcI1ByjgjVU/viewanalytics#start=openform>)

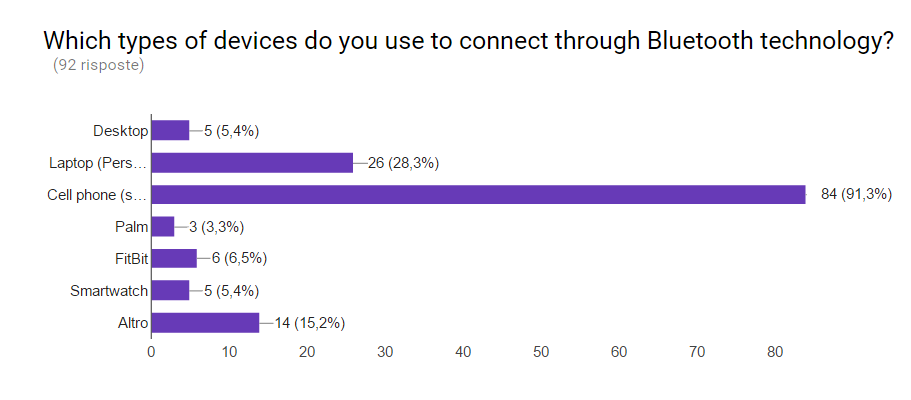


Fig. A19

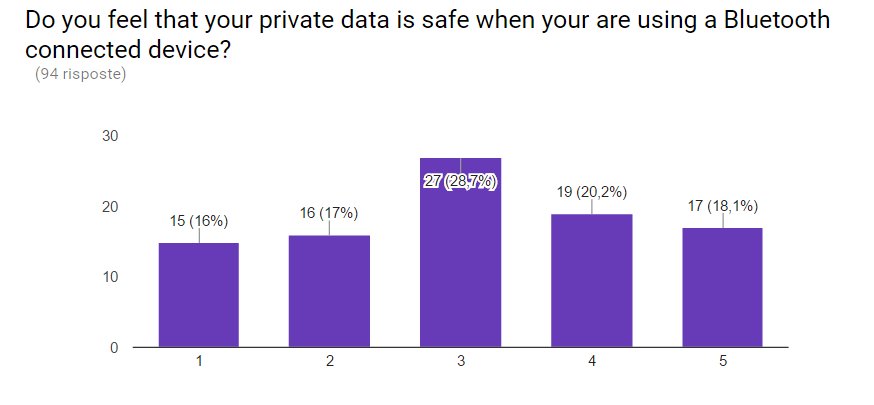


Fig. A20.

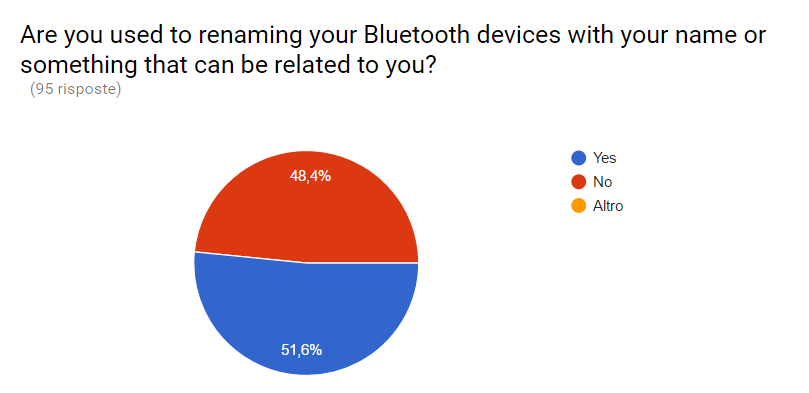


Fig. A21.

APPENDIX B – COLLECTED DATA



Fig.B1

Fig. B2



Fig. B3



Fig.B4



Fig. B5



Fig.B6



Fig.B7

1. 01/03/2017 “This work was supported in part by Mr. Conti , lecturer at the University of Padova ”. Author: Giulio Pilotto student of University of Padova(e-mail:pilotto.giulio@gmail.com). [↑](#footnote-ref-1)