



UNIVERSITÀ DEGLI STUDI DI GENOVA

SCUOLA POLITECNICA

DIBRIS

**Dipartimento di Informatica, Bioingegneria,
Robotica e Ingegneria dei Sistemi**

**TESI DI LAUREA MAGISTRALE IN
INGEGNERIA INFORMATICA**

**Monitoraggio di flussi e densità pedonale presso aree
congestionabili mediante sensori innovativi**

Relatore:

Chiar.^{mo} Prof. Roberto Sacile

Correlatore:

Dott. Luca Zero

Candidato:
Gianluca Ceccoli

Anno Accademico 2018/2019



UNIVERSITÀ DEGLI STUDI DI GENOVA
SCUOLA POLITECNICA
DIBRIS

**Dipartimento di Informatica, Bioingegneria,
Robotica e Ingegneria dei Sistemi**

**TESI DI LAUREA MAGISTRALE IN
INGEGNERIA INFORMATICA**

**Monitoraggio di flussi e densità pedonale presso aree
congestionabili mediante sensori innovativi**

Relatore:

Chiar.ºmo Prof. Roberto Sacile

Correlatore:

Dott. Luca Zero

Candidato:

Gianluca Ceccoli

Anno Accademico 2018/2019

Monitoraggio di flussi e densità pedonale presso aree congestionabili mediante sensori innovativi

Sommario

La Crowd Analysis sta acquisendo maggiore importanza grazie alla sua capacità di prevenire situazioni pericolose dovute ai sempre più frequenti raduni di persone in ambienti che non prevedono un tale sovraffollamento. Al contempo, il boom tecnologico dei dispositivi mobili consente metodi di monitoraggio alternativi alle classiche videocamere. In questo elaborato verrà presentato lo sviluppo e l'installazione di un sistema di monitoraggio basato su sensori Wi-Fi presso il Parco Nazionale delle Cinque Terre, in quanto area fortemente congestionabile. Sulla base del conteggio degli indirizzi MAC rilevati dai sensori, il sistema è in grado di classificare il traffico in tre categorie di viabilità: *libera*, *trafficata* e *a rischio*. La realizzazione del progetto va a mettere in discussione la randomizzazione dell'indirizzo MAC, comune misura per la protezione della privacy adottata dai più noti sistemi operativi per dispositivi mobili, misura che dovrebbe impedirne il tracking. La soluzione sviluppata per contrastare la randomizzazione si basa sulla costruzione di chiavi derivate dalle caratteristiche fisiche dei dispositivi per discriminare gli indirizzi MAC rilevati e sul controllo di aggregazione tra indirizzi appartenenti alla stessa categoria. In un futuro prossimo questo sistema potrebbe essere in grado di ricoprire il ruolo di strumento di predizione a disposizione dei turisti, analogamente alle previsioni meteo.

Monitoring pedestrian flow and density at congestable areas using innovative sensors

Abstract

Crowd Analysis is gaining relevance thanks to its ability to prevent potentially dangerous situations due to overcrowding in environments that do not have safety plans in place. The technological boom in mobile devices allows alternative monitoring methods to classic video cameras. This paper will illustrate the development and installation of a monitoring system that uses Wi-Fi sensors in the highly congested area of the Cinque Terre National Park.

Based on the count of the MAC addresses detected by the sensors, the system can classify three categories to measure traffic: *free*, *busy* and *at risk*.

The realization of the project discredits the randomization of the MAC address, a common measure for privacy protection adopted by most mobile devices' operating systems, as a measure that should prevent tracking.

The solution developed to counter randomization is based on the construction of signatures derived by the physical characteristics of the devices, in order to discriminate the detected MAC address and on the aggregation check between addresses belonging to the same category.

In the future, this system should be able to supply the tourists with real-time information just like a weather forecast app.

Ringraziamenti

Prima di lasciar spazio alla trattazione, vorrei ringraziare coloro che mi hanno permesso di raggiungere questo traguardo.

In primis ringrazio il Prof. Roberto Sacile che mi ha guidato nella stesura della tesi e senza il quale questo elaborato non avrebbe visto la luce. Ringrazio anche il mio correlatore Dott. Luca Zero e i suoi colleghi del laboratorio di automazione per avermi assistito durante la parte tecnica. Non meno prezioso è stato l'aiuto di Emanuele Raso e Daniele Moggia del Parco Nazionale delle Cinque Terre.

Ringrazio i miei familiari, in particolare i miei genitori. Grazie per il sostegno morale ed economico che mi ha permesso di arrivare dove sono ora e grazie per i vostri saggi consigli durante la mia permanenza all'università.

Grazie a Chiara Leoni e Ulisse Quartucci per aver contribuito direttamente a questo lavoro e ad Andrea Maffia, Cristiano Contu e Mattia Marchetti per l'interessamento al progetto.

Per ultimi ma non meno importanti, i miei compagni e i miei amici. Persone fantastiche con cui ho condiviso il banco per anni tra fatiche, sconforti e momenti di gioia, fino al raggiungimento dei traguardi. Grazie per aver alleggerito le sessioni d'esame che sembravano durare in eterno e per l'aiuto pratico e morale che ci siamo scambiati. Questa mia laurea è sicuramente in parte anche la vostra.

Un sentito grazie a tutti.

Gianluca

Indice

Sommario.....	I
Abstract.....	II
Ringraziamenti.....	III
Indice	IV
1. Introduzione.....	1
2. Stato dell'arte.....	3
2.1. Tecnologie utilizzate.....	4
2.1.1. Videocamere	5
2.1.2. Infrarossi e laser	6
2.1.3. Sniffer.....	8
2.2. Progetti di monitoraggio	9
2.2.1. SAIL Amsterdam (2015).....	10
2.2.2. Pellegrinaggio Hajj (2016).....	11
2.2.3. Il caso Venezia (2018)	13
3. Le connessioni Wi-Fi	14
3.1. MAC Address	15
3.2. Randomizzazione.....	17
3.2.1. Android	18
3.2.2. Apple	20
3.2.3. Windows	23
3.2.4. Linux	24
4. Tecnologia proposta	25
4.1. Architettura del sistema	25
4.2. Algoritmo.....	27

4.2.1.	Fingerprinting.....	29
4.2.2.	Aggregazione	30
4.3.	Caratteristiche dello strumento	32
4.3.1.	Modello del problema	32
4.3.2.	Accuratezza nell'aggregazione	33
4.3.3.	Penetration rate.....	34
4.3.4.	Portata massima	34
4.4.	Autonomia delle stazioni	35
4.4.1.	RAM.....	35
4.4.2.	Crontab.....	36
4.4.3.	Monitor mode.....	38
4.5.	Database.....	43
4.6.	Applicazione Web.....	47
5.	Applicazione al caso di studio	50
5.1.	Le Cinque Terre	50
5.2.	Collocamento degli sniffer.....	52
5.2.1.	Monterosso.....	53
5.2.2.	Vernazza.....	54
5.2.3.	Corniglia.....	55
5.2.4.	Manarola	56
5.2.5.	Riomaggiore.....	57
5.3.	Applicazione Web GIS	59
5.4.	Test di monitoraggio	62
6.	Considerazioni finali	63
7.	Sviluppi futuri.....	64
	Bibliografia	65

1. Introduzione

La crescita sempre più rapida della popolazione nel mondo e la conseguente urbanizzazione hanno portato a un incremento nel numero di attività come eventi sportivi, raduni politici, dimostrazioni pubbliche ecc. che risultano in più frequenti raduni di folle. In questi scenari, la Crowd Analysis sta ricevendo crescente attenzione per poter assicurare miglior gestione della folla o riconoscimento automatico di situazioni potenzialmente pericolose. Il rilevamento di situazioni anormali può migliorare significativamente l'efficienza della sorveglianza, risparmiando l'attenzione umana per i casi più importanti, previa segnalazione di un sistema.

Il costo sempre più accessibile dei sistemi di videosorveglianza ha come conseguenza che i metodi tradizionali per il monitoraggio del traffico pedonale facciano affidamento su tecnologie basate su acquisizione video o contapersone (fotocellule o barriere laser) che tuttavia sollevano questioni estetiche o di sicurezza per i dispositivi stessi, oltre a quelle legate alla privacy delle immagini registrate. Come ogni altro problema di Computer Vision, la Crowd Analysis comporta molte sfide da affrontare come occlusioni, distribuzione non uniforme delle persone, illuminazione variabile o prospettive sfavorevoli che rendono l'analisi particolarmente difficile.

D'altra parte, l'ammodernamento dei dispositivi mobili e la propensione dell'uomo a possederne uno, ha permesso che il monitoraggio dei cellulari possa competere nel campo della stima dei movimenti del traffico pedonale. Se il sistema di rilevazione è dotato di ricevitori Wi-Fi, è possibile catturare l'origine-destinazione (O-D) di un pedone, il suo travel time o il wait time, ricavando informazioni sul flusso di un sottoinsieme di pedoni, possessori di devices con modulo Wi-Fi acceso.

Lo scopo principale di questo elaborato è quello di studiare la fattibilità di implementazione di un sistema di monitoraggio di flussi impiegando degli sniffer Wi-Fi i cui dati saranno integrati con quelli disponibili da altri mezzi di monitoraggio preesistenti, per consentire, in tempo reale o semi-reale, la stima dell'afflusso pedonale in aree particolarmente predisposte a congestimenti come i borghi delle Cinque Terre. Lo scopo secondario è quello costruire un dataset in modo da essere in grado in futuro di effettuare predizioni circa l'andamento dell'affluenza. Per realizzare il progetto è necessario studiare lo stato dell'arte attuale nel campo della randomizzazione del MAC address per garantire

quanto più possibile la veridicità dei dati raccolti dai sensori e al contempo mettere alla prova la sicurezza riguardo la privacy che i produttori di hardware e software vorrebbero garantire impiegando questa tecnica.

L’idea è quella di installare più stazioni Wi-Fi, inizialmente in posizioni strategiche presso i cinque borghi per poi valutare una possibile specializzazione presso uno dei paesi. Una prima fase consiste nella definizione del network di sensori atti al monitoraggio dei flussi turistici. In secondo luogo, è prevista la progettazione e realizzazione di una base di dati per gestire in modo centralizzato i dati provenienti dalla rete di sensori. Come ultimo passo, sarà necessario rendere i dati disponibili in forma grafica agli utenti accreditati mediante un ambiente WebGIS.

Nel secondo capitolo verrà discusso lo stato dell’arte riguardo le tecnologie impiegate nel monitoraggio e riguardo i passati esperimenti di stima del flusso pedonale in grandi eventi. Nel terzo capitolo verrò approfondito il discorso del Wi-Fi, cosa è il MAC address e cosa si intende per randomizzazione. Nel quarto capitolo si scenderà nel dettaglio della realizzazione del sistema proposto, verrà discussa l’architettura, i limiti del sistema di misurazione e le sue capacità. Il capitolo numero cinque tratterà l’applicazione al caso di studio della tecnologia proposta, con una panoramica delle Cinque Terre, dove sono stati posizionati i sensori e la visualizzazione dei dati sull’applicazione web. Il sesto capitolo sarà dedicato alle conclusioni finali e il settimo parlerà delle future applicazioni possibili.

2. Stato dell'arte

La Crowd Analysis può essere effettuata su due livelli: quello microscopico e quello macroscopico. Il primo si focalizza sul singolo individuo all'interno di una folla e come, per esempio, questo interagisca con gli altri pedoni o con l'ambiente in cui è immerso con modelli che sono computazionalmente più onerosi rispetto a quelli macroscopici. Per contro, i modelli macroscopici tendono a considerare la folla come se fosse un fluido [1] e pertanto forniscono misure più approssimate. Ad ognuno dei due punti di vista, corrisponde un'interpretazione diversa dello stesso problema. Le due questioni a cui siamo interessati sono il conteggio di persone e il tracking dei loro movimenti che, secondo l'interpretazione microscopica, assumono il significato rispettivamente di LOI (line-of-interest counting) e del tracking individuale di ogni pedone. LOI conta il numero di persone che passano attraverso una data linea virtuale mentre per quanto riguarda il tracking, le misure d'interesse si trovano nella velocità istantanea del pedone e nella sua direzione.

Se invece si considera la prospettiva macroscopica questi due problemi prendono il nome di ROI (region-of-interest counting), che conta il numero di persone in una data regione, e con il tracciamento generico della folla, registrando la sua velocità media e direzione del flusso.

Un importante aspetto delle folle è la loro densità ed è naturale pensare che folle di diverse densità dovrebbero ricevere diversi livelli di attenzione. Gli studi effettuati prima da Fruin [2] e poi ripresi da Polus et al. [3] ci forniscono una chiara idea del problema del *livello del servizio* per il traffico pedonale definito come: flusso libero, flusso limitato, flusso denso e flusso interrotto in accordo con una misura della densità definita come il numero di pedoni per unità di area. Weidmann inoltre dimostrò empiricamente che lo scopo dello spostamento influisce sulla velocità di moto [4] che per il traffico turistico o ricreativo stimò essere 0.99 m/s. Le stime sulla densità e sul movimento della folla ci possono aiutare a stabilire se siano in atto o meno comportamenti anormali o allarmanti [5] e a gestire situazioni potenzialmente pericolose.

Le soluzioni a LOI condividono tre step fondamentali: rilevazione del foreground per discriminare i pixel in movimento rispetto a quelli statici dello sfondo, riconoscimento di ogni singola persona nella scena e tracciamento della sua traiettoria per aumentare il contatore nel momento in cui questa interseca la linea virtuale d'interesse. Le differenze tra

i vari metodi ricadono nell'approccio utilizzato per ogni step e nel tipo di dispositivo utilizzato. I primi tentativi nel conteggio delle persone sono progettati per processare i video catturati da telecamere RGB. Molti di questi metodi utilizzano combinazioni di rilevazione del foreground come frame differencing o background subtraction.



Figura 1 Frame originale



Figura 2 Estrazione foreground

Il problema principale di queste soluzioni è che presentano diverse difficoltà in scenari molto affollati e soggetti a cambiamenti di luce. Per limitare l'impatto di questi problemi, cominciarono a sorgere studi su come svolgere un'analisi 3D per sfruttare l'informazione riguardo la profondità. I primi metodi che seguono questa linea di pensiero furono basati sull'utilizzo di telecamere stereoscopiche anche se con l'uscita di Microsoft Kinect, che utilizza la luce strutturata, nuove opportunità si sono aperte proprio con l'utilizzo di questo device [6], [7].

In questo capitolo verrà presentata una panoramica dei dispositivi esistenti da impiegare nel monitoraggio del traffico pedonale e delle esperienze di significativa importanza in cui sono stati utilizzati.

2.1. Tecnologie utilizzate

Esistono diversi tipi di dispositivi che vengono quotidianamente utilizzati per la Crowd Analysis: dispositivi che acquisiscono immagini, laser, fotocellule e sniffer. Le prime tre categorie riescono a gestire più o meno bene l'informazione riguardo il numero di ingressi/uscite da un varco, quindi problemi di tipo LOI. Le telecamere, inoltre, possono essere impiegate anche per la stima della densità pedonale nelle immagini

catturate (ROI). Gli sniffer, in particolare WiFi, si distinguono per la loro versatilità, specialmente se utilizzati in sinergia con molteplici esemplari.

2.1.1. Videocamere

I metodi convenzionali si approcciano al problema della Crowd Analysis utilizzando generiche telecamere a colori o in scala di grigio e si possono dividere in tre categorie: rilevamento e tracking dei pedoni, regressione sulla base di feature e clustering delle traiettorie. I fotogrammi possono anche essere utilizzati per rilevare comportamenti anomali nella folla [5] o, grazie a metodi recenti basati su deep learning, per fornire una stima della densità di questa [8], [9], [10].

Nonostante siano stati fatti grandi progressi, i problemi menzionati sono ben lontani dall'essere risolti a causa di svariate sfide che includono cambiamenti nell'illuminazione, ombre, occlusioni e diversità nell'aspetto dei pedoni. Inoltre, molte delle soluzioni si basano su immagini ad alta risoluzione e modelli complessi che rendono difficile l'applicabilità real-time. Per migliorare le performance del conteggio e per ridurre la complessità computazionale, sono stati proposti diversi metodi che sfruttano sensori di profondità differenti (Kinect, telecamere a tempo di volo, telecamere stereoscopiche ecc) in sinergia o in sostituzione alle telecamere RGB [6], [11], [12].

Sul mercato vi sono molteplici soluzioni, relativamente a basso costo, molte delle quali sfruttano appunto telecamere stereoscopiche [13], [14].



Figura 3 Telecamera stereoscopica Hikvision iDS-2CD6810F/C

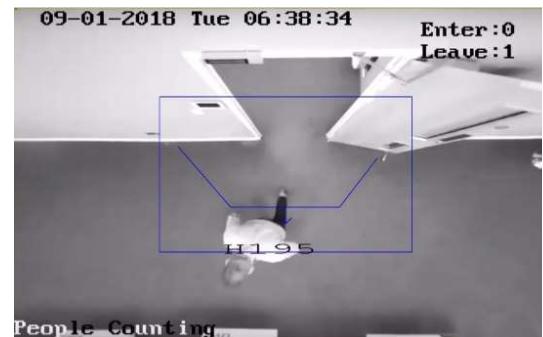


Figura 4 Frame tratto dalla telecamera a sinistra [15]

Da tenere in considerazione l'impatto estetico non indifferente che l'installazione comporta ed il fatto che in generale, per grandi masse, l'accuratezza dell'analisi cala. Trattando immagini, se il software non considera il problema, possono sorgere controversie riguardo la privacy e per conteggiare il numero di ingressi/uscite da un varco vi è il vincolo del posizionamento presso la verticale dello stesso.

2.1.2. Infrarossi e laser

Infrarossi e laser forniscono un'alternativa efficiente ai problemi sopra menzionati. In studi passati, telecamere ad infrarossi, scanner laser o una combinazione dei due sono stati impiegati in questo campo. Grazie all'abilità di acquisire immagini tramite radiazioni di calore, le camere IR sono in grado di discriminare meglio i soggetti mentre gli scanner laser si possono ottenere maggiori informazioni sulla profondità di campo. In aggiunta, con device laser, non sono necessarie immagini ad alta risoluzione, riducendo di molto la complessità computazionale e rendendo le operazioni idonee ad un ambiente real-time. Wu et al. [16] mostrano un approccio innovativo basato su un laser ad infrarossi multipunto per ricavare l'informazione sulla profondità e costruire un modello spaziale dello sfondo per l'individuazione delle persone senza il bisogno di immagini ad alta risoluzione.

STMicroelectronics [17] ci dimostra come la ricerca in questo campo stia dando vita a soluzioni sempre più convenienti e minimali [18]. LZR-SIGMA, per esempio, è un dispositivo laser ad infrarossi con misura a tempo di volo [19] facilmente reperibile sul mercato europeo.



Figura 5 LZR®-SIGMA. Soluzione contapersone laser con campo di rilevazione ad alta risoluzione (250 punti per battente)



Figura 6 SIGMA possiede un sistema di conteggio affidabile che offre una misurazione accurata del flusso di adulti, bambini e gruppi di persone.

2.1.3. Sniffer

Considerata la notevole importanza che gli smartphone stanno a mano a mano acquisendo per via della loro capacità di accedere al web utilizzando dati mobili e reti Wi-Fi, l'analisi del traffico wireless si è rivelata un'ottima alternativa ai metodi convenzionali sopracitati. Si definisce sniffing (eavesdropping) l'attività di intercettazione passiva dei dati che transitano in una rete. I software che sono in grado di captare i singoli pacchetti del traffico Internet prendono il nome di sniffer e quelli Wi-Fi necessitano di un modulo wireless che stia in ascolto del traffico. Laddove i metodi di videosorveglianza o quelli basati su laser e IR si applicano molto bene ad analisi di tipo LOI, facendo leva sull'attaccamento che l'uomo ha per il proprio dispositivo mobile, gli sniffer Wi-Fi trovano efficace impiego nella stima delle densità, flussi e tempo di viaggio. Recenti studi vedono impiegati molteplici sniffer che in sinergia rendono possibile la stima della posizione di un dispositivo mobile all'interno di un'area schematizzabile in celle grazie a tecniche di machine learning [20], la stima della densità di persone all'interno di un centro commerciale [21], la stima del flusso e dei percorsi pedonali [22], [23], [24] o il monitoraggio di grandi eventi come il pellegrinaggio Hajj verso la Mecca [25].

Esistono alcuni prodotti specializzati [26] ma la maggior parte degli scienziati, che in questi anni hanno cercato di sfruttare questo metodo di monitoraggio emergente, si affida ancora a soluzioni “artigianali”, spesso costituite da una o più stazioni composte da un single-board computer (come Raspberry Pi [27]) con un modulo Wi-Fi usb esterno. Il motivo di questa scelta risiede nell'installazione di un sistema che, a differenza di quelli di videosorveglianza, risulta economico, semplice e senza la necessità di hardware aggiuntivo.

L'ottenimento di informazioni tramite lo sniffing Wi-Fi di per sé non è un processo complicato anche perché una delle caratteristiche delle reti wireless è proprio quella di essere broadcast. I dispositivi con il modulo Wi-Fi acceso, infatti, sono regolarmente in cerca di connettività. Mandano continuamente richieste per cercare reti a cui hanno già effettuato un accesso in passato. Queste richieste contengono tra le altre cose, informazioni che consentono l'identificazione del dispositivo inviante. Lo scopo però che in questo caso si vuole raggiungere tramite

lo sniffing, ovvero la Crowd Analysis, comporta uno studio non banale, dovuto a varie ragioni. In primo luogo, i dati ottenuti sono quelli di un sottoinsieme del traffico reale, in cui il penetration rate dei dispositivi con Wi-Fi attivato gioca un ruolo fondamentale. In secondo luogo, la rilevazione di un dispositivo dipende in primis da esso stesso poiché, nonostante il corretto funzionamento del Wi-Fi, può risultare comunque momentaneamente invisibile. Infine, a seconda dei casi, i dispositivi potrebbero non lasciare un'impronta univoca e quindi il conteggio andrebbe a considerare dei doppioni o peggio triple o quadruple copie come entità separate.

2.2. Progetti di monitoraggio

Oggiorno, molti eventi di larga scala sono ospitati in aree urbane che non sono progettate per ricevere questo grande afflusso di spettatori. Oltre al numero di visitatori fuori misura, il comportamento di queste folle è spesso imprevedibile e non contemplato. Mai come in questo periodo sono necessari sistemi di monitoraggio delle folle che forniscano informazioni real-time riguardo lo stato del traffico pedonale, per garantire la sicurezza dei visitatori, prevenendo situazioni di pericolo e colli di bottiglia o identificandole tempestivamente.

Di seguito sono riportati alcuni tentativi di monitoraggio effettuati in passato durante eventi di larga scala i quali hanno utilizzato più tecnologie contemporaneamente per avere una visione il più completa possibile della situazione.

2.2.1. SAIL Amsterdam (2015)

SAIL è un evento marittimo quinquennale che si tiene ad Amsterdam, in Olanda [28]. Ha il vanto di essere una delle più grandi (se non la più grande) manifestazione nautica al mondo e senz’altro il più grande evento pubblico Olandese. Ogni cinque anni più di 600 navi tra velieri, fregate, natanti, navi moderne, militari e repliche di imbarcazioni storiche navigano lungo il canale del mare del Nord per poi ormeggiare al porto di IJhaven ad Amsterdam. Il legame tra Amsterdam e il mare è



Figura 7 La Amerigo Vespucci italiana

profondo e SAIL non è solo un’esibizione di grandi navi. In poche parole, comprende “qualsiasi cosa che galleggi” e durante l’edizione del 2000 si raggiunse l’impressionante numero di 8000 imbarcazioni. Accanto alle attività in acqua ci sono molte attrazioni anche a terra come

musica, arte, cultura e tante attività per bambini.

SAIL fu organizzato per la prima volta nel 1975 e visto il grande successo, si è deciso di ripetere l’evento in forma quinquennale, riunendo nelle recenti edizioni circa 2 milioni di persone sul territorio.

Durante SAIL 2015, una combinazione di sistemi di conteggio video, tracciatori GPS e sensori Wi-Fi fu impiegata per determinare lo stato del traffico pedonale [29], [30]. I dati vennero raccolti utilizzando 8 telecamere per conteggio, 20 sensori Wi-Fi e 324 tracciatori GPS furono distribuiti ai visitatori per registrare le loro traiettorie. Il problema fondamentale dello studio però fu che non c’era una vera e propria conoscenza assoluta (in termini tecnici ground-truth) del traffico reale a cui fare riferimento, ma i calcoli vennero comparati alle immagini registrate da altre telecamere aeree. Le 8 telecamere ebbero il compito di calcolare il numero di pedoni che attraversano una sezione trasversale. In teoria, i dispositivi impiegati avevano un’accuratezza del 98% decrescente a 92% in situazioni di alta densità ma nessuna considerazione su velocità o percorsi può essere fatta con questi dati.

I sensori Wi-Fi invece ebbero lo scopo di derivare i flussi e il tempo di percorrenza, acquisendo informazioni riguardo i pedoni dotati di un qualsiasi dispositivo mobile (es. smartphone) all'interno del loro raggio d'azione. I dati contenenti i gli indirizzi MAC dei dispositivi, l'ID del sensore e il primo e ultimo timestamp di rilevamento di ogni dispositivo rendono possibile il calcolo cumulativo dei singoli dispositivi e dei tempi di attraversamento da sensore a sensore (supponendo che ad un dispositivo corrispondesse un pedone).

Confrontando i dati ottenuti dai sensori Wi-Fi e dalle telecamere, si nota che i sensori percepirono circa un terzo dei pedoni conteggiati dalle telecamere e la metà di questi (quindi 1/6) corrispondono a dispositivi unici. Nonostante un penetration rate così basso, i calcoli eseguiti su approssimativamente 2.3 milioni di visitatori sono sufficienti per avere un'idea dei percorsi effettuati e sui tempi di tragitto per alcune tratte. Per impiegare i sensori Wi-Fi nella stima della densità tuttavia, sono necessarie stime accurate sul penetration rate (rapporto tra pedoni totali e numero di dispositivi mobili rilevati) in modo da poter stimare in maniera più accurata il traffico reale.

Nello studio citato vengono illustrati nel dettaglio 3 algoritmi di stima dello stato per stimare travel time, densità e flussi [30].

2.2.2. Pellegrinaggio Hajj (2016)

Hajj è il tradizionale pellegrinaggio annuale verso la Mecca. Essendo il quinto pilastro dell'Islam e momento di purificazione per i fedeli, l'Hajj è considerato il pellegrinaggio più grande al mondo in cui si riuniscono 2-3 milioni di musulmani provenienti da tutto il globo. L'Hajj è una sequenza di riti praticati in tempi e luoghi prestabiliti e questi vincoli spazio-temporali rendono la gestione dell'Hajj un processo complesso. È con lo studio di come si comportano i pellegrini, dei loro pattern, delle interazioni, delle necessità e domande che si può fornire un livello di servizio soddisfacente. Con questo pensiero, A. Basalamah [25] installò un sistema composto da 8 sniffer Wi-Fi autoalimentati a

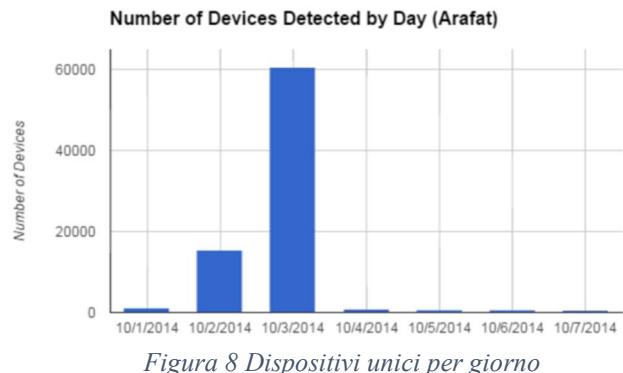


Figura 8 Dispositivi unici per giorno

energia solare in una delle aree destinate ad ospitare i pellegrini per un giorno, Arafat, 12 km a sud est della Sacra Città della Mecca. Arafat è un'area deserta per tutto il resto

dell'anno in cui durante il pellegrinaggio vengono allestite delle tende per i fedeli e nel 2016 furono stimate 185.000 persone. La collezione dei dati cominciò due giorni prima del picco stimato e continuò per i successivi tre giorni. I risultati ottenuti mostrarono che il sistema fu in grado di rilevare circa il 37.5% della folla, corrispondente a 69.467 dispositivi unici. Dalla rete di sensori, Basalamah fu in grado di identificare le aree più affollate, stabilire gli orari dei picchi di mobilità, estrapolare informazione riguardo lo stato economico dei pellegrini avendo a disposizione le marche delle case produttrici dei dispositivi e diverse altre statistiche.

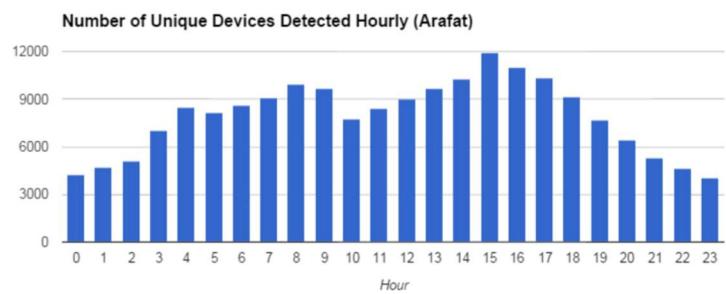


Figura 9 Dispositivi unici per ora (giorno di picco)

2.2.3. Il caso Venezia (2018)

VENIS - Venezia Informatica e Sistemi S.p.A. è l'azienda di servizi ICT e l'operatore locale di comunicazioni elettroniche del Comune di Venezia. Nei primi mesi del 2018, VENIS ha aperto un bando rivolto a soggetti privati ad offrirsi per una sperimentazione gratuita di impianti atti al conteggio dei flussi pedonali nelle aree limitrofe alla Stazione Santa Lucia ed in Piazza San Marco. Le sperimentazioni hanno visto come oggetto l'installazione gratuita e il test di impianti tecnologici e/o sensoristica e/o videorilevazione nelle aree del Centro Storico di Venezia. In sei mesi, i prodotti di sei ditte sono stati messi a confronto e per ognuno è stato stilato un resoconto riguardo tipologia del sensore, attendibilità, flessibilità, installazione, estetica e disponibilità dei dati. I dispositivi presi in analisi si riferiscono a laser, telecamere stereoscopiche, sniffer Wi-Fi, telecamere ad infrarossi e telecamere con annesso software per il riconoscimento di transiti. Di seguito viene riportata una tabella riassuntiva, ma è disponibile alla lettura il report completo [31].

Ditta	Sensore	Attendibilità	Flessibilità	Installazione	Estetica	Disponibilità dati
Hitachi	3D ToF	Molto buona	Portata limitata	Semplice	Ridotte dimensioni ma richiede mensola di protezione e box apparati	Real-time
Nicoli	Barriera LASER	non valutata	Portata elevata	Va installata in quota, installazione semplice	Dimensioni generose ma installata ad altezza elevata non si nota, richiede box apparati	Real-time
Nicoli	Telecamera stereoscopica	Molto buona	Buona	Semplice	Poco impattante, richiede box apparati	Real-time
Nicoli	Sniffer Wi-Fi	non valutata	Elevata	Semplice	Poco impattante (può essere anche nascosto)	Anche real time, in funzione della informazione richiesta
ISCOM	Telecamera	Molto buona	Elevata	Semplice	Poco impattante	Anche real-time (con precisione ridotta)
AXIANS	Telecamera stereoscopica	Eccellente	Elevata	Semplice	Poco impattante, richiede box apparati	Real-time
ZIBONI	Sensore PIR a batteria	Buona, non adatto a situazioni di congestione	Molto Elevata	Semplice	Poco impattante	Differita di un giorno
ZIBONI	Telecamera IR	Buona	Portata limitata	Semplice	Piccolo ma evidente, richiede box apparati	Real-time
Telecom Italia	Telecamera (transiti)	Molto buona	Elevata	Semplice	Poco impattante, richiede box apparati	Real-time
Telecom Italia	Telecamera (presenze)	non valutata	Elevata	Semplice	Poco impattante, Richiede box apparati	Real-time

Figura 10 Sintesi dei risultati

3. Le connessioni Wi-Fi

Lo standard IEEE (Institute of Electrical and Electronics Engineers) 802.11 definisce tre tipi di frame che vengono scambiati dai dispositivi Wi-Fi: Management frames, Control frames e Data frames. I sistemi sniffer in esame fanno leva sulle informazioni contenute nei Management frames per cui gli altri due tipi di non verranno menzionati oltre. Per quanto riguarda i Management frames, esistono quattro principali sottotipi: Beacons, Probes, Association e Authentication a loro volta divisi in ulteriori classi.

I Beacon frames (sottotipo 0x08) sono messaggi che gli AP (access point) inviano periodicamente in modalità broadcast per segnalare ai dispositivi nei dintorni la propria esistenza e il proprio SSID. Le Probe requests (sottotipo 0x04) sono dei frame speciali inviati da una stazione client che richiede informazioni da un particolare AP specificato da un SSID o da tutti quelli nei dintorni con un SSID broadcast. L'invio di una Probe request avviene quando la scheda wireless del dispositivo sta eseguendo il cosiddetto active scanning, cioè la ricerca attiva di AP, mentre il semplice ascolto dei frame Beacon è denominato passive scanning. In seguito ad una Probe request, gli AP emettono una Probe response (sottotipo 0x05) contenente le informazioni necessarie a stabilire una connessione.

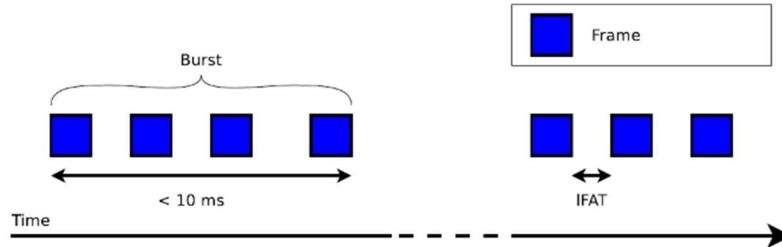


Figura 11 Schema di una sequenza di trasmissione. Un burst è un gruppo di frame di tipo probe request inviati da un device nell'arco di 10ms. L'inter-frame arrival time (IFAT) indica il lasso di tempo tra due frame. [32]

Ogni dispositivo elettronico porta con sé un'informazione univoca, un indirizzo MAC (media access control) che può essere intercettato quando esso tenta di collegarsi ad una rete wireless o il collegamento è già stato effettuato perché è contenuto in chiaro nelle Probe requests. È quindi mediante lo sniffing di queste richieste, durante la fase di active scanning, che si possono collezionare informazioni riguardo i client (nel nostro caso gli smartphone dei turisti) come indirizzo MAC, potenza del segnale (RSSI), timestamp, sequence number della richiesta e svariati altri dati. Come intuibile, le Probe requests sono più frequenti

quando il dispositivo non è connesso a nessuna rete e la loro frequenza diminuisce drasticamente quando esso si connette.

In risposta alla possibilità di ottenere facilmente informazioni dalle Probe requests, molti dei sistemi operativi per smartphone stanno via via implementando sistemi di randomizzazione dell'indirizzo MAC in modo da proteggere la privacy degli utenti e impedire il tracking dei loro spostamenti. In particolare, come constatato empiricamente in un precedente progetto del sottoscritto [33] che conferma quanto trovato in [34], i dispositivi iOS sono quelli che effettuano la randomizzazione in modo più ostico per quello che riguarda l'attribuzione di indirizzi randomici allo stesso dispositivo.

3.1. MAC Address

L'indirizzo MAC (Media Access Control) è un codice di 48 bit (6 byte) assegnato in modo univoco dal produttore ad ogni scheda di rete ethernet o wireless prodotta al mondo, tuttavia modificabile a livello software. Al fine di garantire l'univocità dell'indirizzo MAC tra i dispositivi, l'IEEE assegna blocchi di indirizzi alle organizzazioni in cambio di una tassa. Questi blocchi, conosciuti più comunemente come OUI (Organizationally Unique Identifier), possono essere comprati e registrati tramite l'IEEE che da all'organizzazione il controllo e la responsabilità per tutti gli indirizzi con un dato prefisso a tre byte. I produttori a questo punto sono liberi di assegnare i rimanenti tre byte a qualsiasi valore essi vogliano con l'unica accortezza di non utilizzare lo stesso indirizzo due volte.

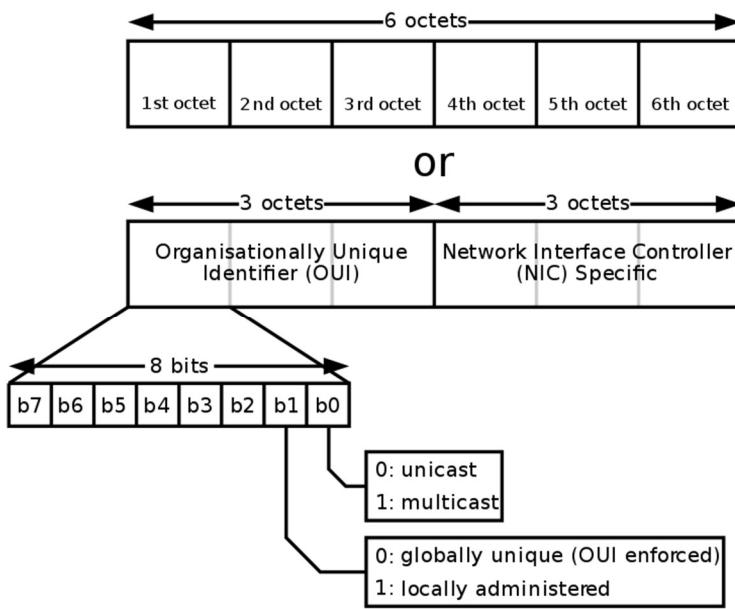


Figura 12 Diagramma che mostra la struttura di un indirizzo MAC-48, con indicate esplicitamente le posizioni del multicast/unicast bit e OUI/local bit

Oltre all'indirizzo MAC globale del produttore e di dominio pubblico, i dispositivi moderni utilizzano frequentemente un indirizzo localmente assegnato che è contraddistinto da un local bit nel byte più significativo dell'indirizzo. Gli indirizzi localmente assegnati non danno la garanzia di univocità e generalmente non sono utilizzati in maniera persistente. Questi infatti, possono essere utilizzati per generare indirizzi MAC randomici come misura aggiuntiva per la privacy. Poiché, in fase di active scanning, i dispositivi comunicano la loro presenza agli AP nelle vicinanze e le Probe request richiedono l'indirizzo MAC della sorgente, se un device utilizza il suo indirizzo globale allora sta effettivamente facendo broadcast della sua identità. Analogamente agli OUI, i produttori possono acquistare da IEEE dei prefissi CID (Company Identifier) con la promessa di non utilizzarli in forma globale per i dispositivi. Di conseguenza, i CID hanno sempre il local bit asserito. Un esempio di questo tipo è dato dal prefisso DA:A1:19 posseduto da Google. Con l'impiego di MAC address randomizzati che cambiano nel corso del tempo, il tracking dei dispositivi risulta non più banale. Nel momento in cui un device però si connette ad un AP, smette di randomizzare ed utilizza il suo MAC reale. Per contro, come già accennato, l'evento che ci interessa sniffare, ovvero la fase di active scanning, si verifica con frequenza nettamente inferiore.

3.2. Randomizzazione

La randomizzazione dell’indirizzo MAC, a cui da questo momento in poi ci riferiremo semplicemente con “randomizzazione”, è una comune misura di protezione per la privacy applicata da buona parte dei sistemi operativi di uso quotidiano. Viene impiegata per prevenire il tracking degli utenti tramite le probe requests che sono trasmesse durante le scansioni network IEEE 802.11.

Questo significa che, per esempio, può essere utilizzato un nuovo MAC address per ogni scansione, dove una scansione consiste nell’invio di probe requests su tutti i canali utilizzabili. Tuttavia, siccome non esiste ancora uno standard sulla randomizzazione, iOS, Android, Windows e Linux implementano le proprie varianti. Questo fa porre la domanda circa la sicurezza che la randomizzazione dovrebbe garantire.

Come precedentemente accennato, la randomizzazione si ferma una volta che un dispositivo viene associato con un AP, per questo nella letteratura si trovano diversi attacchi il cui scopo è risalire al MAC univoco a partire da quello randomico. Un esempio è il cosiddetto Karma attack in cui si costruisce un fake AP il cui SSID corrisponde a uno di quelli ricercati dal dispositivo di cui si vuole risalire all’indirizzo globale, costringendolo così a comunicare il suo vero MAC [34], [35].

Visto che, con le recenti implementazioni di randomizzazione i dispositivi trasmettono più raramente i propri SSID conosciuti, (riducendo l’efficacia del Karma attack), sia perché questo lavoro vuole essere del tutto passivo, senza violazioni di privacy o collezione di dati sensibili, non verranno utilizzati attacchi sui singoli dispositivi ma verrà studiato un modo per aggregare quanto più possibile i MAC address randomici sotto uno stesso dispositivo.

In nessun momento si è tentato di decifrare qualsivoglia dato o di intraprendere azioni attive per modificare il comportamento della rete e soprattutto non abbiamo mai avuto modo di collegare indirizzi MAC alle identità degli individui possessori dei dispositivi.

3.2.1. Android

I sistemi Android cominciarono ad impiegare la randomizzazione con la versione 6.0 che veniva effettuata durante le scansioni in background se l'hardware la supportava. Fino all'uscita di Android 8.0 però questa era utilizzata da una piccolissima percentuale di dispositivi. Google acquistò poi il proprio CID DA:A1:19 utilizzato tutt'ora dalla stragrande maggioranza dei dispositivi Android che randomizzano, lasciando ai software la randomizzazione dei restanti 3 byte dell'indirizzo MAC (2^{24}). Con la versione 9.0 del robottino verde l'opzione aggiuntiva si fece ancora più accessibile comparendo direttamente tra le impostazioni sviluppatore dei dispositivi ma di default non attiva. Nonostante Google abbia gettato tutte le basi per la randomizzazione sui dispositivi, e che praticamente ogni altro cellulare all'infuori degli iPhone monti una versione di Android, questo non è il sistema operativo in cui si nota la maggior quantità di randomizzazione. Per problemi di compatibilità tra software e hardware o per decisione dei produttori stessi che customizzano le versioni del sistema operativo di Google, la maggior parte della randomizzazione arriva da Apple che produce sia hardware che software. Secondo tutti gli studi svolti fino ad oggi, per esempio, non si è mai ancora registrato un dispositivo Samsung che effettuasse la randomizzazione, nonostante questo sia il maggior produttore di dispositivi Android sul mercato.

Un esempio di randomizzazione da parte di Android è mostrato dalla seguente tabella, ottenuta da un esperimento in un ambiente quanto più controllato possibile. Con il tool Wireshark [36] mi sono messo in ascolto dei pacchetti wireless in transito con l'intenzione di filtrare le probe request inviate da uno Xiaomi Redmi Note 7 in mio possesso, non associato ad alcuna rete. Il dispositivo monta la versione 9 di Android e come si può notare, in accordo con quanto descritto precedentemente, Wireshark sostituisce automaticamente il CID conosciuto “DA:A1:A19” con il tag “Google_” mostrando i 3 byte finali dell'indirizzo su cui il sistema operativo effettua la randomizzazione.

Time (s)	Source	Destination	Lenght	Info
0.000000000	Google_0e:fe:86	Broadcast	153	SN=1246
0.020056565	Google_0e:fe:86	Broadcast	153	SN=1247
20.478772870	Google_16:71:b9	Broadcast	153	SN=1272
20.519636137	Google_16:71:b9	Broadcast	153	SN=1275
40.607593376	Google_fc:02:ca	Broadcast	153	SN=1305
40.627627762	Google_fc:02:ca	Broadcast	153	SN=1306
60.529810066	Google_00:b4:8a	Broadcast	153	SN=1333
60.550520580	Google_00:b4:8a	Broadcast	153	SN=1334
80.558958993	Google_a9:08:4f	Broadcast	153	SN=1362
80.578992854	Google_a9:08:4f	Broadcast	153	SN=1363
100.575174121	Google_c8:d7:94	Broadcast	153	SN=1393
100.597433718	Google_c8:d7:94	Broadcast	153	SN=1394
126.177714102	Google_62:54:e4	Broadcast	153	SN=1416
126.197759739	Google_62:54:e4	Broadcast	153	SN=1417
153.825444973	Google_e7:d1:31	Broadcast	153	SN=1441
153.996834953	Google_e7:d1:31	Broadcast	153	SN=1442
157.493772620	Google_e2:48:82	Broadcast	153	SN=1466
157.513863436	Google_e2:48:82	Broadcast	153	SN=1467
177.401985590	Google_fb:a2:0c	Broadcast	153	SN=1489
177.449180018	Google_fb:a2:0c	Broadcast	153	SN=1492
201.581892344	Google_27:46:70	Broadcast	153	SN=1520
201.603682772	Google_27:46:70	Broadcast	153	SN=1521
221.572785571	Google_c8:88:a9	Broadcast	153	SN=1548
221.592971051	Google_c8:88:a9	Broadcast	153	SN=1549
243.207523855	Google_e5:43:1c	Broadcast	153	SN=1576
243.227686911	Google_e5:43:1c	Broadcast	153	SN=1577
266.572511158	Google_0e:dc:d6	Broadcast	153	SN=1603
266.592636835	Google_0e:dc:d6	Broadcast	153	SN=1604
286.593584674	Google_71:26:8e	Broadcast	153	SN=1631
286.613684433	Google_71:26:8e	Broadcast	153	SN=1632

Tabella 1 Randomizzazione Xiaomi Redmi Note 7

La tabella riporta cinque minuti (trecento secondi) di ascolto durante i quali il dispositivo emette 30 probe request sul canale di ascolto, corrispondenti a 15 indirizzi MAC differenti. Si nota subito che i burst intercettabili su un canale sono composti da due frame e che l'indirizzo cambia ad ogni burst.

Intervallo medio di randomizzazione (inter-burst)	20.625 s
Minimo intervallo inter-burst	3.7 s
Massimo intervallo inter-burst	33.9 s
Minimo Sequence number inter-frame	1
Minimo Sequence number inter-burst	22
Massimo Sequence number inter-frame	3
Massimo Sequence number inter-burst	30

Tabella 2 Resoconto randomizzazione Android

La lunghezza dei frame in ogni caso è sempre la stessa, 153 bytes, caratteristica su cui si basa il meccanismo di fingerprinting descritto nel capitolo 4, mentre gli SSID conosciuti non compaiono mai. Il destinatario di ogni richiesta è sempre ff:ff:ff:ff:ff:ff (broadcast) ovvero qualsiasi AP in ascolto.

3.2.2. Apple

Apple introdusse la randomizzazione dell'indirizzo MAC a partire da iOS 8 ma non fu un grande successo. Essa si manifestava a patto che fossero soddisfatte tre condizioni: il wifi fosse attivo ma non connesso ad alcuna rete, il dispositivo fosse in “sleep mode” (schermo spento per inattività) e i servizi di localizzazione fossero disattivati. Neanche a dirlo quasi nessuno avrebbe mantenuto il dispositivo in questo stato. La maggior parte delle app non funziona nemmeno se i servizi di localizzazione non sono attivati e sarebbe del tutto inutile se la randomizzazione terminasse con l'accensione del display. Da iOS 9, Apple estese questa funzionalità a quello che definisce come locations and auto-join scans, che persiste all'attuale iOS 12 [37] e garantisce che la randomizzazione venga effettuata anche a schermo acceso. Al contrario di Google che ha acquistato il proprio CID, gli indirizzi Randomizzati di Apple sembrano non avere un prefisso predefinito e la randomizzazione avviene su

tutti i bit ad eccezione del local bit settato a 1 e del multicast bit a 0. Ne risulta una randomizzazione su 2^{46} bit.

Come per Android è disponibile un esempio di randomizzazione iOS registrato da un iPhone 6 con iOS 11.2.6 in mio possesso, nelle medesime condizioni del precedente esperimento.

Time (s)	Source	Destination	Length	Info
0.000000000	1e:69:da:39:e9:7a	Broadcast	176	SN=912
0.003260863	1e:69:da:39:e9:7a	Broadcast	176	SN=913
3.229036747	ae:ef:ff:f5:f4:b6	Broadcast	176	SN=949
3.249793790	ae:ef:ff:f5:f4:b6	Broadcast	176	SN=950
17.124327398	d6:5c:53:c5:ce:c8	Broadcast	176	SN=986
17.144741222	d6:5c:53:c5:ce:c8	Broadcast	176	SN=987
20.062601984	06:2b:1d:2d:d8:f9	Broadcast	176	SN=1023
20.082064693	06:2b:1d:2d:d8:f9	Broadcast	176	SN=1024
26.086563110	b6:45:d3:ef:c3:7a	Broadcast	176	SN=1062
26.107303803	b6:45:d3:ef:c3:7a	Broadcast	176	SN=1063
30.028257945	8a:29:9b:41:79:f2	Broadcast	176	SN=1097
30.048418686	8a:29:9b:41:79:f2	Broadcast	176	SN=1098
38.093549750	92:c9:45:16:67:a1	Broadcast	176	SN=1134
38.114193199	92:c9:45:16:67:a1	Broadcast	176	SN=1135
51.215853754	be:20:90:e8:0f:6b	Broadcast	176	SN=1170
51.217149709	be:20:90:e8:0f:6b	Broadcast	176	SN=1171
61.820964072	9e:e9:d2:62:c4:b5	Broadcast	176	SN=1207
61.841800507	9e:e9:d2:62:c4:b5	Broadcast	176	SN=1208
66.156033562	d2:32:e7:09:15:0d	Broadcast	176	SN=1246
66.176003065	d2:32:e7:09:15:0d	Broadcast	176	SN=1247
69.123112913	e2:16:2c:ec:3e:77	Broadcast	176	SN=1281
69.149998832	e2:16:2c:ec:3e:77	Broadcast	176	SN=1282
71.975087397	8a:62:b6:21:9b:c7	Broadcast	176	SN=1318
72.003422778	8a:62:b6:21:9b:c7	Broadcast	176	SN=1319
80.086840657	4a:5e:15:5a:7a:6e	Broadcast	176	SN=1357

80.107596215	4a:5e:15:5a:7a:6e	Broadcast	176	SN=1358
84.106216095	a2:4d:3a:85:67:2c	Broadcast	176	SN=1392
84.126922600	a2:4d:3a:85:67:2c	Broadcast	176	SN=1393
92.158398074	42:b8:63:4c:33:bb	Broadcast	176	SN=1429
92.176315718	42:b8:63:4c:33:bb	Broadcast	176	SN=1430
105.246452275	b6:67:97:0a:d6:bf	Broadcast	176	SN=1466
105.266143503	b6:67:97:0a:d6:bf	Broadcast	176	SN=1467
123.258235152	4a:9a:8d:19:95:29	Broadcast	176	SN=1503
123.278961637	4a:9a:8d:19:95:29	Broadcast	176	SN=1504
141.275726553	b6:70:f1:80:7a:44	Broadcast	176	SN=1540
141.296125659	b6:70:f1:80:7a:44	Broadcast	176	SN=1541
174.300246061	9a:00:33:27:84:5e	Broadcast	176	SN=1583
174.320940406	9a:00:33:27:84:5e	Broadcast	176	SN=1584
178.774824196	46:60:42:eb:93:ed	Broadcast	176	SN=1622
178.794451431	46:60:42:eb:93:ed	Broadcast	176	SN=1623
182.868590586	2e:68:86:13:f4:af	Broadcast	176	SN=1657
182.888304434	2e:68:86:13:f4:af	Broadcast	176	SN=1658
187.339999041	4a:62:68:7d:cb:b4	Broadcast	176	SN=1693
187.361313909	4a:62:68:7d:cb:b4	Broadcast	176	SN=1694
195.365936991	de:d6:a1:d6:cf:e7	Broadcast	176	SN=1730
195.386577098	de:d6:a1:d6:cf:e7	Broadcast	176	SN=1731
208.365194904	42:6c:41:97:08:6e	Broadcast	176	SN=1767
208.385796120	42:6c:41:97:08:6e	Broadcast	176	SN=1768
220.761622918	7e:18:0d:89:00:98	Broadcast	176	SN=1804
220.782187109	7e:18:0d:89:00:98	Broadcast	176	SN=1805
226.370413892	3a:09:8f:26:a0:de	Broadcast	176	SN=1841
226.391156374	3a:09:8f:26:a0:de	Broadcast	176	SN=1842
240.587248682	4a:0e:f2:f6:d3:10	Broadcast	176	SN=1880
240.606907671	4a:0e:f2:f6:d3:10	Broadcast	176	SN=1881
244.577034944	7e:6c:c6:72:25:bf	Broadcast	176	SN=1915
244.597226528	7e:6c:c6:72:25:bf	Broadcast	176	SN=1916

252.612058797	72:25:53:5c:7a:36	Broadcast	176	SN=1952
252.631753251	72:25:53:5c:7a:36	Broadcast	176	SN=1953
265.605307175	6a:bf:00:70:18:f1	Broadcast	176	SN=1989
265.625013224	6a:bf:00:70:18:f1	Broadcast	176	SN=1990
283.639501897	9e:6c:30:79:09:a2	Broadcast	176	SN=2032
283.659220793	9e:6c:30:79:09:a2	Broadcast	176	SN=2033
288.725051371	1e:04:51:b8:4f:9d	Broadcast	176	SN=2071
288.745717831	1e:04:51:b8:4f:9d	Broadcast	176	SN=2072

Tabella 3 Randomizzazione iPhone 6

La prima cosa che salta all'occhio è sicuramente la maggior quantità di record registrati nello stesso arco di tempo, 64 per la precisione. La seconda sostanziale differenza riguarda gli indirizzi randomici stessi. L'esperimento conferma che Apple non ha acquistato un CID, di conseguenza l'indirizzo viene generato casualmente in tutta la sua interezza (ad eccezione del local bit e del multicast bit).

Analogamente al caso precedente, iOS non rivela i propri SSID conosciuti, le lunghezze dei frame rimangono invariate e l'indirizzo viene randomizzato ad ogni burst.

Intervallo medio di randomizzazione (inter-burst)	9.4 s
Minimo intervallo inter-burst	3.22 s
Massimo intervallo inter-burst	33 s
Minimo Sequence number intra-burst	1
Minimo Sequence number inter-burst	34
Massimo Sequence number intra-burst	1
Massimo Sequence number inter-burst	40

Tabella 4 Resoconto randomizzazione iOS

3.2.3. Windows

Microsoft supporta la randomizzazione a partire da Windows 10. È possibile attivarla direttamente dalle impostazioni a patto che l'hardware e i driver lo

consentano. Quello che è interessante della randomizzazione di Microsoft è che utilizza MAC address randomici anche quando il dispositivo si sta connettendo ad una rete wireless, al contrario di Android ed Apple che randomizzano solamente quando effettuano la scansione. Windows è in grado di raggiungere questo obiettivo perché utilizza sempre lo stesso MAC randomico ogni volta che si connette ad una rete specifica. Per assicurare che il client venga sempre riconosciuto, l'indirizzo per rete viene calcolato nel seguente modo:

```
address = SHA-256(SSID, macaddr, connectionId, secret) [35]
```

Equazione 1 Creazione indirizzo randomico per rete su Windows

dove *SSID* è ovviamente il nome della rete, *macaddr* indica il vero indirizzo MAC del dispositivo, *connectionId* è un parametro che cambia ogni qualvolta l'utente dimentica la rete e *secret* è un parametro crittografico a 256 bit, unico per interfaccia che viene generato durante l'inizializzazione del sistema e mantenuto uguale ad ogni reboot.

3.2.4. Linux

Anche Linux supporta la randomizzazione che è stata introdotta nella versione del kernel 3.18. A seconda della versione del firmware il MAC cambia più o meno rapidamente durante le scansioni. La durata di default per un MAC randomico in *wpa_supplicant* è di 60 secondi.

4. Tecnologia proposta

In questo capitolo verrà illustrato il sistema proposto nella sua interezza: la sua architettura, lo script di sniffing, l'algoritmo di aggregazione, le caratteristiche dello strumento, la struttura del database e la realizzazione dell'applicazione web.

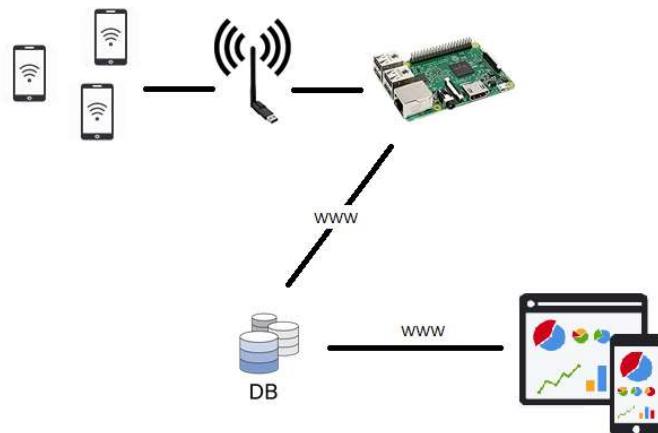


Figura 13 Schema operativo del sistema

4.1. Architettura del sistema

Un primo prototipo di sniffer Wi-Fi è stato realizzato durante l'anno accademico 2017/2018. Il progetto prevedeva l'ammodernamento di uno script Python preesistente che permettesse lo sniffing del traffico Wi-Fi. Il software era destinato a girare su una stazione indipendente composta da un Raspberry Pi dotato di modulo Wi-Fi esterno per lo sniffing, che rendesse disponibili i risultati su un database esterno attraverso la rete.

Sulla base di test sul campo e su [34], [35] e [38] è stato stabilito un criterio per attribuire ad un unico dispositivo i MAC address randomizzati generati dalla stessa sorgente. L'output del sistema sul database coincide con il numero di MAC address unici (globali + randomici aggregati) conteggiati nell'intervallo di tempo che intercorre tra una scrittura e l'altra e quindi con il numero di dispositivi con il modulo Wi-Fi acceso e tracciabili nel raggio d'azione dell'antenna wireless.

A seguire le schede tecniche dell'antenna wireless esterna e dei Raspberry utilizzati.

- Scheda di rete ALFA Network AWUS036NH



Chipset	Ralink RT3070
WiFi Standards	IEEE 802.11/b/g/n
WiFi Frequency	2.4GHz
Antenna Type	2.4GHz 5dBi dipole antenna
Wireless	802.11b: up to 11Mbps
Performance	802.11g: up to 54Mbps 802.11n: up to 150Mbps
Wireless Security	WEP, WPA, WPA2, WPA Mixed, WPS
Interface	USB 2.0, Mini USB, type B female connector
OS Requirement	Windows XP, Vista, 7, 8/8.1 and Windows 10 32/64bit, macOS 10.5 to 10.14 or later Linux

- Raspberry Pi 3 Model B+ & OS



SoC	Broadcom BCM2837B0 quad-core A53 (ARMv8) 64-bit @ 1.4GHz
Networking	Gigabit Ethernet (via USB channel), 2.4GHz and 5GHz 802.11b/g/n/ac Wi-Fi
RAM	1GB LPDDR2 SDRAM
Bluetooth	Bluetooth 4.2, Bluetooth Low Energy (BLE)
GPIO	40-pin GPIO header, populated
Storage	microSD
Ports	HDMI, 3.5mm analogue audio-video jack, 4x USB 2.0, Ethernet, Camera Serial Interface (CSI), Display Serial Interface (DSI)
Dimensions	82mm x 56mm x 19.5mm, 50g
Operating System	Raspbian GNU/Linux 10 (buster)
Kernel	Linux 4.19.66-v7+

4.2. Algoritmo

Ad oggi lo script di sniffing è composto da due thread separati, uno per lo sniffing e la computazione del pacchetto sniffato (ThreadSniffing) e uno per l'invio periodico dei dati al database (ThreadInvio).

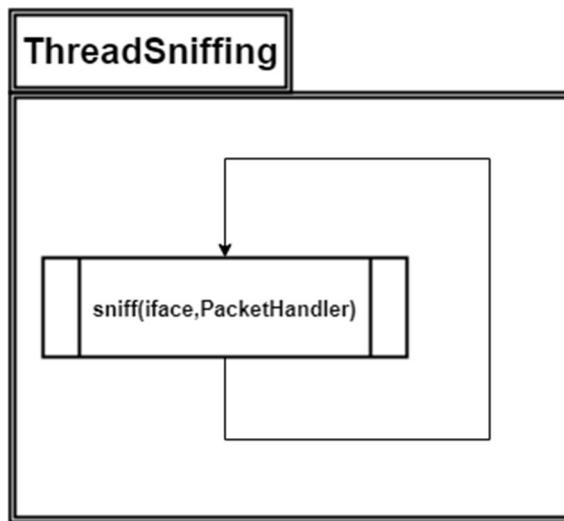
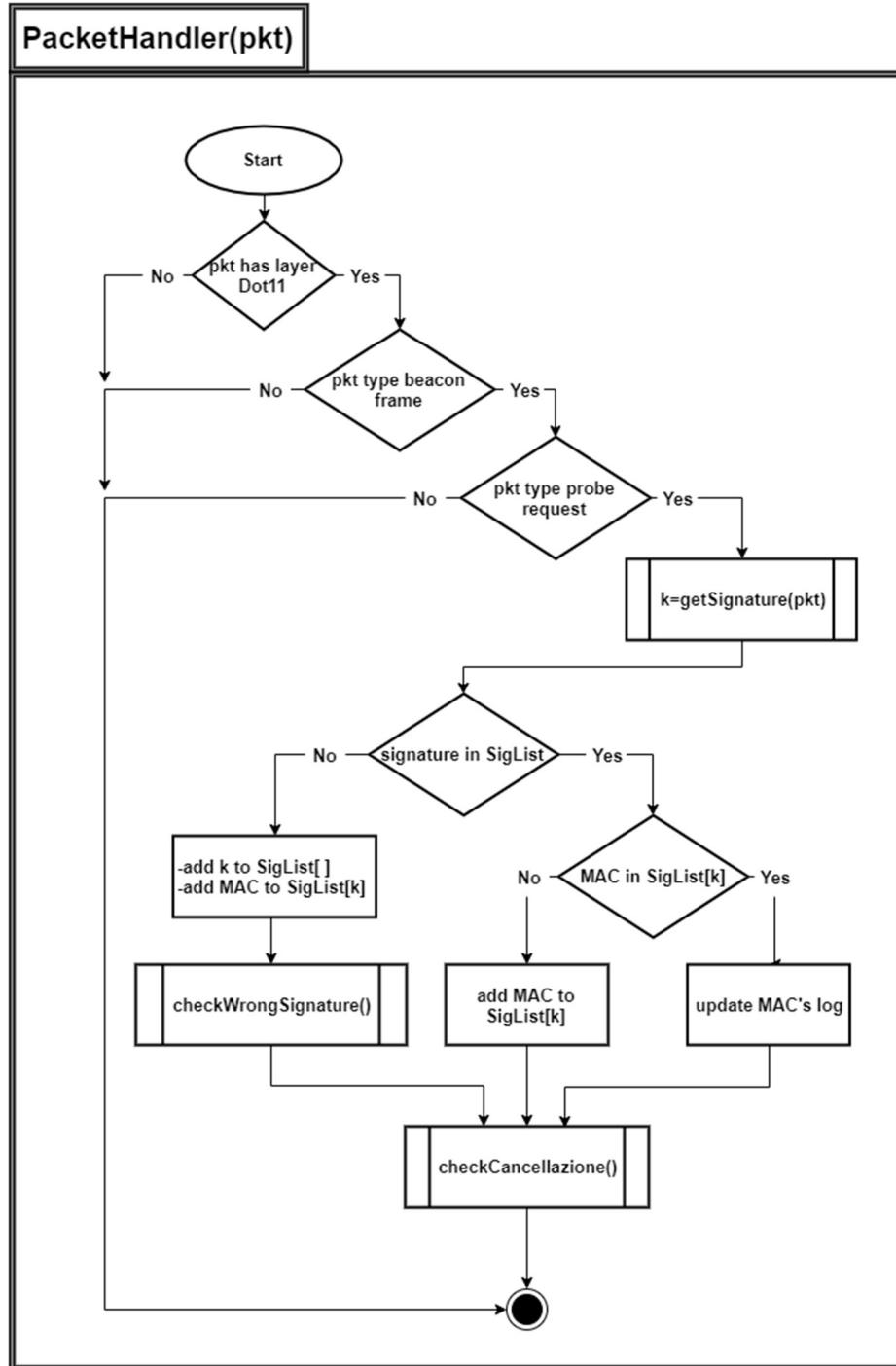


Figura 14 Activity diagram di ThreadSniffing

Il thread in figura è un processo che esegue in loop la funzione sniff fornita dalla libreria scapy. Questa ha bisogno di un parametro che indichi l'interfaccia di rete con cui ascoltare il traffico e di un secondo parametro, una funzione da eseguire dopo aver ricevuto il pacchetto.

L'handler che viene invocato ad ogni ricezione di un pacchetto ha come scopo quello di generare una chiave a partire dalle informazioni contenute nelle probe request, chiave la cui funzione è quella di discriminare le sorgenti delle richieste. Il secondo grande step è quello di associare l'indirizzo MAC contenuto nella richiesta alla chiave appena generata. L'associazione chiave-MAC è di tipo uno-a-molti ed è proprio tra indirizzi aventi la stessa chiave che viene effettuato il test di aggregazione. Non avrebbe senso cercare di aggregare due MAC randomizzati se questi arrivano da due dispositivi diversi.

Di seguito l'Activity diagram di PacketHandler.



Un thread separato, è dedicato all’invio periodico dei dati raccolti tramite la rete al database ospitato su un calcolatore interno all’università, adibito a web server tramite uno stack LAMP. La scrittura può essere vista come una fotografia del sistema relativa all’istante di tempo dedicato all’invio dei dati.

4.2.1. Fingerprinting

L’idea di base dello script è quella di dividere le probe request in gruppi che stiano a identificare le diverse marche e modelli di dispositivi. Per quello che riguarda i MAC globali sniffati non vi è alcun problema, ma per i MAC randomici bisogna verificare se questi debbano essere aggregati. Tramite la libreria scapy, il sistema si mette in ascolto dei pacchetti in transito e recupera le informazioni in essi contenute al fine di costruire una *signature* che rappresenta la chiave che distingue i gruppi. A questa viene poi associato il MAC address contenuto nel pacchetto da cui era stata generata. La signature può essere quindi considerata come una chiave derivata dalle caratteristiche fisiche di ciascun dispositivo [35], [38], le quali vengono trasmesse all’interno del pacchetto. Questo comporta che eguali dispositivi abbiano signature coincidenti.

Come precedentemente studiato in [35], le probe request includono nel loro corpo dati sotto forma di *Information Elements* (IE) anche chiamati *parametri taggati* o più semplicemente *tags*. Questi IE non sono obbligatori ma vengono inviati dai dispositivi per segnalare agli AP le proprie caratteristiche fisiche. Ad ogni IE corrisponde un ID e un campo per i valori. Siccome sono opzionali, non tutti gli IE sono inclusi da tutti i dispositivi e di conseguenza il set di IE trasmessi varia da device a device. Sulla base della letteratura, che ci ha aiutato a capire quali fossero gli elementi più significativi e discriminanti, sono stati oggetto di nostre attenzioni i seguenti tag:

Information Element	ID
Service Set Identity (SSID)	0
Supported Rates	1
Power Capabilities	33
HT Capabilities	45
Extended Supported Rates	50
Extended Capabilities	127
VHT Capabilities	191
Vendor Id	221

Tabella 5 Elenco IE di interesse

Nonostante lo standard 802.11 definisca che i tag debbano essere dichiarati in ordine crescente, molti dispositivi ignorano quanto stabilito e li trasmettono con un ordine scelto a loro discrezione. Per questo motivo, anche l'ordine in cui compaiono gli IE può essere utilizzato come potenziale fonte di informazione.

4.2.2. Aggregazione

All'interno delle Probe requests, ad accompagnare ogni MAC, vi sono informazioni utili alla derandomizzazione come timestamp, RSSI e sequence number della richiesta che vengono collezionate dallo script in un log da associare all'indirizzo MAC. L'esecuzione del processo porta ad avere una struttura dati in cui ad una signature possono essere associati più indirizzi MAC e ad un indirizzo MAC uno storico di log. Come già accennato, se una signature contiene molteplici MAC randomici, allora tramite uno studio sui log si cerca di capire se questi siano stati generati dallo stesso dispositivo o meno.

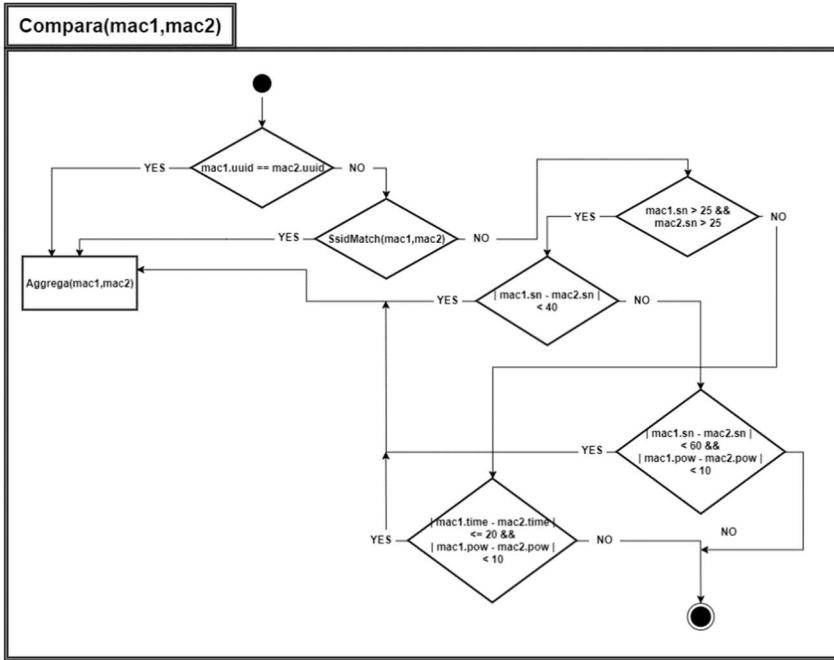


Figura 15 Activity diagram della funzione Compara. Questa funzione viene usata per stabilire se due indirizzi MAC randomici siano stati generati dallo stesso dispositivo.

I criteri per l'aggregazione sono basati sulle differenze tra i log di due MAC differenti per i valori sopracitati:

1. piccole variazioni di sequence number;
 2. medie variazioni di sequence number e potenza;
 3. medie variazioni di timestamp e piccole di potenza;

Seppur rari, vi sono casi in cui un dispositivo può trasmettere le informazioni WPS (Wi-Fi Protected Setup), protocollo supportato da alcuni dispositivi che permette una connessione sicura con un AP anche a dispositivi non autenticati, da cui si ricava l'UUID (Universally Unique Identifier) che, come scoperto in [35], viene generato direttamente dal MAC globale. In questi casi è sufficiente confrontare gli UUID per stabilire se i due MAC randomici provengano dallo stesso dispositivo. Un altro caso particolare si ha quando le richieste inviate dai dispositivi non sono di tipo broadcast ma indirizzate a specifici AP, quindi contenenti gli SSID degli hotspot con cui vorrebbero stabilire una connessione, conosciuti in precedenza. In una situazione del genere, liste di AP uguali portano all'aggregazione dei due MAC random. Una volta trovate le condizioni per cui è possibile l'aggregazione, al MAC con il log più recente vengono aggiunti i log dell'altro, che viene etichettato come "da cancellare".

4.3. Caratteristiche dello strumento

4.3.1. Modello del problema

Lo scopo dello strumento è fornire in output un numero indicante la stima delle persone rilevate, in modo da classificare lo stato del traffico secondo un numero discreto di livelli. Note le informazioni riguardo la randomizzazione degli indirizzi MAC, la misura registrata può essere rappresentata come

$$\tilde{x} = (x_1 + x_2)$$

dove x_1 rappresenta il numero di MAC globali rilevati e x_2 il numero di MAC randomizzati rilevati già aggregati, ad indicare il numero di dispositivi unici.

La misura in uscita dagli sniffer però è una misura rumorosa le cui fonti di disturbo sono da ritrovarsi in:

- accuratezza di x_2 , che potrebbe sovrastimare o sottostimare il numero di dispositivi unici;
- dispositivi con modulo Wi-Fi spento e quindi non rilevabili;
- dispositivi con Wi-Fi acceso ma non rilevabili;
- portata massima dello strumento o comunque limiti dettati dall'architettura.

Dato il grande numero di variabili e la mancanza di una base statistica sufficiente per risolvere un problema di identificazione ben fatto, per il momento si è optato per un semplice modello proporzionale piuttosto che per la realizzazione di un modello di complessità troppo elevata.

$$\tilde{y} = \alpha \tilde{x}$$

In un'ottica di questo tipo, in cui il risultato finale andrà a definire semplicemente una classe di viabilità, si ritiene che l'identificazione del parametro α mediante least squares sia sufficiente ad avere una stima verosimile delle persone presenti

$$\min_{\alpha} (y - \tilde{y})^2$$

con y ground truth, numero reale di persone presenti.

4.3.2. Accuratezza nell'aggregazione

Per testare il comportamento dello script di aggregazione è necessario un ambiente totalmente controllato, in cui posizionare un numero predefinito di dispositivi che effettuino la randomizzazione e verificare che in output venga fornito lo stesso numero. In questo modo però non si avrebbe la certezza che vengano effettuate tutte le aggregazioni in modo corretto, sarebbe disponibile solo “cosa” e non il “come”. Siccome non si conoscerebbe la verità riguardo la sequenza di MAC adottati da ogni singolo device e vista la mancanza di un ambiente completamente schermato, il test effettivamente eseguito ha visto l’impiego di cinque diversi laptop provvisti di una versione di Kali Linux in live boot, grazie alla quale si è potuto randomizzare manualmente. Su ciascun pc girava uno script che ogni quindici secondi impiegava il comando *macchanger -r* per randomizzare il proprio indirizzo MAC e riportava su un file di log la sequenza di indirizzi adottati. Allo stesso tempo, lo script di sniffing è stato modificato per includere una scrittura su file ogni volta che venisse presa la decisione di aggregazione, indicando quali indirizzi fossero stati aggregati. Il dataset così ottenuto è composto da cinque file distinti aventi la struttura

Current MAC:	b2:49:45:aa:23:48
New MAC:	d6:78:bb:9c:15:e1
Current MAC:	d6:78:bb:9c:15:e1
New MAC:	5e:0e:54:91:93:b1

e da un file di log con tutte le aggregazioni effettuate, nel medesimo formato. I risultati ottenuti sono stati soddisfacenti: su quasi 1000 randomizzazioni realmente effettuate dai laptop, lo script ne ha registrate circa l’80%.

Laptop	Aggregazioni rilevate	Accuratezza
Laptop1	133/191	69.6%
Laptop2	134/183	74.1%
Laptop3	165/196	84.2%
Laptop4	129/142	90.8%
Laptop5	170/222	76.6%

Totale	733/934	78.5%
---------------	----------------	--------------

Tabella 6 Risultati test di aggregazione

4.3.3. Penetration rate

Dopo aver trattato l'accuratezza sull'aggregazione, una seconda fonte di rumore è rappresentata dal tasso di moduli Wi-Fi accesi rispetto alla popolazione da monitorare, ovvero il penetration rate. In altri termini, quante persone mediamente hanno il modulo wireless del cellulare attivato e sono rilevabili dal sensore. Come già ampiamente trattato, non importa che il dispositivo sia connesso ad una rete, l'importante è che questo esegua active scanning.

L'ottenimento della cosiddetta *ground truth* riguardo il rate è piuttosto complesso perché che varia a seconda dell'età della popolazione, del luogo di monitoraggio, del contesto e possibilmente da altri fattori non menzionati. Noto il rate tra dispositivi rilevati e ground truth, si può stimare il parametro proporzionale da applicare a seconda della situazione.

4.3.4. Portata massima

La portata massima dell'apparecchio è l'ultima fonte di rumore che verrà studiata. Una misura teorica della portata massima la si può ottenere supponendo che vi siano un numero infinito di persone nel raggio d'azione del sensore. Siccome i dispositivi vengono scartati per vecchiaia dopo che non risultano più rilevazioni per un certo periodo di tempo, la domanda a cui si vuole trovare risposta è: di queste infinite persone, quante di queste vengono rilevate contemporaneamente?

Per rispondere al quesito si è calcolato il tempo medio di esecuzione di un ciclo conseguentemente alla rilevazione di una probe request che, in condizioni di medie affollamento (100 mac address unici in memoria), risulta essere in media di circa 15,8 ms. Per avere un'idea più concreta, questo risultato si traduce nella capacità di conteggiare 4.000 persone nell'arco di un tempo di decadimento di 60 secondi, ovviamente a patto di rilevare richieste sempre diverse.

Una misura pratica invece, la si otterrebbe andando a misurare il traffico immersendo realmente il sensore in mezzo a migliaia di persone per vedere quanti

indirizzi questo riesca a raggiungere contemporaneamente. Purtroppo, una prova pratica non è ancora stata possibile, l'unico dato concreto si ha dai dati registrati fino a ora alle Cinque Terre, in cui i picchi superano i 200 dispositivi.

4.4. Autonomia delle stazioni

Data la natura del progetto, il sistema deve essere in grado di funzionare autonomamente, senza che ci sia il costante bisogno di mantenere manualmente attivo lo script. L'unica accortezza è quella di alimentare i dispositivi e assicurarsi che siano connessi ad internet per l'invio dei dati. In caso di cali di alimentazione, crash inaspettati o interruzioni della rete, la stazione dovrebbe essere in grado di intraprendere azioni correttive.

4.4.1. RAM

Un ostacolo all'autonomia delle stazioni arriva dallo script di sniffing. Questo purtroppo consuma parecchia della poca RAM a disposizione dei Raspberry. La soluzione ottimale sarebbe stata quella di alleggerire il carico computazionale degli sniffer demandando a loro esclusivamente le operazioni di sniffing e invio dei pacchetti al database. Tutto l'onere della computazione lo avrebbe dovuto avere un'applicazione lato server in modo che la potenza limitata dei Raspberry non fosse più un problema.

Dopo aver effettuato lo spostamento della computazione, sono emersi errori nella ricostruzione del pacchetto una volta inviato al server. Per poter scrivere su database il pacchetto grezzo, questo veniva convertito dal tipo `scapy.packet` al tipo `stringa`. L'operazione inversa non è stata possibile a causa della perdita di informazioni durante i passaggi di tipo che, nonostante l'aiuto del correlatore Dott. Ing. Luca Zero, che ringrazio per l'interessamento alla causa, non siamo riusciti a riottenere per via della scarsa documentazione della libreria `scapy`.

A questo punto, i requisiti del progetto sono cambiati in aumentare l'autonomia dei dispositivi cercando di impiegare efficientemente l'utilizzo della memoria principale.

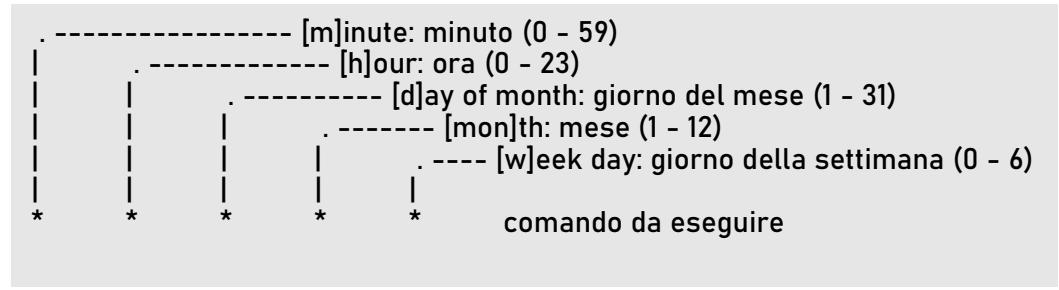
Le modifiche inizialmente vedevano l'implementazione di un altro thread indipendente di cancellazione che periodicamente eliminasse la struttura dati contenente tutte le informazioni dello script e facesse utilizzo manuale del garbage collector di Python. Dopo un'analisi del processo non si notavano significativi svuotamenti di memoria nel momento in cui il thread di cancellazione veniva eseguito, quindi si è deciso di non utilizzare il thread di cancellazione e di migliorare il modo in cui le entries della struttura dati venivano eliminate.

Al termine di una fase di riorganizzazione del codice, questo ha ottenuto piccoli miglioramenti sulla durata dell'esecuzione dello script che raggiungeva circa le quattro/sei ore ma che comunque necessitava in qualche modo di essere riavviato per ottenere la completa indipendenza, soprattutto in zone congestionabili dove la mole di dati sarebbe stata sicuramente più elevata.

4.4.2. Crontab

In questo scenario di indipendenza si è utilizzato il demone **Cron**, un processo di **Unix** che esegue istruzioni o comandi in determinate date e orari, predefinite dall'utente. **Cron** può gestire i comandi, “**jobs**”, in modo da eseguirli una sola volta alla prima occorrenza della data/ora prescelta, ma è anche possibile fargli eseguire ripetutamente i comandi, per esempio nel caso in cui si voglia eseguire una determinata operazione ogni giorno in una data ora, e in questo caso viene usato il ‘**Crontab**’. **Crontab** usa un demone chiamato **Crond**, un processo costantemente in esecuzione in background, che una volta al minuto, legge i contenuti del registro comandi pianificati ed esegue quelli per cui si è esaurito il periodo di attesa. I comandi mandati in esecuzione da **Crontab** vengono chiamati **Cronjob**. Da crontab si possono specificare anche comandi da eseguire in determinate situazioni, ad esempio ad ogni reboot del dispositivo.

Per schedulare un job si può ricorrere a cinque unità di misura:

*Equazione 2 Struttura di un cronjob*

in cui:

- L'operatore virgola , specifica una lista di valori, ad esempio: 1,3,4,7,8.
- L'operatore trattino - specifica un intervallo di valori, ad esempio: 1-6, che equivale a 1,2,3,4,5,6.
- L'operatore asterisco * specifica tutti i possibili valori di un campo. Ad esempio, un asterisco nel campo dell'ora è equivalente a «ogni ora»

Per esempio, per eseguire myScript.py ogni ora tra le 10:00 e le 22:00 in ogni giorno si utilizzerà:

```
0 10-21 * * * python myScript.py
```

Equazione 3

che si traduce in: ogni minuto 0 di ogni ora compresa tra le 10 e le 21 (compresa) di ogni giorno del mese, di ogni mese, in ogni giorno della settimana, esegui myScript.py. Come si può intuire l'unità di misura più piccola che si può controllare è il minuto e per eseguire un cronjob in un intervallo di n minuti, diciamo 30, si può utilizzare l'operatore / posizionando nello slot dei minuti: */30.

Per maggiori spiegazioni riguardo il crontab consultare [39] o [40].

Crontab è stato impiegato per far fronte al problema della memoria principale menzionato nel paragrafo precedente. Schedulando come cronjob lo script di sniffing ed eseguendolo ogni ora con un timeout di 59 minuti, la ram viene svuotata del tutto e lo strumento è in grado di rimanere attivo autonomamente idealmente per sempre. Nel caso di qualsivoglia crash inaspettato del codice, questo riprenderà a funzionare

allo scattare dell'ora successiva. Per quanto riguarda la mancanza di alimentazione, al crontab è stato aggiunto uno script shell che faccia passare in modalità monitor la scheda wireless ad ogni reboot.

4.4.3. Monitor mode

È bene precisare che il Raspberry è in possesso di una scheda wireless integrata che però non può essere messa in monitor mode. Da qui la necessità di una scheda esterna, notevolmente migliore anche per quello che riguarda la portata.

Le schede di rete, nell'utilizzo quotidiano, si trovano nella cosiddetta “managed mode”. Questa modalità permette loro di ricevere i pacchetti che le vedono come destinatarie e di scartare quelli di cui non lo sono. La modalità monitor invece fa sì che non venga scartato alcun pacchetto, permettendo alla scheda, come suggerisce il termine, monitorare la rete, ethernet o wireless che sia.

Prima di passare a commentare la modalità monitor della scheda wireless è bene fare ancora un po' di chiarezza sulle periferiche di rete.

- Un sistema Linux-based mostra le periferiche wireless come wlanX e le periferiche ethernet come ethX dove X corrisponde ad un indice intero crescente ≥ 0 utilizzato per distinguere le interfacce.
- Lo script di sniffing ha bisogno che gli venga passato come parametro il nome dell'interfaccia di rete con la quale sniffare il traffico.
- L'interfaccia da passare deve essere dapprima identificata mediante il suo nome e successivamente essere messa in monitor mode, che tra le altre cose comporta l'aggiunta del suffisso “mon” al suo nome.

Con il comando shell ifconfig, un sistema Linux-based mostra le interfacce di rete. Il comando iwconfig è simile a ifconfig ma è dedicato alle interfacce wireless.

```
pi@raspberrypi:~ $ iwconfig  
eth0      no wireless extensions.  
  
lo       no wireless extensions.  
  
wlan0    IEEE 802.11 ESSID:off/any  
          Mode:Managed Access Point: Not-Associated Tx-Power=31 dBm  
          Retry short limit:7 RTS thr:off Fragment thr:off  
          Power Management:on  
  
wlan1    IEEE 802.11 ESSID:off/any  
          Mode:Managed Access Point: Not-Associated Tx-Power=20 dBm  
          Retry short long limit:2 RTS thr:off Fragment thr:off  
          Power Management:off
```

Equazione 4 output di iwconfig

Come si può notare dal precedente output, entrambe le schede di rete wireless collegate al Raspberry (quella esterna usb e quella integrata sulla scheda madre) hanno come dicitura Mode:Managed, che è la modalità classica per potersi collegare ad una rete e navigare in internet.

Ai fini del progetto, è necessario cambiare modalità alla scheda wireless esterna e metterla in monitor mode. Per farlo è stata utilizzata la suite aircrack-ng. Supponendo che wlan0 si riferisca alla scheda integrata sulla scheda madre e wlan1 alla scheda esterna segue:

```

pi@raspberrypi:~ $ sudo ifconfig wlan1 up
pi@raspberrypi:~ $ sudo airmon-ng start wlan1

PHY Interface Driver Chipset
phy0 wlan0 brcmfmac Broadcom 43430
phy1 wlan1 rt2800usb Ralink Technology, Corp. RT5370

(mac80211 monitor mode vif enabled for [phy1]wlan1 on [phy1]wlan1mon)
(mac80211 station mode vif disabled for [phy1]wlan1)

pi@raspberrypi:~ $ iwconfig
eth0 no wireless extensions.

wlan1mon IEEE 802.11 Mode:Monitor Frequency:2.457 GHz Tx-Power=20 dBm
    Retry short long limit:2 RTS thr:off Fragment thr:off
    Power Management:off

lo no wireless extensions.

wlan0 IEEE 802.11 ESSID:off/any
    Mode:Managed Access Point: Not-Associated Tx-Power=31 dBm
    Retry short limit:7 RTS thr:off Fragment thr:off
    Power Management:on

```

Equazione 5 iwconfig dopo aver messo la scheda in monitor mode

Si noti che wlan1 è passata in Mode:Monitor e ora si chiama wlan1mon. Siccome l'indice si manifesta in ordine crescente, inizialmente si pensava che alla scheda esterna venisse sempre assegnato l'identificatore wlan1. Di conseguenza, all'interno dello script di sniffing il nome dell'interfaccia con cui sniffare era inizialmente hardcoded e sempre lo stesso: "wlan1mon" perché lo script shell, al reboot, eseguiva la procedura su wlan1. Facendo girare il Raspberry per diversi giorni, si è scoperto in realtà che non venivano assegnati sempre gli stessi

identificatori alle interfacce di rete ma capitava che venissero invertiti. Dal momento che il parametro era hardcoded, poteva non esserci corrispondenza tra questo e l'identificativo assegnato dal sistema operativo alla scheda esterna. Lo script shell eseguito al reboot, tentando di mettere in monitor mode wlan1, avrebbe potuto eseguire l'operazione sulla scheda integrata che, non potendo essere messa in monitor mode, non avrebbe permesso la generazione di wlan1mon e lo script di sniffing sarebbe andato incontro ad un errore.

La prima soluzione proposta ha come fine quello di disabilitare la scheda wireless integrata sulla scheda madre del Raspberry, in modo che l'unica abilitata e visibile rimanga quella esterna e di conseguenza prenda sempre il nome wlan0. Non essendoci più altre schede wireless in competizione per l'assegnazione dei nomi, il parametro da passare alla funzione di sniff sarà sempre hardcoded ma questa volta “wlan0mon”.

Per disabilitare il WiFi integrato è necessario modificare un file di configurazione del sistema operativo raggiungibile in */etc/modprobe.d/raspi-blacklist.conf* aggiungendo le seguenti righe:

```
blacklist brcmfmac
blacklist brcmutil
```

Equazione 6 /etc/modprobe.d/raspi-blacklist.conf

Una seconda soluzione ha come fine quello di provare a mettere in monitor mode una delle due interfacce tramite il sopracitato script shell che si esegue ad ogni reboot del dispositivo. Se ifconfig mostra una scheda in monitor mode allora si è eseguita l'operazione sulla scheda corretta e si passa quell'identificativo alla funzione. Se ifconfig non mostra nessuna scheda in monitor mode allora si ripete la stessa operazione sulla seconda interfaccia che sicuramente sarà quella corretta. Lo script in python deve essere esteso con il modulo netifaces che permette di vedere le interfacce di rete e con un semplice ciclo che salva il nome più lungo che riesce a trovare tra queste.

```
@reboot sudo sh monitor-routine.sh
```

Equazione 7 Script shell da eseguire al riavvio all'interno del crontab

```
sudo ifconfig wlan0 up
sudo airmon-ng start wlan0
T2=$(ifconfig | grep "wlan0mon" | cut -d ':' -f1)
#se la scheda era wlan0
if [${T2}="wlan0mon"]
then
    echo "wlan0mon activated"
else
    sudo ifconfig wlan1 up
    sudo airmon-ng start wlan1
    echo "wlan1mon activated"
fi
```

```
import netifaces
x=netifaces.interfaces()
iface=""

for interface in x:
    if(len(interface)>len(iface)):
        iface=interface
```

Equazione 9 Codice aggiuntivo nello script di sniffing in python

Equazione 8 monitor-routine.sh

Siccome nell'ambiente descritto esistono solo wlan0 e wlan1 e soltanto una delle due verrà rinominata in wlanXmon, allora iface conterrà l'identificativo corretto da passare alla funzione di sniff.

4.5. Database

Il progetto si appoggia ad un sistema Debian GNU/Linux 9 (stretch) su cui è stato opportunamente installato uno stack LAMP. Il server ospita il database e l'applicazione web descritta nel sottocapitolo successivo.

Il database è composto da quattro tabelle due delle quali sono riempite dagli sniffer piazzati nelle aree di monitoraggio.

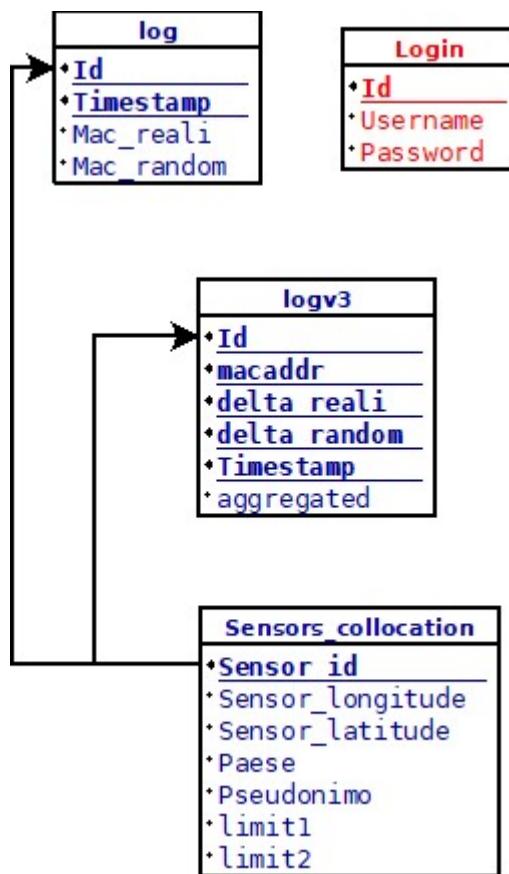


Figura 16 Struttura generale del database

Login

Da specifiche iniziali, l'applicazione avrebbe dovuto consentire la visualizzazione dei dati solamente agli utenti autenticati del parco. Come questi otterranno le credenziali non è ancora stato specificato. In futuro potrebbero essere presenti funzionalità variabili a seconda del livello di accesso o le credenziali

potrebbero essere le stesse per tutti. In ogni caso la tabella è già predisposta e pronta all'utilizzo.

- **Id:** chiave primaria. Indice intero che si autoincrementa;
- **Username:** username;
- **Password:** password.

Field	Type	Null	Key	Default	Extra
Id	int(11)	NO	PRI	NULL	auto_increment
Username	varchar(24)	NO		NULL	
Password	varchar(52)	NO		NULL	

Figura 17 Struttura tabella Login

Sensors_collocation

In Sensors_collocation sono riportati i dati distintivi degli sniffer. Ogni riga è aggiunta manualmente ogni volta che il sistema viene esteso con un nuovo sensore sniffer.

- **Sensor_id:** chiave primaria. Identificatore univoco intero assegnato ad ogni sniffer;
- **Sensor_longitude:** coordinata longitudinale della posizione dello sniffer;
- **Sensor_latitude:** coordinata latitudinale della posizione dello sniffer;
- **Paese:** nome del paese delle Cinque Terre a cui le coordinate si riferiscono;
- **Pseudonimo:** pseudonimo dello sniffer per facilitarne la lettura sulla mappa e la distinzione tra possibili sniffer localizzati all'interno dello stesso paese;
- **limit1:** limite di persone secondo il quale l'area sulla mappa passa da verde a gialla;
- **limit2:** limite di persone secondo il quale l'area sulla mappa passa da gialla a rossa.

Field	Type	Null	Key	Default	Extra
Sensor_id	int(11)	NO	PRI	NULL	
Sensor_longitude	decimal(7,6)	NO		NULL	
Sensor_latitude	decimal(8,6)	NO		NULL	
Paese	varchar(32)	NO		NULL	
Pseudonimo	varchar(32)	NO		NULL	
limit1	int(11)	NO		NULL	
limit2	int(11)	NO		NULL	

Figura 18 Struttura tabella Sensors_collocation

log

Log è la prima tabella che i sensori popolano durante l'esecuzione dello script di sniffing. Ciclicamente vengono inviati i dati relativi a quanti dispositivi sono presenti nel raggio d'azione del sensore riferiti a quell'istante di tempo. I dati che vengono inviati sono:

- **Id:** parte di chiave primaria composta. Indica l'identificativo del sensore da cui provengono i dati della tupla;
- **Timestamp:** seconda parte di chiave primaria composta. Indica l'istante di tempo a cui si riferiscono i dati inviati;
- **Mac_reali:** numero di dispositivi con MAC address globale in memoria in quell'istante;
- **Mac_random:** numero di dispositivi con MAC address randomizzato in memoria in quell'istante. Il dato si riferisce già al numero fornito in output dal processo di aggregazione.

Field	Type	Null	Key	Default	Extra
Id	int(11)	NO	PRI	NULL	
Timestamp	timestamp	NO	PRI	CURRENT_TIMESTAMP	
Mac_reali	int(11)	NO		NULL	
Mac_random	int(11)	NO		NULL	

Figura 19 Struttura tabella log

logv3

L'ultima tabella del database è nuovamente riempita dai sensori e contiene l'informazione relativa all'aggiunta e rimozione di ogni singolo MAC address rilevato dal set degli indirizzi in memoria in un dato istante. In caso di sviluppi futuri, se si volessero costruire gli itinerari percorsi dai turisti rilevati di stazione in stazione, grazie a questa tabella si potrebbero identificare le prime e ultime rilevazioni di uno stesso indirizzo in ogni punto di interesse.

- **Id:** parte di chiave primaria composta. Indica l'identificativo del sensore da cui provengono i dati della tupla;
- **macaddr:** parte di chiave primaria composta. Indica l'indirizzo MAC in ingresso o uscita dal set attualmente in memoria;
- **delta_reali:** parte di chiave primaria composta. Indica la variazione del numero di MAC globali conseguente all'ingresso/uscita di macaddr. Valori ammessi -1, 0, +1.
- **delta_random:** parte di chiave primaria composta. Indica la variazione del numero di MAC randomizzati conseguente all'ingresso/uscita di macaddr. Valori ammessi -1, 0, +1.
- **Timestamp:** parte di chiave primaria composta. Indica l'istante di tempo in cui l'ingresso/uscita è stata rilevata;
- **aggregated:** riferito ad un indirizzo randomizzato. Indica se macaddr sia uscito dal set di indirizzi in memoria perché aggregato ad un altro (valore 1) o se perché trascorso il timeout per cui l'indirizzo viene considerato vecchio (valore 0).

Field	Type	Null	Key	Default	Extra
Id	int(11)	NO	PRI	NULL	
macaddr	varchar(32)	NO	PRI	NULL	
delta_reali	int(11)	NO	PRI	NULL	
delta_random	int(11)	NO	PRI	NULL	
Timestamp	datetime	NO	PRI	NULL	
aggregated	int(11)	NO		NULL	

Figura 20 Struttura tabella logv3

Siccome ci sono molteplici sensori che popolano questa tabella, l'unico modo per garantire una chiave primaria univoca è la combinazione dei primi cinque campi.

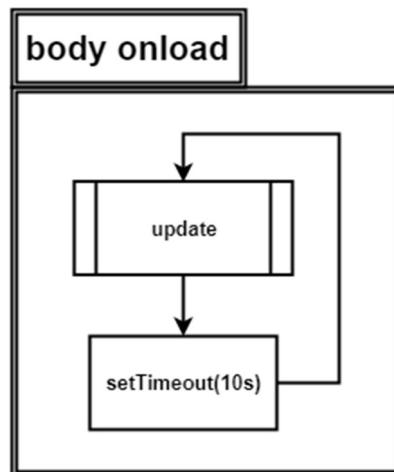
4.6. Applicazione Web

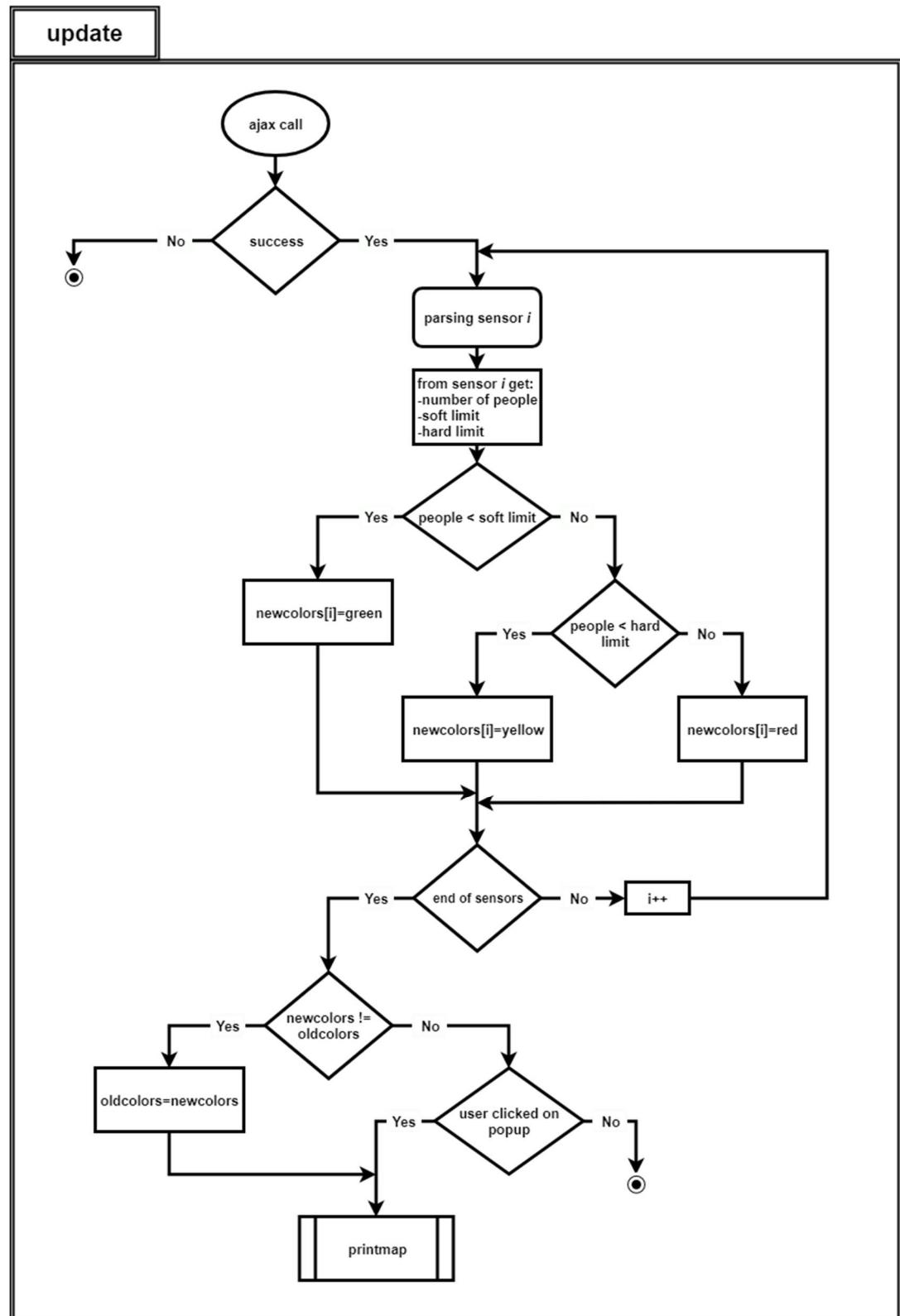
Per riprendere il discorso riguardo l'intasamento dei luoghi congestionabili e il monitoraggio in tempo reale, è stata sviluppata un'applicazione web che permettesse la visualizzazione dei dati in maniera grafica e intuitiva, direttamente su una mappa con marker in corrispondenza delle coordinate geografiche in cui i dispositivi vengono posizionati.

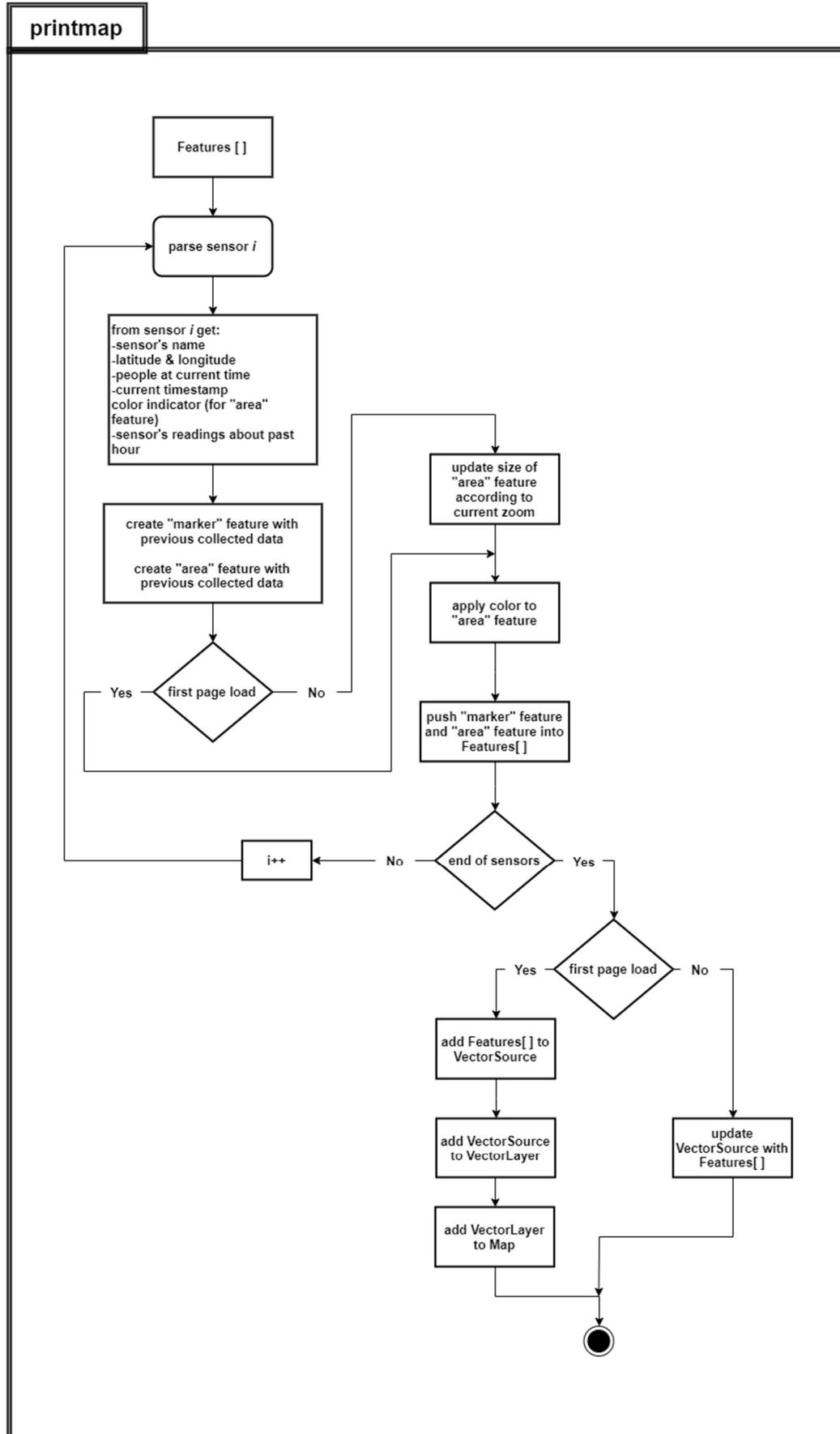
La console di visualizzazione utilizza OpenLayers 3 e OpenStreetMap per costruire la mappa con i marker cliccabili. Attorno ad ogni marker vi è un'area colorata coerentemente con la densità di persone attualmente percepite da quel sensore (verde, giallo, rosso).

Tramite chiamate ajax vengono richiesti periodicamente i dati al database e la mappa si modifica in caso vi siano cambiamenti di stato. Lo stato è quell'insieme di condizioni che determinano i colori di ogni sensore presente sulla mappa. Di conseguenza, nel caso in cui presso uno dei sensori si verifichi una congestione improvvisa, l'utente vedrà cambiare il colore dell'area interessata automaticamente.

Oltre l'aggiornamento periodico, la funzione di update può anche essere invocata manualmente quando uno dei marker indicanti i sensori vengono cliccati per mostrare i dati aggiornati.







5. Applicazione al caso di studio

5.1. Le Cinque Terre



Figura 21 Overview del Parco Nazionale delle Cinque Terre

Il Parco Nazionale delle Cinque Terre è un frastagliato tratto di costa della Riviera ligure di levante situato nel territorio della provincia di La Spezia tra Punta Mesco e Punta di Montenero, nel quale si trovano cinque borghi, da ovest verso est: Monterosso al Mare, Vernazza, Corniglia, Manarola e Riomaggiore. Nel 1997 l'UNESCO ha inserito le Cinque Terre nella lista del Patrimonio Mondiale dell'Umanità come "paesaggio culturale".

Gran parte dei parcheggi sono riservati ai residenti e in ragione della mancanza di spazi adeguati, la discesa ai borghi è possibile solo a piedi. Esiste tuttavia un servizio di bus elettrici gestito dal Parco Nazionale che nei periodi di maggior afflusso turistico garantisce i collegamenti. Nella stagione turistica un servizio di battelli di linea collega tutti i borghi ad eccezione di Corniglia (che non è situata sulla costa) con Porto Venere, a sua volta collegata con La Spezia, Lerici e con gli altri paesi della riviera di Levante. Data la situazione geografica piuttosto impervia dei borghi, il mezzo di trasporto di gran lunga più utilizzato per raggiungerli è la ferrovia che fa tappa in tutte le terre e vede una notevole frequenza di transiti. Una volta nei pressi dei paesi, esistono due sentieri principali che i turisti percorrono per visitare il Parco. Il primo (sentiero numero 1) è conosciuto come Alta via delle Cinque Terre e collega Levanto con Portovenere. Permette di percorrere un tragitto più elevato e

5. Applicazione al caso di studio

nell'entroterra rispetto al secondo ma è attraversato da molti sentieri trasversali che consentono di raggiungere la costa. Il secondo sentiero (sentiero numero 2), meglio conosciuto come Sentiero Azzurro, è quello più apprezzato dai turisti. Collega le Cinque Terre con un percorso sul litorale e regala parecchi scorci panoramici. Visto il suo successo, l'accesso è regolato tramite l'acquisto della Cinque Terre Card.

A causa della storicità dei sentieri e per motivi di sicurezza, alcuni tratti vengono spesso chiusi/manutenuti secondo le necessità più o meno impreviste. [41]

5.2. Collocamento degli sniffer

Il passo fondamentale per il corretto funzionamento dell'applicazione web è stato quello dell'individuazione di spot idonei al posizionamento dei dispositivi. I luoghi di interesse possono essere piazze, stazioni, o passaggi che si comportino come collo di bottiglia per il flusso pedonale e che quindi sarebbe bene tenere sotto controllo.

Siccome i dati devono essere disponibili su un database online per poter essere visualizzati dall'applicazione, è opportuno che gli sniffer abbiano l'accesso ad Internet, tramite LAN o tramite GPRS. Siccome il progetto ha visto in questo periodo la sua prima vera realizzazione, in accordo con l'Ente Parco, i luoghi che inizialmente sono stati individuati per la messa in atto corrispondono alle cinque stazioni ferroviarie dei cinque paesi, o meglio ai 5 Terre Point in corrispondenza delle stazioni. Oltre a consentire una visione globale delle cinque frazioni, questa strategia ci ha permesso di connettere i dispositivi tramite LAN agevolando il processo di connessione ad Internet.

Visto che il progetto in sé sarebbe stato, ed è tutt'ora, un esperimento per determinare la fattibilità di monitoraggio e di predizione, una volta piazzati i dispositivi questi avrebbero dovuto essere accessibili tramite reti esterne in modo tale da poter effettuare aggiornamenti software senza dover accederci da una rete locale, quindi senza effettivamente recarsi in loco. Per sopperire a questo problema, ogni dispositivo è stato registrato sul sito remot3.it [42] ed equipaggiato con una versione dell'applicazione dedicata che permette di eseguire il reverse ssh tunneling da qualsiasi rete esterna.

Nei prossimi paragrafi seguiranno le posizioni nel dettaglio di ogni singolo dispositivo installato.

5.2.1. Monterosso

Monterosso è la prima stazione raggiungibile se si viaggia da nord verso sud ed è anche una delle località più trafficate assieme a Riomaggiore in quanto, a differenza dei borghi interni, il primo e l'ultimo sono soggetti ad un numero maggiore di soste da parte dei treni regionali.



Figura 22 Stazione di Monterosso



Figura 23 5 Terre Point di Monterosso - esterno



Figura 24 5 Terre Point di Monterosso - interno



Figura 25 Passeggiata di Monterosso

Come si può notare dalle immagini, il dispositivo è posizionato in modo tale da poter ascoltare il traffico sia all'interno della stazione che lungo la passeggiata sottostante.

5.2.2. Vernazza

Vernazza è il secondo borgo che si incontra da nord a sud e, nonostante sia uno di quelli interni, anch'esso è solito ricevere un grande afflusso di turisti. L'ufficio del 5 Terre Point di Vernazza è proprio tra i due binari della stazione, il che permette un monitoraggio completo di questa. Siccome l'accesso alla ferrovia è possibile solo tramite una scala che scende sotto al ponte della stazione, anche l'area sottostante è raggiungibile dallo sniffer.



Figura 26 Stazione di Vernazza



Figura 27 Stazione di Vernazza



Figura 28 Vicolo sottostante la stazione

5.2.3. Corniglia

Proseguendo il tragitto verso sud si arriva al terzo paese, Corniglia. Anche in questo caso, come a Vernazza, la stazione ha un'unica via d'uscita dove i turisti si incamminano verso il paese e i sentieri. Il 5 Terre Point si trova proprio accanto a questo vicolo di cui ne fa angolo, costringendo i passanti a transitare all'interno del raggio d'azione dello sniffer.



Figura 29 Stazione di Corniglia



Figura 30 Via d'uscita della stazione di Corniglia

5.2.4. Manarola

Il penultimo paese della lista è Manarola. Come le due frazioni precedenti, per uscire dalla stazione i turisti sono costretti a transitare presso la piazzetta antistante l'edificio adibito a 5 Terre Point. Nel caso di Manarola, un tunnel piuttosto stretto, il cui ingresso è sito di fronte all'edificio in questione, s'interpone tra la stazione e il paese. Motivo in più per monitorare la suddetta area.



Figura 31 Piazzetta di Manarola antistante lo sniffer



Figura 32 5 Terre Point di Manarola

5.2.5. Riomaggiore

L'ultimo borgo, che assieme a Monterosso delimita i confini delle Cinque Terre, è Riomaggiore in cui, come già accennato, il traffico si intensifica. Ci sono diversi punti d'ingresso o uscita dalla stazione, alcuni non proprio vicini al 5 Terre Point, ma questo ha la fortuna di trovarsi nella piazza in cui i turisti transitano per raggiungere il paese, anche questa volta, tramite un tunnel.



Figura 33 5 Terre Point di Riomaggiore - esterno



Figura 34 Accesso al tunnel di Riomaggiore raggiungibile dallo sniffer

5. Applicazione al caso di studio



Figura 35 5 Terre Point di Riomaggiore – interno

5.3. Applicazione Web GIS

Dopo una schermata di login, viene presentata all'utente la seguente schermata:

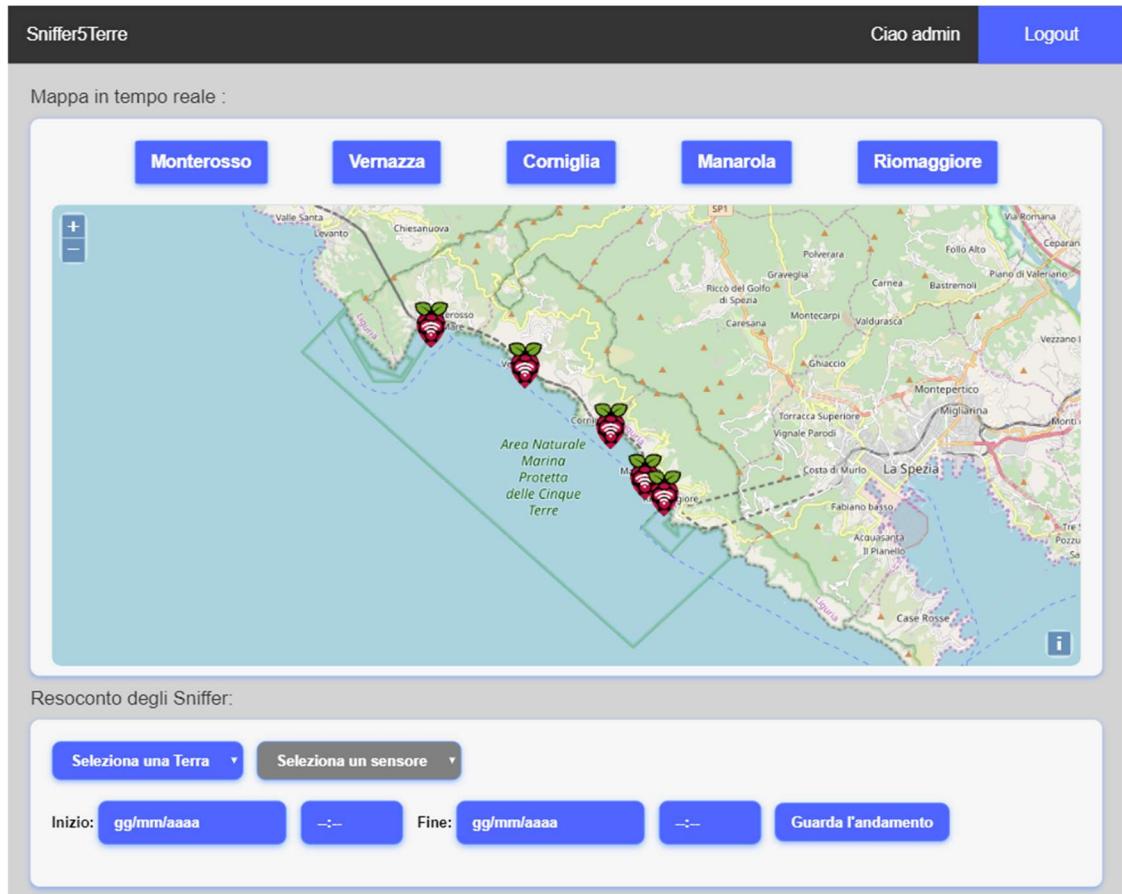


Figura 36 Console sniffer5terre

La mappa mostra una panoramica del Parco delle Cinque Terre in cui sono visibili i marker indicanti le posizioni degli sniffer che sono attivi. Cliccando su uno dei bottoni si viene catapultati all'interno del paese scelto e il marker ingrandito mostra intorno a sé quello che dovrebbe essere il suo raggio d'azione, colorato a seconda della densità turistica percepita in quel momento (verde poco affollato, giallo mediamente affollato, rosso molto affollato). La pagina web controlla periodicamente se ci sono cambiamenti di stato, dove lo “stato” è la condizione tale per cui l’area si riempie di un colore piuttosto che un altro, e se questo cambia, l’utente vede cambiare davanti a sé il colore automaticamente.

Il marker e la sua area sono cliccabili e permettono la visualizzazione di un popup che mostra maggiori informazioni quali il nome associato allo sniffer, il numero di turisti stimato e l'orario di ultimo aggiornamento a cui il dato si riferisce. Inoltre, all'interno della nuova scheda, viene rappresentato sottoforma di grafico lo storico di rilevamenti riferiti all'ultima ora trascorsa rispetto a quella attuale. Nel caso in cui per qualche malfunzionamento non fossero disponibili dati nell'ultima ora, il popup conterrà la dicitura "Nessun dato disponibile nell'ultima ora" e visualizzerà l'ultimo rilevamento disponibile con la relativa marca temporale.

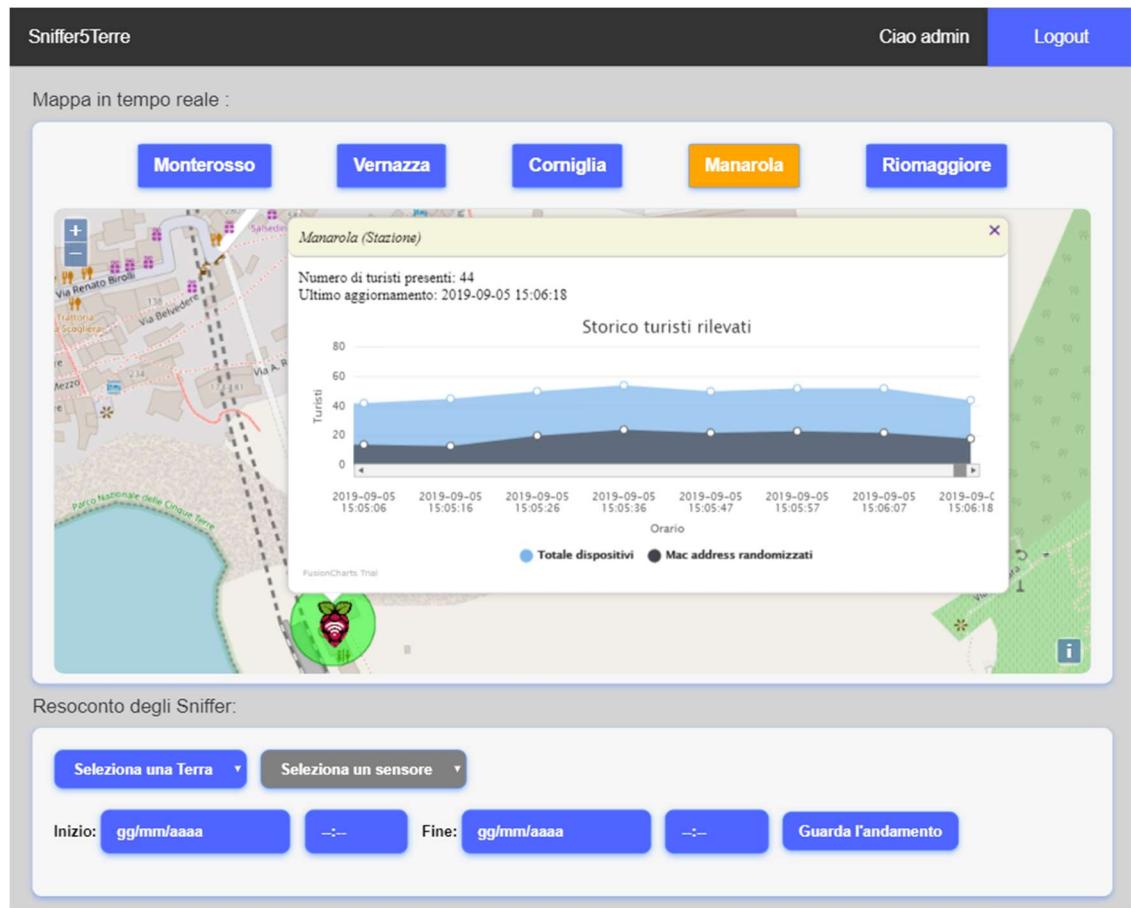


Figura 37 Vista del popup

Sotto alla scheda contenente la mappa, vi è una seconda scheda utile a visualizzare i dati raccolti da uno sniffer a scelta tra quelli disponibili e in un intervallo di tempo personalizzato.

5. Applicazione al caso di studio

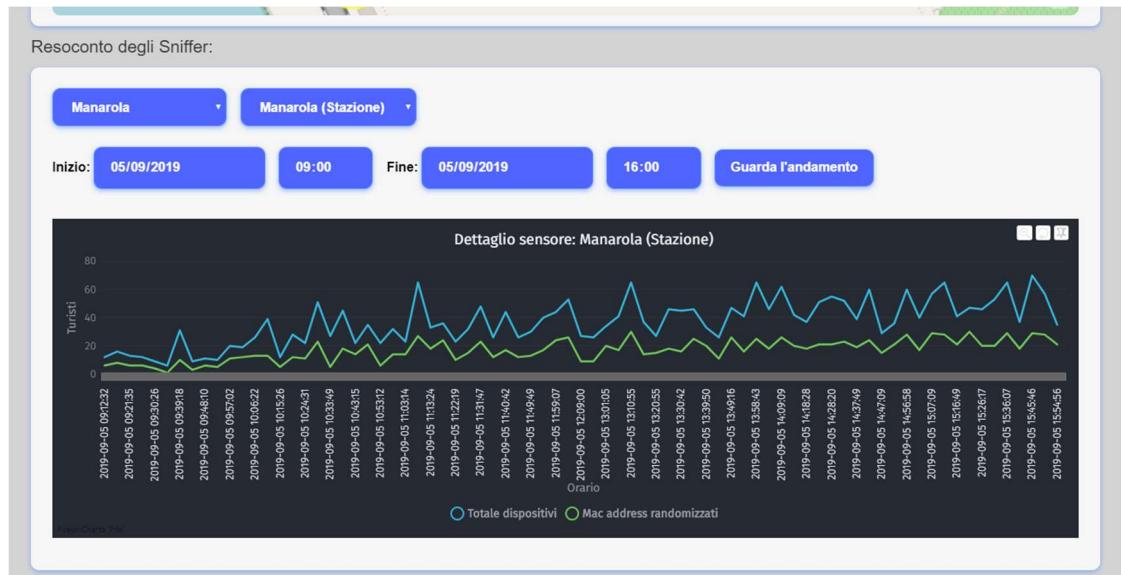


Figura 38 Selezione arbitraria di un intervallo di tempo

Siccome, a differenza del popup, l'intervallo di tempo può essere più esteso di un'ora, questo grafico permette lo zoom di un'area a scelta per ottenere una vista più dettagliata all'interno dell'intervallo selezionato.



Figura 39 Zoom sul grafico



Figura 40 Dettaglio a risoluzione maggiore

5.4. Test di monitoraggio

Stabilire il rapporto tra dispositivi rilevati e persone realmente presenti, implica avere a disposizione il dato esatto riguardo il secondo valore al momento di una misurazione da parte del sensore. Siccome questa informazione è piuttosto difficile da reperire, inizialmente si è pensato di calcolare il tasso di rilevamento presso il parcheggio di Manarola, zona di arrivo dei pullman. Essendo a conoscenza degli orari di arrivo e del numero di turisti trasportati, si pensava che questo tipo di rilevazione fosse verosimile a quella ottenibile presso le stazioni, con il vantaggio di una posizione isolata che consente di non avere disturbi e continuando ad avere un target di popolazione come quello in cui sono piazzati i sensori.

In realtà si è scoperto, una volta sul campo, che la popolazione che preferisce il pullman al treno è significativamente più avanti con l'età e non rispecchia appieno quella rilevata nelle stazioni. Un altro punto a sfavore del parcheggio è che questo non possiede una rete wireless a cui i turisti sono spinti ad accedere. Presso i Cinque Terre Point, dove si ricorda essere posizionati i sensori, gli operatori del Parco annunciano sempre la presenza di Wi-Fi gratuito, compreso con l'acquisto di una 5 Terre Card. Si presuppone quindi che in aggiunta a coloro che già hanno attivato il modulo Wi-Fi, buona parte lo attivi appositamente per usufruire dei servizi appena acquistati.

L'unico modo per ottenere una stima del parametro proporzionale è stato quello di effettuare manualmente rilevazioni, con l'aiuto di contapersone meccanici.

I test eseguiti sino a questo momento mostrano che un parametro proporzionale equivalente a 1,35 può essere un valore iniziale ragionevole.

6. Considerazioni finali

Nonostante la semplicità del modello, questo si è rivelato sufficiente per discriminare il traffico nelle tre classi di viabilità. I sensori sono attivi ventiquattr'ore su ventiquattro e riflettono intuitivamente sull'applicazione web la situazione in loco istante per istante. Gli unici test che sono stati effettuati, dato che il sistema è online da relativamente poco tempo, sono sopralluoghi di persona per stabilire se i dati riportati dai sensori fossero coerenti con la situazione reale.

Sebbene i sistemi operativi si stiano muovendo sul campo della randomizzazione per proteggere la privacy degli utenti, questo meccanismo possiede falte evidenti che permettono la sua neutralizzazione. Innanzitutto, sarebbe opportuno che all'interno delle probe request i dispositivi comunicassero solamente le informazioni essenziali a stabilire una connessione. Aggiungere IE permette una distinzione dei modelli piuttosto efficace. In secondo luogo, il sequence number, informazione su cui si basa fortemente l'algoritmo di aggregazione illustrato, dovrebbe essere azzerato ad ogni cambio di indirizzo. In questo modo, con pochi IE a formare la fingerprint e l'informazione inutilizzabile riguardo il numero di sequenza, bisognerebbe ricorrere a soluzioni molto più elaborate e non alla portata dei più.

7. Sviluppi futuri

Sarebbe opportuno sviluppare un modulo previsionale basato sullo storico delle rilevazioni, registrando eventuali feature come meteo, giorno della settimana, festività e periodo dell'anno. L'applicazione web sarebbe quindi uno strumento a disposizione dei turisti per decidere quando visitare un luogo oltre che un sistema di monitoraggio in tempo reale che affianchi i gestori del parco. Per procedere nella realizzazione di questo strumento sono necessari però mesi di rilevazioni, dati di cui ovviamente in questo momento non si dispone.

Manarola, SP, Italia

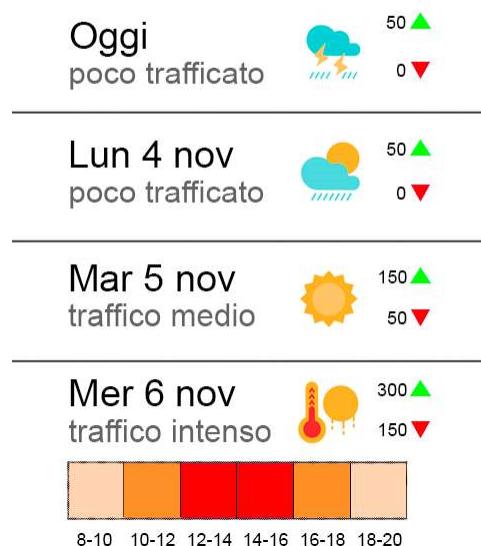


Figura 41 Esempio di applicazione previsionale

In secondo luogo, si potrebbe utilizzare l'informazione derivante dai MAC address globali per ricostruire a scopo informativo i tragitti e i percorsi più battuti. Sarebbe utile equipaggiare ogni paese con più di un sensore in modo da calcolare gli spostamenti e le velocità medie dei turisti all'interno di uno stesso borgo.

Bibliografia

- [1] D. Helbing, «Models for Pedestrian Behavior,» 1998.
- [2] J. J. Fruin, «Designing for pedestrians: a level of service concept,» *Highway Research Board*, pp. 1-15, 1971.
- [3] A. Polus, J. L. Schofer e A. Ushpiz, «Pedestrian Flow and Level of Service,» *Journal of Transportation Engineering*, vol. 109, n. 1, pp. 46-56, 30 12 2008.
- [4] U. Weidmann, «Transporttechnik der Fussgänger Transporttechnische Eigenschaften des Fussgängerverkehrs, Literaturauswertung,» 1992.
- [5] G. Xiong, J. Cheng, X. Wu, Y. L. Chen, Y. Ou e Y. Xu, «An energy model approach to people counting for abnormal crowd behavior detection,» *Neurocomputing*, vol. 83, pp. 121-135, 15 4 2012.
- [6] P. Vera, S. Monjaraz e J. Salas, «Counting pedestrians with a zenithal arrangement of depth cameras,» *Machine Vision and Applications*, vol. 27, n. 2, pp. 303-315, 1 2 2016.
- [7] X. Zhang, J. Yan, S. Feng, Z. Lei, D. Yi e S. Z. Li, «Water filling: Unsupervised people counting via vertical kinect sensor,» in *Proceedings - 2012 IEEE 9th International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2012*, 2012.
- [8] V. A. Sindagi e V. M. Patel, «A survey of recent advances in CNN-based single image crowd counting and density estimation,» *Pattern Recognition Letters*, vol. 107, pp. 3-16, 1 5 2018.
- [9] Q. Chang, Y. Qi, W. Zhou e J. Liu, «An Auto-adaptive CNN for Crowd Counting in Monitor Image,» in *Proceedings of 2018 6th IEEE International Conference on Network Infrastructure and Digital Content, IC-NIDC 2018*, 2018.

- [10] Y. Zhang, D. Zhou, S. Chen, S. Gao e Y. Ma, «Single-image crowd counting via multi-column convolutional neural network,» in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016.
- [11] L. Del Pizzo, P. Foggia, A. Greco, G. Percannella e M. Vento, «Counting people by RGB or depth overhead cameras,» *Pattern Recognition Letters*, vol. 81, pp. 41-50, 10 2016.
- [12] Terabee, «Time-of-Flight principle: Technologies and advantages - Terabee,» [Online]. Available: <https://www.terabee.com/time-of-flight-principle/>.
- [13] L. Hikvision Digital Technology Co, «IP Camera | Network Camera - Hikvision,» [Online]. Available: <https://www.hikvision.com/en/Products/Network-Camera>.
- [14] VIVOTEK Inc., «People Counting :: VIVOTEK ::,» [Online]. Available: <https://www.vivotek.com/solutions/people-counting>.
- [15] DVSLTD, «Hikvision People Counting Camera Set Up - YouTube,» [Online]. Available: <https://www.youtube.com/watch?v=zAK90GxVn5s>.
- [16] H. Wu, C. Gao, Y. Cui e R. Wang, «Multipoint infrared laser-based detection and tracking for people counting,» *Neural Computing and Applications*, vol. 29, n. 5, pp. 1405-1416, 1 3 2018.
- [17] STMicroelectronics, «Time of Flight (ToF) Sensors - STMicroelectronics,» [Online]. Available: <https://www.st.com/en/imaging-and-photonics-solutions/proximity-sensors.html?querycriteria=productId=SC1934#>.
- [18] STMicroelectronics, «People Counting Using a Single ST Time-of-Flight Sensor - YouTube,» [Online]. Available: <https://www.youtube.com/watch?v=c91Ve-g0J2U>.
- [19] BEA, «LZR®-SIGMA | BEA Europe,» [Online]. Available: <https://www.bea-sensors.com/it/prodotto/lzr-sigma/>.
- [20] A. E. Redondi e M. Cesana, «Building up knowledge through passive WiFi probes,» *Computer Communications*, vol. 117, pp. 1-12, 1 2 2018.

- [21] X. Tang, B. Xiao e K. Li, «Indoor Crowd Density Estimation Through Mobile Smartphone Wi-Fi Probes,» *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 25 4 2018.
- [22] Y. Fukuzaki, N. Nishio, M. Mochizuki e K. Murao, «A pedestrian flow analysis system using Wi-Fi packet sensors to a real environment,» 2014.
- [23] J. Weppner, B. Bischke e P. Lukowicz, «Monitoring crowd condition in public spaces by tracking mobile consumer devices with wifi interface,» 2016.
- [24] A. Kurkcu e K. Ozbay, «Estimating Pedestrian Densities, Wait Times, and Flows with Wi-Fi and Bluetooth Sensors,» *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2644, n. 1, pp. 72-82, 7 7 2017.
- [25] A. Basalamah, «Crowd Mobility Analysis using WiFi Sniffers,» *International Journal of Advanced Computer Science and Applications*, vol. 7, n. 12, 4 1 2017.
- [26] Hak5, «WiFi Pineapple - Hak5,» [Online]. Available: <https://shop.hak5.org/products/wifi-pineapple>.
- [27] Raspberry Pi Foundation, «Buy a Pi – Raspberry Pi,» [Online]. Available: <https://www.raspberrypi.org/products/>.
- [28] «Home | Sail Amsterdam,» [Online]. Available: <https://www.sail.nl/?lang=en>.
- [29] W. Daamen, Y. Yuan, D. Duives e S. Hoogendoorn, «Comparing three types of real-time data collection techniques: Counting cameras, Wi-Fi sensors and GPS trackers,» *Proceedings of Pedestrian and Evacuation Dynamics* , pp. 568-574, 2016.
- [30] Y. Yuan, W. Daamen, D. Duives e S. Hoogendoorn, «Comparison of three algorithms for real-Time pedestrian state estimation - Supporting a monitoring dashboard for large-scale events,» in *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2016.
- [31] E. Boni e E. Piccin, «Risultati della sperimentazione dei sistemi di monitoraggio dei flussi pedonali,» 14 Settembre 2018. [Online]. Available:

https://www.venis.it/sites/www.venis.it/files/redazione/Trasparenza/762_AT_180925_all1.pdf.

- [32] C. Matte, M. Cunche, F. Rousseau e M. Vanhoef, «Defeating MAC Address Randomization Through Timing Attacks,» in *Proceedings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks - WiSec '16*, New York, New York, USA, 2016.
- [33] G. Ceccoli, C. Leoni e U. Quartucci, «Relazione progetto Wi-Fi sniffing Cinque Terre,» Luglio 2018. [Online]. Available: https://drive.google.com/file/d/1uJT7iVB9G1NUSr3NJQTa27_o4kqaklH/view.
- [34] J. Martin, T. Mayberry, C. Donahue, L. Foppe, L. Brown, C. Riggins, E. C. Rye e D. Brown, «A Study of MAC Address Randomization in Mobile Devices and When it Fails,» 2017.
- [35] M. Vanhoef, C. Matte, M. Cunche, L. Cardoso, F. Piessens, L. S. Cardoso, †. Iminds-Distrinet e K. U. Leuven, «Why MAC Address Randomization is not Enough: An Analysis of Wi-Fi Network Discovery Mechanisms An Analysis of Wi-Fi Network Discovery Mechanisms,» 2016.
- [36] Wireshark Foundation, «Wireshark · Go Deep.,» [Online]. Available: <https://www.wireshark.org/>.
- [37] Apple Inc., «iOS 12.3 Security,» 2019.
- [38] D. Gentry e A. Pennarun, «Passive Taxonomy of Wifi Clients using MLME Frame Contents».
- [39] Wikipedia.org, «crontab - Wikipedia,» [Online]. Available: <https://it.wikipedia.org/wiki/Crontab>.
- [40] Envato Tuts+, «Scheduling Tasks with Cron Jobs,» [Online]. Available: <https://code.tutsplus.com/tutorials/scheduling-tasks-with-cron-jobs--net-8800>.

- [41] Consorzio Turistico Cinque Terre, «Benvenuto su 5terre | 5terre,» [Online]. Available: <https://www.cinqueterre.it/>.
- [42] remot3.it Inc, «remote.it - Virtual Private Internet,» [Online]. Available: <https://remote.it/>.