

Indoor tracking e analisi dati Wi-Fi in ambito GDO

Alban Gorreja^{1[0000-0003-0016-418X]}, Eva Savina Malinverni e^{2[0000-0001-6582-2943]} Roberto Pierdicca^{3[0000-0002-9160-834X]}

¹ Università Politecnica delle Marche, a.gorreja@pm.univpm.it

² Università Politecnica delle Marche, e.s.malinverni@staff.univpm.it

³ Università Politecnica delle Marche, r.pierdicca@staff.univpm.it

Parole chiave. Posizionamento e Location intelligence, indoor, retail, analisi dati, WiFi.

Abstract:

L'analisi del comportamento umano è al centro di moltissime ricerche e studi, di molteplici ambiti tra cui quelli di carattere scientifico ed economico-commerciale. Questo lavoro di ricerca è focalizzato nell'individuare il posizionamento dinamico in ambienti indoor delle persone, basandosi sul segnale WiFi, in quanto l'uso del solo segnale GPS, negli odierni smartphone ad oggi non è sufficiente per la geolocalizzazione. Successivamente, una breve analisi dei dati verrà fatta per dimostrare il corretto funzionamento dell'impianto, per sottolineare l'importanza che la conoscenza dei dati e l'uso che se ne può fare. In letteratura si trovano diversi modi per localizzare gli utenti in uno spazio indoor, ma quasi sempre bisogna dotare la persona interessata di uno strumento wireless (come ad esempio gli RFID) oppure far interagire la persona con una specifica applicazione in grado di comunicare con la strumentazione installata nel luogo di interesse attraverso Bluetooth.

Tuttavia, la localizzazione delle persone tramite WiFi può avvenire sia senza il download di applicazioni per lo smartphone sia senza collegamento alla rete WiFi. Questo permette la geolocalizzazione delle persone pur non raccogliendo informazioni su di essi, garantendo quindi l'anonimato totale e rispettando le norme sulla privacy vigenti.

1 Introduzione

La conoscenza e la gestione delle informazioni e dei dati sulle persone riveste primaria importanza ormai da decenni, soprattutto in ambiti dove il marketing gioca un ruolo fondamentale [1]. In un mondo in continuo sviluppo e sempre più digitalizzato diventa necessario raccogliere e gestire i dati in modo efficiente. Le grandi aziende del web profilano i loro utenti per poter ottenere il massimo in termini di ricavi, vendendo i dati o fornendo annunci mirati sulla base della profilazione fatta. In questo modo gli investimenti sono incentrati su una precisa fetta di mercato, composta da quelle persone veramente interessate a uno specifico prodotto o servizio [2]. Così come nel web lo stesso concetto si applica nella vita fuori dal mondo di internet. Cartelloni pubblicitari, schermi e comunicazioni di ogni tipo ci travolgono durante tutto il corso della giornata. Quindi è fondamentale per chi vende, che queste informazioni arrivino solo alle persone

potenzialmente interessate, senza dover perdere in inutili investimenti. Nel mondo fisico senza l'ausilio degli strumenti informatici di profilazione, è difficile accedere alle preferenze delle persone e quindi fidelizzare. In questo articolo tratteremo di come si possono ottenere informazioni su un generalizzato comportamento umano all'interno di un centro commerciale. L'obiettivo è quello di descrivere un sistema in grado di studiare il comportamento in funzione delle preferenze all'interno di un luogo chiuso. Questo sistema non solo permette di conoscere quali sono i poli attrattori più forti, ma anche di capire come questi possono cambiare nel tempo e cosa fare per orientare l'attenzione dei consumatori verso un preciso percorso o verso certi prodotti che vendono meno di altri.

Vi sono diverse tecnologie per localizzare le persone. Tra le più comuni troviamo il GPS dei nostri smartphone, ma questo trova grandi limiti negli spazi indoor e soprattutto diventa necessario richiedere l'accesso al dispositivo tramite un'applicazione apposita. Si perde in questo modo la possibilità di mappare tutti gli spostamenti di coloro che frequentano lo spazio. In letteratura si possono osservare diversi esempi per giungere alla geolocalizzazione all'interno degli edifici, spesso anche sovrapponendo diverse tecniche e tecnologie [3] [4]. Ognuno di questi sistemi offre delle buone prestazioni, ma si basano soprattutto nella localizzazione puntuale o statica delle persone che frequentano un ambiente indoor, senza definirne la traiettoria. Per questo tipo di soluzione invece, bisogna affidarsi a strumenti che usano tecnologia Bluetooth [5]. Come mostrato in [6], l'utilizzo di questi dispositivi richiede, anche in questo caso, che l'utente scarichi e utilizzi un App appositamente creata per far trovare e riconoscere la propria posizione. Un'altra tecnologia che può essere utilizzata per questo scopo è il LiFi, in grado di definire con alta precisione la posizione dei dispositivi, si parla di pochi centimetri di errore ma questa tecnologia si presenta inadatta a coprire uno spazio esteso, inoltre non risulta efficace perché come già visto in precedenza, anche in questo caso bisognerebbe munire le persone di un dispositivo in grado di interagire con il LiFi. Oltretutto i costi eccessivamente elevati non giustificano quasi mai questa tecnologia se non per installazioni spot con tutti i limiti che ne conseguono. [7]

Al contrario, grazie al WiFi il riconoscimento dei dispositivi e quindi delle persone può avvenire senza che ci sia bisogno di scaricare applicazioni per lo smartphone e neppure effettuando l'accesso alla rete WiFi, mantenendo l'anonimato totale quindi rispettando le norme sulla privacy vigenti. Questa è la caratteristica principale, nonché il punto di forza di questo sistema basato sull'individuazione del dispositivo cellulare da parte della rete WiFi. È sufficiente che i dispositivi abbiano attivo il loro WiFi, per essere rilevati e localizzati all'interno dell'ambiente indoor (Fig.1). anche senza essere collegati alla rete interna.



Fig. 1 Mappa di calore dell'area presa in esame

2 Metodologia

Nello specifico, i test sono stati condotti in un centro commerciale di Milano, analizzando i dati per tutto l'anno 2020. Sono stati utilizzati otto Access Point (AP), tipo Cambium cn pilot e600, che hanno permesso la copertura di un'area di 1600 mq, rappresentativa, dell'intera struttura. Si fa quindi notare che non tutta la struttura è stata coperta dal sistema di localizzazione. Questo avrà chiaramente effetti sui dati raccolti, nello specifico nelle zone non coperte dal sistema di tracking. L'area interessata è stata suddivisa in 30 celle, le quali definiscono la sensibilità dell'output, ovvero la localizzazione mappata dei visitatori del centro commerciale. Le persone rilevate verranno conteggiate all'interno di una di queste celle che possiamo immaginare come punti all'interno della mappa. Poiché il segnale WiFi si estende oltre l'area di 1600mq interessata all'interno del centro commerciale, tutte le persone rilevate al di fuori di quest'area di studio verranno inserite in queste celle di "estremità" andando a riempirle di persone che in realtà non sono presenti. Questo problema si risolverebbe coprendo tutto il centro commerciale con il sistema di tracking. Poiché questo non è stato possibile in questa fase di studio, basta non tener conto di queste celle specifiche.

Il sistema strutturato con questi specifici Access Point permette di rilevare tutti i dispositivi i quali WiFi è attivo, non necessariamente collegati alla rete dell'ambiente interno, ma anche quelli allacciati alla rete tramite i propri dati. Anche il dispositivo in stand-by, non in utilizzo viene rilevato. Poiché il sistema WiFi deve poter riconoscere i dispositivi, e tramite essi i visitatori del centro mentre si muovono, l'algoritmo deve poter leggere un nome univoco da ogni dispositivo. Questo nome univoco è il MAC

address di cui ogni smartphone è dotato. Questo dato è univoco e non è legato ai dati privati di una persona. Oltretutto i moderni smartphone per garantire un maggior livello di sicurezza relativa alla privacy, possono cambiare il proprio MAC address, il quale viene letto dal sistema WiFi. Si parla in questo caso di MAC address randomizzato. Per poter garantire la traiettoria percorsa dal visitatore, il MAC address deve essere sempre lo stesso altrimenti l'algoritmo non può riconoscere il nuovo MAC address se dovesse cambiare all'improvviso. Tuttavia, il MAC address randomizzato non rappresenta un problema serio, perché questo si aggiorna ogni volta che il dispositivo viene riavviato o che lo stesso WiFi dello smartphone viene spento e riaccessso. Casistica molto improbabile durante la visita in un centro commerciale, poiché non solo gli utenti possono non utilizzare mai il telefono ma anche quando lo fanno, difficilmente hanno motivo di riavviare il proprio WiFi, soprattutto chi naviga attraverso i dati del proprio operatore.

L'algoritmo di localizzazione, attraverso la triangolazione dei punti, riesce a stabilire, successivamente ad un'attenta calibrazione in sito, dove si trova il dispositivo rilevato, quindi registra la presenza, quante volte e dove, presso una delle 30 celle. Più tempo una persona passa fermo in una cella e più volte viene letto il suo valore all'interno di una cella, definendo quindi la permanenza dell'utente. È possibile ottenere una maggiore sensibilità aumentando il numero di celle, ma per gli scopi da noi prefissati (marketing e gestione emergenze) si ritiene opportuno una quantità di 30 celle. Tutto questo ha permesso di studiare non solo l'andamento durante l'anno dei visitatori del centro commerciale, ma anche di fare analisi per aree, fornendo dati per stilare valutazioni sulla frequentazione delle zone e identificare i maggiori poli attrattori ad ogni orario della giornata, così da poter fare analisi di mercato oltre che valutazioni sulla sicurezza in caso di emergenza.

Per poter utilizzare il servizio di WiFi gratuito per chi si trova nel centro commerciale, bisogna effettuare una registrazione nella quale viene chiesta attraverso un'informativa, il consenso di essere tracciati. A questo punto avviene una distinzione tra i "visitatori" ovvero tutti coloro che entrano nel centro commerciale, ma non usufruiscono della rete interna e quindi non acconsentono a essere profilati, e gli "utenti" ovvero le persone che accettano di essere tracciate, le quali possono inoltre fornire dati personali, quali nome, età, genere. Vedremo successivamente come anche una raccolta di poche informazioni sugli utenti, permetta una prima profilazione molto utile per azioni di marketing.

3 Risultati

In questo anno segnato dagli effetti della pandemia, questo studio ha sottolineato l'importanza di gestire l'informazione dando prova anche del corretto funzionamento del sistema di tracciamento. I dati sono stati raccolti in una piattaforma di business intelligence, denominata Tableau, la quale ha permesso di elaborare i dati dall'inizio della pandemia rendendoli leggibili da tutti. Un esempio è mostrato in Fig. 2 dove è possibile osservare nel grafico l'andamento del numero di visitatori nell'arco di sette

mesi, a partire da gennaio 2020 registrati dalla rete WiFi a servizio del centro commerciale preso in considerazione per lo studio. Il periodo di rilevamento dei dati è stato interessato dagli effetti delle misure di contenimento della pandemia Covid-19, che hanno causato la chiusura di tante attività produttive. Considerando ciò è subito chiaro come si sia passati dai picchi di quasi novemila persone rilevate nei weekend, delle prime settimane del 2020 alle poche centinaia registrate nei giorni di lockdown, dove comunque sono rimasti aperti i supermercati per il rifornimento alimentare.

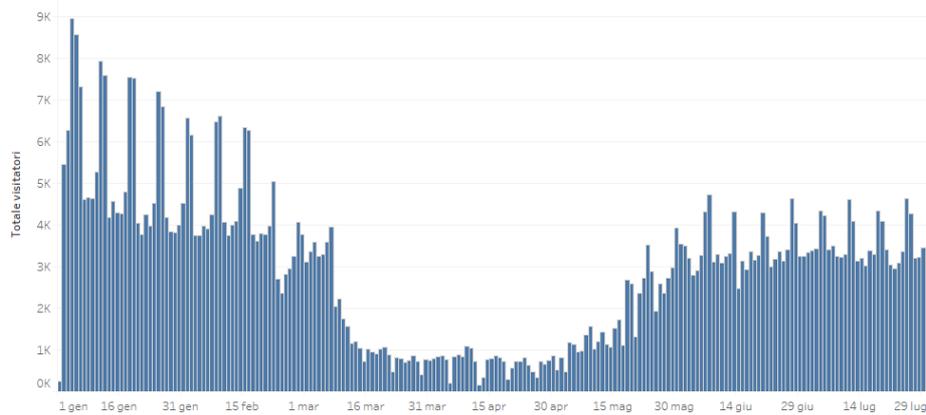


Fig. 2 Andamento dei visitatori nel periodo gennaio - luglio 2020

I dati raccolti sono stati raggruppati secondo diversi indicatori chiave di prestazioni (KPI), scelti in base alle possibili esigenze dei destinatari. Come spiegato in precedenza, questo sistema permette il tracciamento delle traiettorie delle persone presenti all'interno di uno spazio indoor. Generalmente, i dati analizzati sono stati raggruppati in dettagli dei singoli utenti, statistiche giornaliere, statistiche mensili e analisi di comportamento globale. Con i dati dei visitatori non profilati, possiamo costruire una mappa di calore (Heatmap) che esprime la distribuzione nello spazio delle persone così da evidenziare le celle con maggior affluenza. Da qui si possono fare ulteriori valutazioni sugli orari o giorni della settimana con più affluenza così come i giorni del mese. Invece, per quanto riguarda gli utenti profilati, si possono derivare informazioni relative alle aree dove l'utente ha speso maggior tempo, la frequenza delle visite nell'arco del mese, quanto tempo passa nel centro commerciale e in quali orari della giornata. In figura 3 vediamo un esempio delle aree più frequentate da un utente registrato, il quale ha dato il consenso al trattamento dei dati personali. In tabella 1 e 2 vediamo inoltre le informazioni relative allo stesso utente nei quindici giorni precedenti al 20 giugno 2020.



Fig. 3 Celle più frequentate da un singolo utente in un giorno

Tabella 1. Statistiche dal 06/06/2020 al 20/06/2020.

Sesso	Frequenze visite	Frequenza max	Permanenza media	Giorno preferito
Donna	7	01:55:56	01:13:48	Giovedì

Tabella 2. Dati heatmap relative al 20/06/2020

Giorno	Ora primo ingresso	Ora ultima uscita	Permanenza
2020-06-20	07:59:39	12:23:34	01:14:39
2020-06-18	12:22:38	19:22:51	01:55:56
2020-06-17	08:21:16	13:09:08	00:43:32
2020-06-16	08:04:02	13:01:08	00:58:17
2020-06-13	13:14:58	20:08:26	01:43:22
2020-06-12	12:14:52	18:58:16	01:28:07
2020-06-11	12:24:12	19:01:34	01:37:05

4 Conclusioni

Analisi di questo genere permetterebbero ai centri commerciali di sfruttare al massimo le loro potenzialità di vendita, di capire cosa attira maggiormente i propri clienti, di riorganizzare gli spazi pubblicitari e di gestire meglio le aree sovraffollate. La possibilità di fare osservazioni e avere un quadro degli spostamenti delle persone in

ambienti interni può essere un valore aggiunto per la gestione dell'odierna situazione di pandemia nonché fondamentale per la gestione di un gran afflusso di persone in situazioni di emergenza.

Questo lavoro di ricerca è in progresso e punta a migliorare sia l'algoritmo sia la strumentazione in sito, ampliando e unendo le tecnologie citate in precedenza, con altre quali dispositivi Beacon, Raspberry e telecamere contapersone per ottenere un risultato complessivo più accurato. La precisione ad oggi non è assoluta e le variabili da considerare sono molte. Un importante passo avanti su cui questo lavoro si sta concentrando è dato dall'integrazione di questi dati in modelli informativi 3D come il BIM, dove i dati relativi al tracciamento verranno integrati in modelli in grado di sfruttare le potenzialità intrinseche di un sistema informativo. Tra queste si evidenziano la possibilità di far interagire il modello stesso con diversi stakeholders e/o fare analisi sulle aree più frequentate dello spazio preso in esame. Inoltre, stiamo lavorando per realizzare un modello virtuale dello spazio con strumentazione Matterport. Questa strumentazione ci permetterebbe di ottenere uno spazio virtuale 3D navigabile nel quale vorremo inserire i dati precedentemente ricavati.

Riferimenti bibliografici

- [1] N. A. Morgan, R. J. Slotegraaf, and D. W. Vorhies, "Linking marketing capabilities with profit growth," *Int. J. Res. Mark.*, vol. 26, no. 4, pp. 284–293, Dec. 2009, doi: 10.1016/j.ijresmar.2009.06.005.
- [2] P. J. Wisniewski, B. P. Knijnenburg, and H. R. Lipford, "Making privacy personal: Profiling social network users to inform privacy education and nudging," *Int. J. Hum. Comput. Stud.*, vol. 98, pp. 95–108, Feb. 2017, doi: 10.1016/j.ijhcs.2016.09.006.
- [3] A. Yassin *et al.*, "Recent Advances in Indoor Localization: A Survey on Theoretical Approaches and Applications," *IEEE Commun. Surv. Tutorials*, vol. 19, no. 2, pp. 1327–1346, 2017, doi: 10.1109/COMST.2016.2632427.
- [4] F. Zafari, A. Gkelias, and K. K. Leung, "A Survey of Indoor Localization Systems and Technologies," *IEEE Commun. Surv. Tutorials*, vol. 21, no. 3, 2019, doi: 10.1109/COMST.2019.2911558.
- [5] S. S. Chawathe, "Beacon placement for indoor localization using Bluetooth," 2008, doi: 10.1109/ITSC.2008.4732690.
- [6] E. Osaba, R. Pierdicca, E. S. Malinverni, A. Khromova, F. J. Lvarez, and A. Bahillo, "A smartphone-based system for outdoor data gathering using a wireless beacon network and GPS data: From cyber spaces to senseable spaces," *ISPRS Int. J. Geo-Information*, vol. 7, no. 5, 2018, doi: 10.3390/ijgi7050190.
- [7] M. A. Arfaoui *et al.*, "Invoking Deep Learning for Joint Estimation of Indoor LiFi User Position and Orientation," *IEEE J. Sel. Areas Commun.*, 2021, doi: 10.1109/JSAC.2021.3064637.

#AsitaAcademy2021