



**Consiglio Nazionale delle Ricerche
Istituto di Calcolo e Reti ad Alte Prestazioni**

*Posizionamento indoor con tecnologia Wi-Fi
Ottimizzazione del metodo "fingerprinting"
in ambienti chiusi e confinati o strutturati.*

C. Lupascu, A. Machì, L. Minneci, G. Tarantino, M. Tripiciano
ICAR-CNR UOS Palermo

RT-ICAR-PA-14-02

Giugno 2014



Consiglio Nazionale delle Ricerche, Istituto di Calcolo e Reti ad Alte Prestazioni (ICAR)
– Sede di Cosenza, Via P. Bucci 41C, 87036 Rende, Italy, URL: www.icar.cnr.it
– Sezione di Napoli, Via P. Castellino 111, 80131 Napoli, URL: www.na.icar.cnr.it
– Sezione di Palermo, Viale delle Scienze, 90128 Palermo, URL: www.pa.icar.cnr.it



Consiglio Nazionale delle Ricerche
Istituto di Calcolo e Reti ad Alte Prestazioni

Progetto FESR Sicilia GEOPOS
Deliverable Attività di Ricerca
Maggio 2014 - Giugno 2014

Team di progetto

Dott. A. Machì	ICAR-CNR	Responsabile Scientifico
Dott. M. Tripiciano	ICAR-CNR	Ricercatore
Dott. C. Lupascu	ICAR-CNR	Assegnista Ricerca
Ing. L. Minneci	ICAR-CNR	Assegnista Ricerca
Ing. G. Tarantino	ICAR-CNR	Assegnista Ricerca

* si ringraziano il Prof. C. Giaconia del DEIM UniPa e l'Ing. A. Scianna dell'IBAM/CNR per le stimolanti discussioni tecniche e gli utili suggerimenti nella esecuzione del progetto.

RT-ICAR-PA-14-02

Data Ultimo Aggiornamento: 07 Luglio 2014



Consiglio Nazionale delle Ricerche, Istituto di Calcolo e Reti ad Alte Prestazioni (ICAR)
– Sede di Cosenza, Via P. Bucci 41C, 87036 Rende, Italy, URL: www.icar.cnr.it
– Sezione di Napoli, Via P. Castellino 111, 80131 Napoli, URL: www.na.icar.cnr.it
– Sezione di Palermo, Viale delle Scienze, 90128 Palermo, URL: www.pa.icar.cnr.it

I rapporti tecnici dell'ICAR-CNR sono pubblicati dall'Istituto di Calcolo e Reti ad Alte Prestazioni del Consiglio Nazionale delle Ricerche. Tali rapporti, approntati sotto l'esclusiva responsabilità scientifica degli autori, descrivono attività di ricerca del personale e dei collaboratori dell'ICAR, in alcuni casi in un formato preliminare prima della pubblicazione definitiva in altra sede.

1	Introduzione	5
2	Localizzazione tramite fingerprinting	5
3	Campagna di misure ed esperimenti per il controllo della variabilità ed il miglioramento dell'accuratezza.....	8
3.1	Variabili influenti sull'accuratezza di localizzazione	8
3.2	Campagne di misure e dataset.....	8
3.2.1	Sito PoloP1.....	8
3.2.2	Sito AreaP2	9
3.2.3	Sito AreaPT.....	11
3.3	Valutazioni empiriche ed esperimenti.....	12
3.3.1	Test A (non effettuato): incertezza sulla misura d'intensità del segnale acquisito al variare del modello del terminale smartphone.....	12
3.3.2	Test B: fluttuazione temporale istantanea del segnale radio e della sua distribuzione spaziale.....	13
3.3.3	Test C: effetto della fluttuazione interna delle componenti del pattern di fingerprint... ..	18
3.3.4	Test E: Influenza della densità di campionamento dei punti di riferimento	20
3.3.5	Test F: scelta dell'algoritmo di localizzazione	22
3.3.6	Test G: incertezza sulla stima della posizione dell'operatore durante l'acquisizione della radiomappa.	37
3.3.7	Controllo network-side della variabilità lenta del campo a radiofrequenza	39
3.4	Setting:	43
3.5	Esperimenti effettuati:.....	43
3.6	Valutazione dei risultati degli esperimenti e problemi ancora aperti.....	57
4	Conclusioni	59
4.1	Tecniche di ottimizzazione e regole di best practice	59
4.2	Varianti alle regole CISCO per la pianificazione della posizione AP	60
4.3	Ulteriori test da sviluppare.....	60
5	Appendice A: struttura della Base di Dati ed esempi di query in PostgreSQL	61
5.1	Query di esempio	62
6	Appendice B: mappe di distribuzione spaziale del campo degli AP nel sito Area2P (porte aperte).....	69
7	Appendice C: sorgente documentato del codice dei moduli mobile logger, radiomap LAN server e mobile localizer	73
7.1	Funzione che implementa l'algoritmo di localizzazione KNN	73
7.2	Funzione che implementa l'algoritmo di localizzazione KMEANS	74
8	Appendice D: sorgente documentato degli script di simulazione di algoritmi su piattaforma MATLAB.....	80
8.1	Modulo estrazione dati dal RDBMS	80
8.2	Modulo statistiche errore con algoritmo KNN	82
8.3	Modulo statistiche errore con algoritmo K-Means	84
8.4	Modulo statistiche errore con algoritmo di apprendimento del rank	86
9	Appendice E: sorgente documentato del codice dei moduli della applicazione di monitoraggio network-side.....	89
10	Bibliografia	96

1 Introduzione

Il presente Rapporto Tecnico descrive l'analisi della variabilità del segnale e della accuratezza della metodologia di localizzazione client-side tramite fingerprinting descritte nel Deliverable GEOPOS_1 [1].

Sono anche descritti i dettagli realizzativi delle tecniche di ottimizzazione ed analisi messe in opera sia nel preprocessing delle misure di ampiezza del segnale WiFi utilizzate per la creazione di radiomappe, che sugli algoritmi di localizzazione che fanno uso delle stesse. Vengono inoltre riportati i risultati ottenuti con l'applicazione delle metodologie di ottimizzazione implementate, confrontandone il livello di accuratezza e precisione (varianza) con riferimento al fingerprinting standard.

2 Localizzazione tramite fingerprinting

Dati un ambiente chiuso più o meno strutturato ed un insieme di n dispositivi di accesso alla rete (AP) emettenti periodicamente pacchetti di segnalazione di persistenza della rete (beacon) in accordo al protocollo 802-11xxx WiFi, ed opportunamente disposti in modo da coprire con il proprio campo elettromagnetico l'intero ambiente, viene indicato come "fingerprint" di un punto del campo, la nupla costituita dai valori medi ordinati delle intensità del segnale di ogni AP (componente), come rilevata da un dispositivo mobile calibrato e posizionato nel punto [1 pp. Xx].

Ciascuna di tali componenti del fingerprinting è di per se una variabile aleatoria soggetta a fluttuazioni statistiche, anche significative, dovute ad un insieme di fattori quali inaccuratezza della determinazione della posizione di misura, variazione delle caratteristiche di assorbimento del mezzo interposto, presenza di segnali riflessi, complesse funzioni direzionali del guadagno di antenna, interferenze fra segnali di emettitori diversi sullo stesso canale di trasmissione, etc [4].

La distribuzione di intensità del campo risultante dalla interazione del segnale radio emesso dall'AP con l'ambiente confinante risulta complessa, generalmente decrescente con la distanza, ma non radialmente monotona, con massimi e minimi locali e difficilmente rappresentabile con modelli di path-loss basati su funzioni decrescenti secondo leggi di potenza, anche a distanza dalla zona di prossimità all'antenna [16], specialmente in caso di ambienti strutturati.

Il metodo del fingerprinting consiste nel campionare opportunamente i valori del campo di ogni AP nella zona di interesse, in un numero significativo di punti di riferimento e costruire una radio mappa dell'ambiente costituita dalla matrice a $n+2$ dimensioni contenente i valori medi dei valori di segnale associati alle coordinate dei punti di riferimento [17].

Il metodo presuppone che, se la geometria dell'ambiente non subisce trasformazioni, se i valori medi dell'intensità rappresentano ragionevolmente il fingerprint di un punto di riferimento e se i punti di riferimento campiono opportunamente o con frequenza spaziale elevata l'ambiente, la radio mappa costituisca una buona approssimazione spaziale dei campi nell'area campionata.

Dato il vettore delle componenti istantanee rilevato in un qualunque punto di test, esso dovrebbe essere più simile in media ai fingerprint dei punti di riferimento spazialmente vicini che a quelli spazialmente lontani. La presenza di un numero elevato di AP, e quindi di componenti della misura, dovrebbe irrobustire la somiglianza media limitando l'influenza di irregolarità nella distribuzione del campo del singolo AP.

Gli algoritmi di localizzazione tramite radiomappa tendono quindi a selezionare un numero minimale di punti di riferimento sulla radio mappa simili nel fingerprinting e stimare la posizione del punto di test attraverso interpolazione o media pesata delle coordinate dei punti individuati.

Col metodo k-nearest-neighbor (kNN) vengono cercati i k fingerprint di riferimento più vicini (dal punto di vista della distanza euclidea nello spazio delle componenti di RSSI) ai fingerprint di test. Successivamente vengono mediate le k posizioni più vicine, ottenendo una stima della posizione del punto di test.

Col metodo clustering con k-means, i fingerprint di riferimento vengono divisi in due cluster con l'algoritmo kmeans. L'obiettivo dell'algoritmo kmeans è di minimizzare la varianza totale intra-cluster. Ogni cluster viene identificato mediante un centroide. L'algoritmo segue una procedura iterativa:

- inizialmente crea K partizioni e assegna ad ogni partizione i punti di riferimento casualmente
- calcola il centroide di ogni partizione
- costruisce una nuova partizione associando ogni punto di riferimento al cluster il cui centroide è più vicino ad esso
- vengono ricalcolati i centroidi per i nuovi cluster e così via, finché l'algoritmo non converge (i centroidi non cambiano più).

Per ogni punto di test si cerca il fingerprint di riferimento più vicino (dal punto di vista della distanza euclidea nello spazio delle componenti di RSSI). I “vicini “ del punto di test saranno i punti di riferimento appartenenti allo stesso cluster del punto di riferimento trovato più vicino al punto di test. La procedura continua iterativamente con la riclusterizzazione dei “vicini” del punto di test in due cluster. I nuovi “vicini “ del punto di test saranno i punti di riferimento appartenenti allo stesso cluster del punto di riferimento trovato più vicino al punto di test. La procedura converge quando il numero di “vicini” è minore di k (k=10).

Nelle versioni pesate (Weighted KNN e Weighted KMeans) invece del baricentro si calcola il baricentro pesato in base all'inverso delle distanze euclidee nello spazio degli RSS.

Nel Deliverable GEOPOS_1 [1] sono riportati un insieme di esperimenti preliminari volti ad acquisire evidenza degli effetti sopra descritti sul campo di radiazione degli AP ed a valutarne l'importanza. E' descritta una campagna di mappatura di un largo ambiente indoor parzialmente strutturato presso il Polo Didattico dell'Università di Palermo (PoloP1), adattando regole di buona pratica suggerite in letteratura ed in particolare, quelle suggerite dalla società CISCO [19]. Sono, infine, descritte le modifiche apportate al package open-source AirPlace [18] utilizzato per la acquisizione dei segnali RSSI su terminale mobile, il calcolo delle radio mappe su server LAN e la localizzazione su terminale mobile.

La tabella sotto riportata indica i valori di accuratezza media di localizzazione di 54 punti di test, con le relative deviazioni standard medie, ottenute utilizzando 24 AP per coprire una superficie di 172x54 mq circa, e campionando quasi uniformemente con 82 punti di riferimento la parte centrale di 150 x 20 metri circa, con punti di riferimento su una griglia quadrata di 4 metri di lato.

Algoritmo	parametro	Accuratezza media	Deviazione
-----------	-----------	-------------------	------------

		(m)	standard (m)
kNN	k=4	6.739	3.475
Weighted kNN	k=5	6.678	3.506
k-Means	cardinalità cluster finale=4	7.406	4.069
Weighted k-Means	cardinalità cluster finale=4	7.244	4.087

Tabella 2-1 - Variazione accuratezza al variare dell'algoritmo di clustering in assenza di pubblico

Algoritmo	parametro	Accuratezza media (m)	Deviazione standard (m)
kNN	k=12	7.067	4.377
Weighted kNN	k=12	6.950	4.170
k-Means	cardinalità cluster finale=4	8.365	5.204
Weighted k-Means	cardinalità cluster finale=4	8.504	5.251

Tabella 2-2 - variazione accuratezza al variare dell'algoritmo di clustering in presenza di folto pubblico

I valori di accuratezza ottenuti (errore medio fra 7 ed 8 m \pm 4 m) sono compatibili con quelli descritti in letteratura per sistemi robusti commerciali in condizioni operative reali [17] ma risultano meno performanti rispetto ad alcuni sistemi descritti in letteratura in ambienti di dimensioni e struttura simili [20]. I valori ottenuti sono nettamente superiori alla distanza media dei punti di riferimento (4 metri), e sono affette da alta variabilità.

Nel resto del rapporto vengono descritto lo studio analitico delle componenti di tale variabilità effettuato attraverso una nuova campagna di misura ed esperimenti mirati (AreaP2 ed AreaPT). Vengono anche descritte le correzioni agli algoritmi di analisi apportate per ottimizzare il metodo.

In particolare, nel Capitolo 3 si riesaminano le variabili che influiscono sull'accuratezza e si analizzano i risultati di una campagna di acquisizione e di esperimenti mirati di verifica. Nel Capitolo 4 si discutono un insieme di regole di buona pratica nella pianificazione del posizionamento degli AP desunte dalla esperienza empirico/analitica fin qui accumulata. Nel Capitolo 5 si comparano varianti dell'algoritmo di clustering dei punti di riferimento. Infine, nel Capitolo 6 si discute una metodologia di controllo della variabilità temporale lenta della radio-mappa attraverso misure incrociate di intensità RSSI effettuate dagli stessi AP in modalità di acquisizione network side.

L'Appendice A descrive la struttura della Base di Dati ed esempi di query in Posgres SQL utilizzate per la conservazione ed filtraggio dei dati, l' Appendice B mostra in formato grafico esempi di distribuzione spaziale dei campi degli AP della campagna di misure AreaP1, le Appendici C,D e E riportano il codice sorgente dei moduli software della applicazione mobile di logging, del server di radio mappa e della app mobile di localizzazione. Infine l'Appendice F riporta dettagli delle misure dello studio di cui al Capitolo 6.

3 Campagna di misure ed esperimenti per il controllo della variabilità ed il miglioramento dell'accuratezza

3.1 Variabili influenti sull'accuratezza di localizzazione

Nel corso della campagna di misure al Polo Didattico, descritta nel Deliverable GEOPOS_1 [1], è stato possibile individuare una serie di fattori che concorrono direttamente e/o indirettamente alla variabilità dei valori di campo rilevati e alla accuratezza della stima della posizione.

Fra queste le principali sono:

- la tipologia di terminale utilizzato per l'acquisizione,
- la distribuzione spaziale della fluttuazione temporale istantanea del segnale radio,
- la composizione del fingerprint,
- l'errore nella stima della posizione dovuta all'algoritmo di localizzazione,
- la densità di campionamento della radiomappa
- l'accuratezza con cui operatore segue il protocollo di misura mentre acquisisce i valori di campo nei punti di calibrazione e/o di test,
- le variazioni nella geometria dell'ambiente e/o l'assorbimento dovuto alla presenza di pubblico..

Nel seguito vengono riportate esperienze ed esperimenti effettuati in due nuovi siti, per valutare i singoli contributi all'errore. Laddove l'esperimento non è stato effettuato, viene fornita una stima empirica dell'ordine di grandezza del contributo di errore atteso, in base alla esperienza empirica acquisita degli sperimentatori.

3.2 Campagne di misure e dataset

Sito	Tipologia	Area	Dist muri	AP	Dens AP
1. PoloP1	Semi-indoor confinato	160x30	8-12	24	1/200 mq
2. AreaP2	Indoor strutturato	26x26	3-5	8	1/85 mq
3. AreaPT	Semi-indoor confinato	25x4	5-12	1	1/100

3.2.1 Sito PoloP1

Il sito PoloP1, consiste nell'intero primo piano del complesso multi-didattico dell'Università di Palermo. (v. foto vuoto/affollato) descritto nel Deliverable GEOPOS_1 [1].

E' costituito da 3 gruppi di 4 aule principali ciascuno, prospicienti un ampio corridoio centrale, intervallate da ballatoi con ampie vetrate verso l'esterno e da scale di accesso. Le pareti delle aule sono in cemento armato. L'area d'interesse è limitata alla zona centrale di corridoio.

Il sito è esemplificativo di una situazione "semi-indoor" o di locale "semi-strutturato", "confinato" da pareti a geometria irregolare.

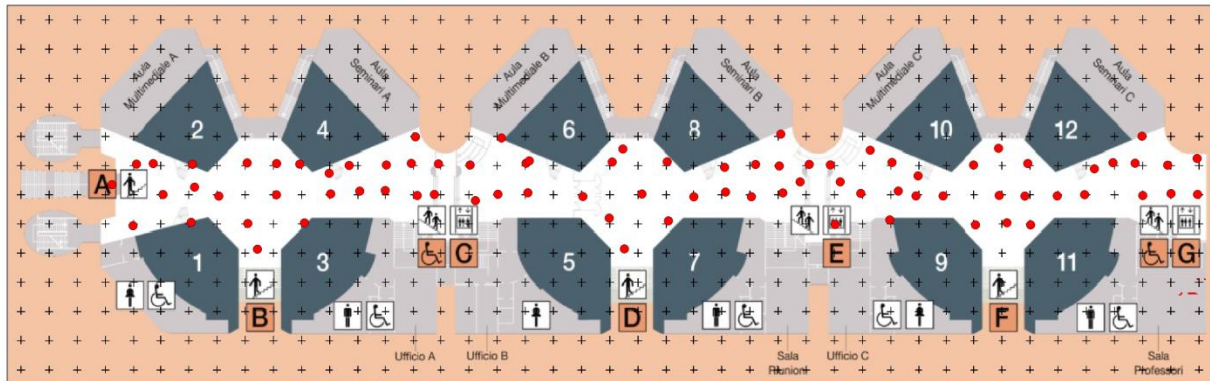
Sono state acquisite misure di RSSI in punti di riferimento e di test in condizioni di scarsa presenza di pubblico (dataset "PoloP1_VR" e "PoloP1_VT"), ed in condizioni di alto affollamento (dataset

“PoloP1_PR” e “PoloP1_PT”, con i suffissi VR,PR e VT,VR che stanno ad indicare rispettivamente i Punti di Riferimento ed i Punti di Test).

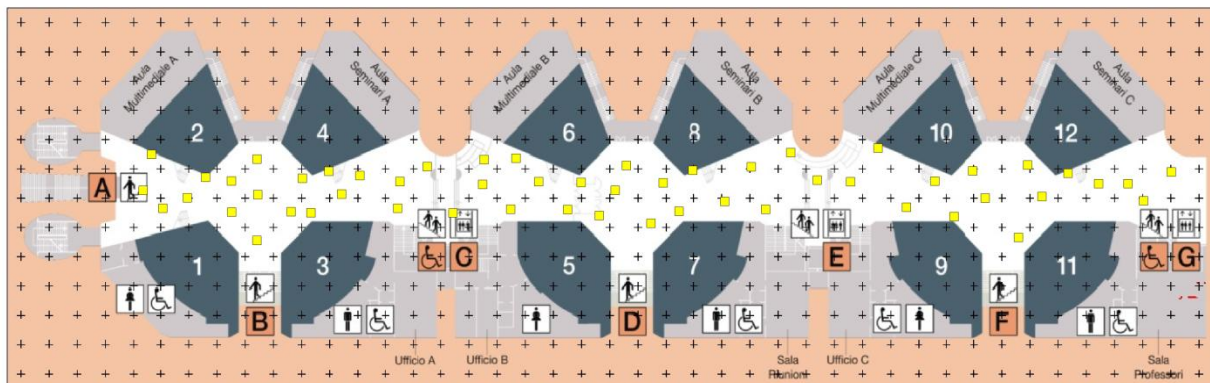
I punti di riferimento sono stati acquisiti sui vertici di una griglia a maglia quadrata di 4 metri, quelli di test preferibilmente in posizioni intermedie.

La rete wi-fi interna è costituita da 12 access point (AP) a 2.4 GHz collocati su una parete interna di ciascuna aula e da 3 AP a doppia emissione (2.4 e 5 GHz), posizionati su totem nelle rotonde del corridoio. Ulteriori 6 AP a 2.4 GHz sono stati aggiunti secondo il modello di net planning descritto nel Deliverable GEOPOS_1 nelle aree a minor copertura.

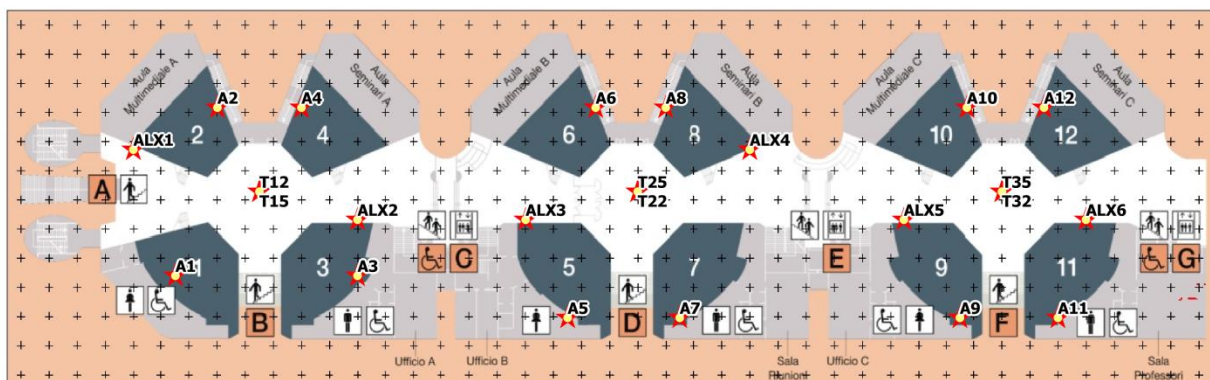
Posizionamento punti di riferimento



Posizionamento punti di test



Posizionamento Access Points



3.2.2 Sito AreaP2

Il sito AreaP2, consiste nell'intero secondo piano del corpo B dell' Area delle Ricerca CNR. Ha pianta quadrata di m 25x25 circa, ed è costituito da una dozzina di studi, prospicienti un

corridoio a U che circonda la zona centrale contenente locali di servizio. Le pareti dei locali sono costituite da tramezzi di gesso. Le porte dei locali sono in metallo pesante tagliafuoco.

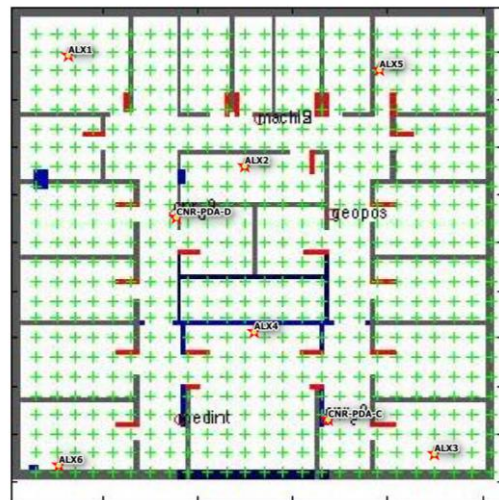
Il sito è esemplificativo di una situazione “indoor” o di locale “strutturato”.

Sono state acquisite misure di RSSI in punti di misura a porte aperte ed a porte chiuse.

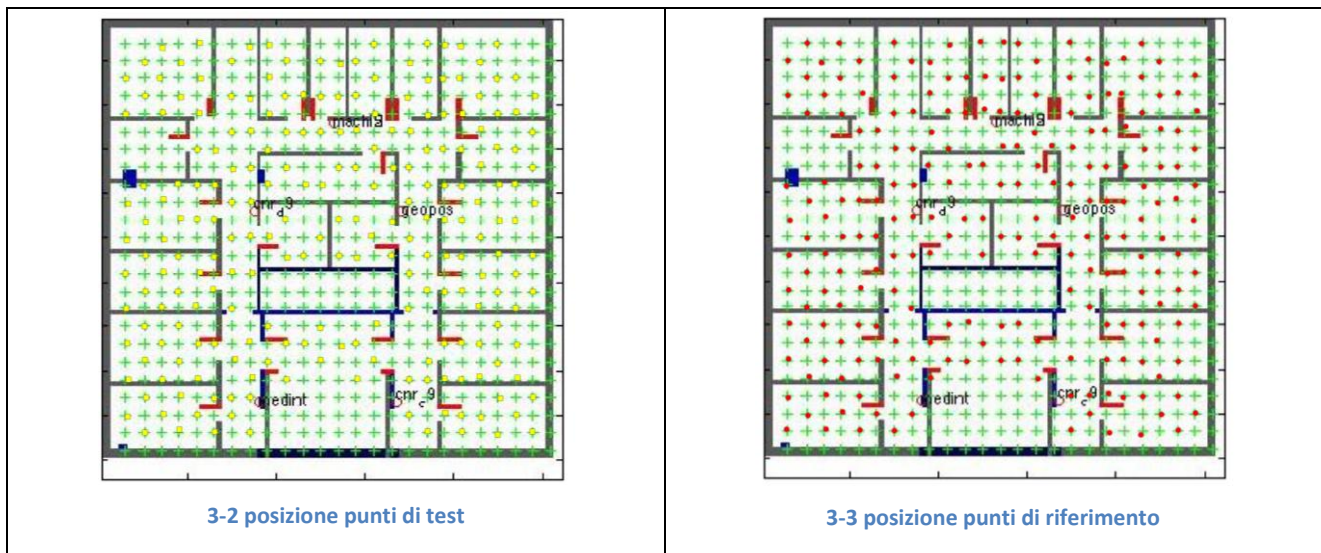
I punti di misura sono stati acquisiti sui vertici di una griglia a maglia quadrata di 1 metro con markers sul pavimento nei locali accessibili e di 2 metri con stima visuale nei restanti.

Dai dataset di misure sono stati estratti per campionamento spazialmente uniforme e decimazione, 5 dataset di punti riferimento (4 a porte aperte a densità decrescente ed 1 a porte chiuse) e due di test come indicato in tabella

DataSet	Condizioni	Punti
AreaP2_AR1	Riferimento, porte aperte, copertura 4 mq	200 approx
AreaP2_AR2	Riferimento, porte aperte, copertura 8 mq	100 approx
AreaP2_AR3	Riferimento, porte aperte, copertura 16 mq	50 approx
AreaP2_AR4	Riferimento, porte aperte, copertura 32 mq	25 approx
AreaP2_CR	Riferimento, porte chiuse, copertura 8 mq	100 approx
AreaP2_AT	Test, porte aperte, copertura 4 mq	200 approx
AreaP2_CT	Test, porte chiuse, copertura 8 mq	100 approx



3-1 posizione Access Points

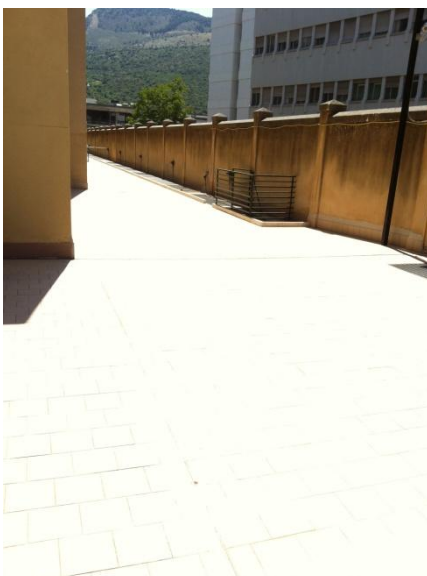


3.2.3 Sito AreaPT

Il sito AreaPT, consiste in una sezione del piazzale a Piano Terra dell’Area delle Ricerca, confinato dalla parete esterna del corpo di fabbrica A dell’edificio e da un muro perimetrale .

La sezione di riferimento è costituita da un’ampia area seguita da un ampio corridoio. Il sito è esemplificativo di una situazione “semi-indoor” o di locale “confinato”.

I punti di misura sono stati acquisiti su un asse latitudinale con origine un AP, a intervalli di uno o due metri utilizzando sia un device mobile (acquisizione client-side, DataSet APT_CS) che un secondo AP (acquisizione network-side DataSet APT_NS); in aggiunta è stato effettuato un set di misure utilizzando il device mobile in tethering come trasmettitore.



3.3 Valutazioni empiriche ed esperimenti

3.3.1 Test A (non effettuato): incertezza sulla misura d'intensità del segnale acquisito al variare del modello del terminale smartphone.

Problema: Il protocollo 802.11.xx non impone una calibrazione della misura del segnale RSSI riportata dai device compliant al di sopra del livello fisico. Tale calibrazione, è effettuata internamente al driver del sistema operativo dello smartphone e risente del posizionamento della antenna di ricetrasmisione del particolare modello di device mobile.

Ipotesi: I valori di RSSI acquisiti con modelli di smartphone diversi possono mostrare bias (spostamento dello zero), differenza di scala per differente guadagno di amplificazione o guadagno di antenna, non linearità nella misura dei livelli di segnale RSSI..

Metodo di stima: (esperimento non effettuato) Misura della differenza nei valori medi del segnale RSSI in punti di riferimento della radio mappa presi con il modello standard (Samsung Galaxy) e con un terminale differente, rapportata alla stima dell'effetto della fluttuazione intrinseca del segnale RSSI misurato con lo smartphone ufficiale.

Evidenze: Comparazioni sporadiche effettuate dagli operatori riportano una differenza fra modelli, indistinguibile dalla fluttuazione intrinseca del segnale e dalla influenza dovuta alla modalità di impugnatura dello stesso.

Valutazione: Nessuna valutazione, allo stato attuale. Si assume che i livelli di segnale acquisiti come interi dall'hardware siano opportunamente siano calibrati da parte del firmware di device e riportati correttamente in dB dal driver del sistema operativo (Android nel nostro caso).

3.3.2 Test B: fluttuazione temporale istantanea del segnale radio e della sua distribuzione spaziale.

Problema: In ciascun un punto di calibrazione, la radio mappa è costruita mediando l'ampiezza istantanea del segnale radio. La localizzazione è falsata dalla fluttuazione incoerente delle ampiezze dei singoli valori RSSI componenti il fingerprint. La misura dell'intensità del segnale è legata a diversi fattori tra cui il riconoscimento (in hardware) del pattern di header del pacchetto di beacon, la gestione dell'antenna diversity in trasmissione [2] ed eventuali procedure di power saving proprietarie delle diverse Network Interface Card [3]. In condizioni di fading differenti pacchetti ad intensità degradata e time shift variabile possono interferire col pacchetto principale ed essere con questo confusi. In condizione di collisione accentuata il segnale può anche non essere sincronizzato nella finestra temporale prevista, dando luogo ad una mancata rilevazione.

Ipotesi: a) La mancanza di rilevazione di segnale è da considerarsi come indicativa di basso segnale disturbato e non di assenza dello stesso; b) punti ad alta variabilità del segnale sono indicativi di disuniformità del campo e non adeguati a essere presi come riferimenti (near neighbours) nella determinazione della posizione tramite interpolazione.

Metodo di stima: esperimento controllato di acquisizione della radio mappa ad alta densità spaziale con elevato numero di samples, calcolo della distribuzione statistica del segnale, filtraggio dei valori improbabili (soglia)

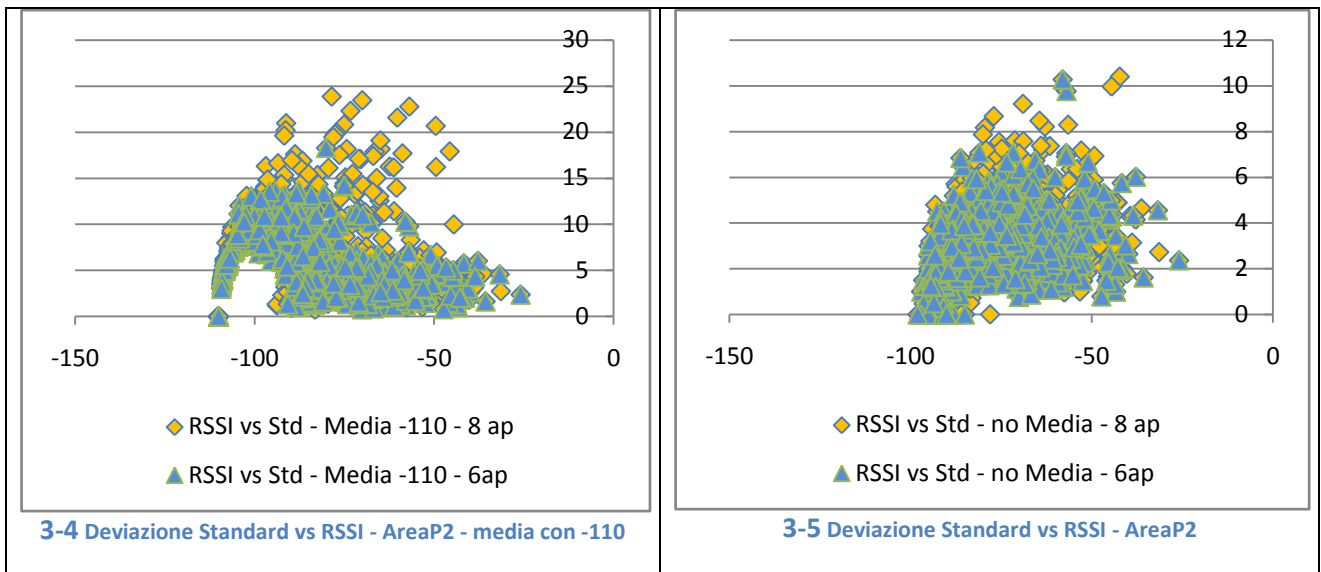
Settings: DataSets AreaP2_AR1(20 campioni per punto), PoloP1_VR (5 campioni per punto)

Evidenze: Le figure 3-4 e 3-5 a seguire mostrano la distribuzione dei valori di deviazione standard in funzione della media del valore acquisito in ogni punto di riferimento, per entrambi i dataset. Per il DataSet AreaP2 è anche evidenziata la maggiore dispersione dei valori

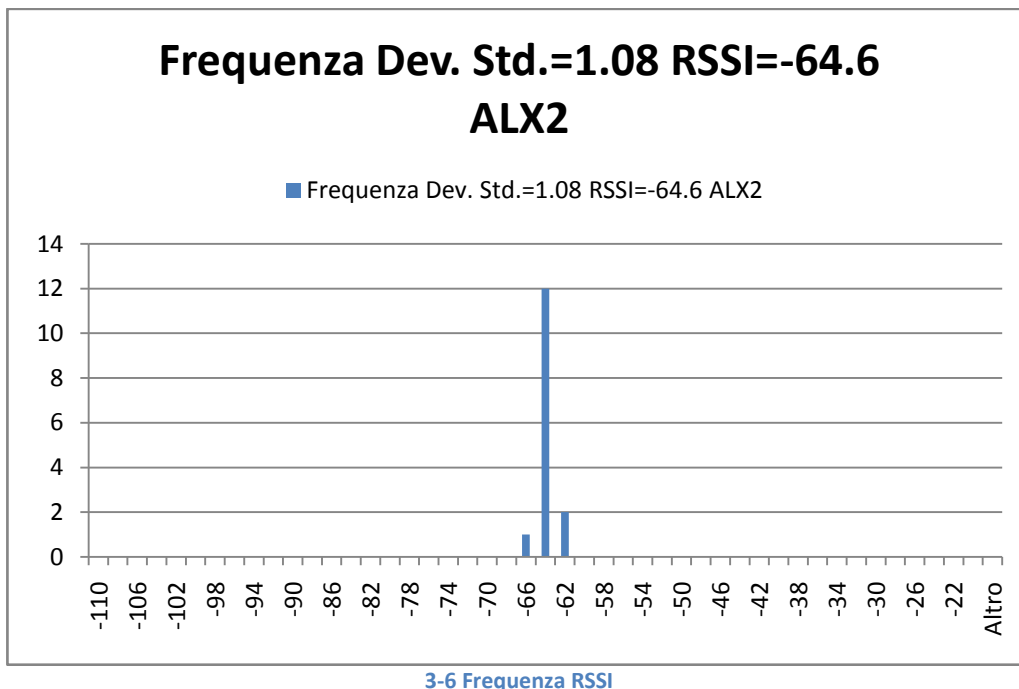
La figura 3-4 mostra invece per il dataset AreaP2_A1 la minore variabilità dei segnali dei 6 AP Alix (in blu), collocati entro gli studi e a distanza dalle pareti, rispetto ai 2 AP SMC (in giallo) collocati sulle pareti del corridoio. Tale grafico è stato ottenuto mediando con un valore fittizio di fondoscala (-110) i campioni effettivamente rilevati per ciascun punto e per ciascun AP, e compensando i valori non rilevati sul totale previsto. Ciò comporta l'attribuzione di un valore di deviazione standard tanto più grande quanto più sono distanti i valori rilevati rispetto al fondoscala e quanto maggiori sono il numero di campioni non rilevati.

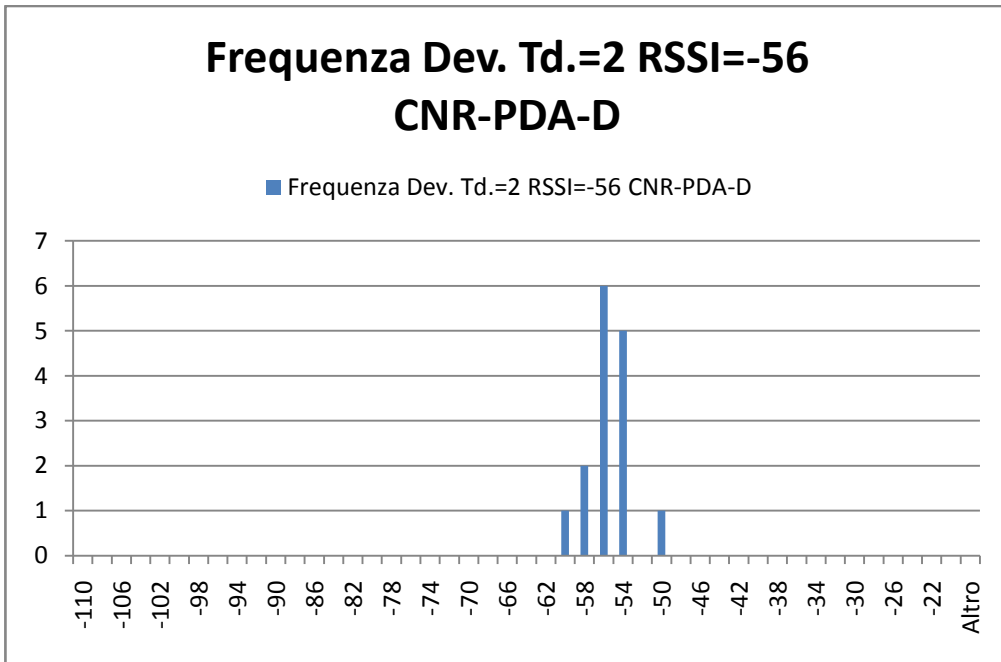
La figura 3-5 mostra invece per il dataset AreaP2_A1 le distribuzioni ottenute non assegnando alcun valore ai campioni non rilevati anziché il valore fittizio di fondo scala (-110 db).

Il confronto tra le due figure permette di rilevare visivamente come la deviazione standard sui valori mediati con il valore fittizio di fondoscala, consenta di discriminare per ogni punto, quegli ap che forniscono un contributo troppo variabile e conseguentemente poco attendibile, esaltando la deviazione standard laddove vi sono pochi valori rilevati o troppo saltuari.

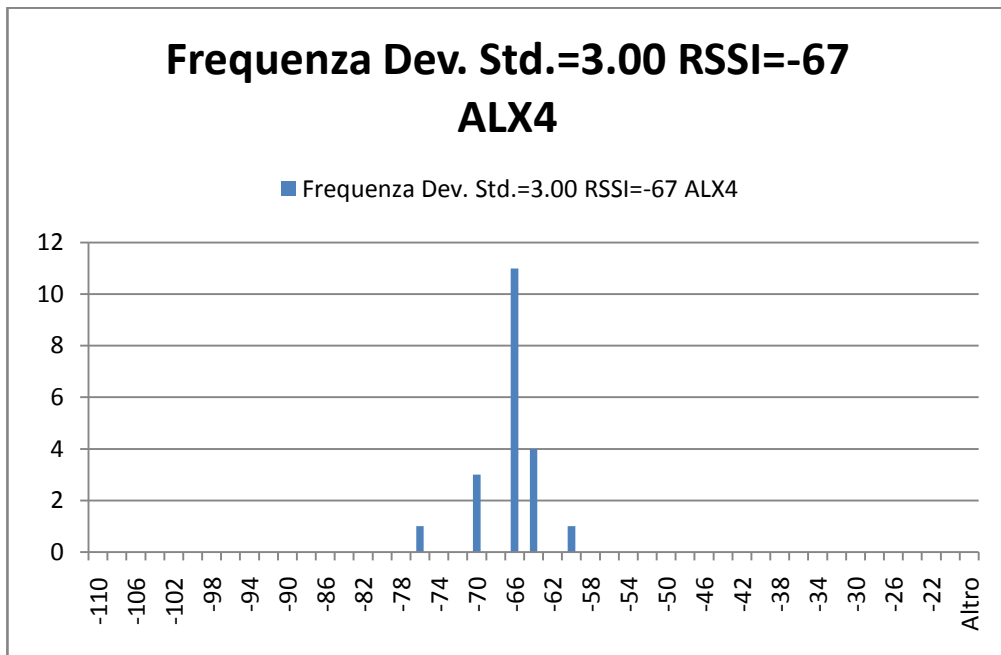


Le figure 3-6, 3-7, 3-8, 3-9, 3-10 3-11 mostrano invece come, all'aumentare della deviazione standard la distribuzione tenda ad apparire multimodale suggerendo una non unicità di popolazione. Questo suggerisce delle modalità di trattamento delle componenti del fingerprint differenziate, a secondo della varianza riscontrata con riferimento ai diversi AP. Risulta chiaro che la media dei valore di RSSI tende a differire dai valori più probabili non appena la deviazione standard aumenta, venendo meno le ipotesi di gaussianità della distribuzione e di conseguenza anche l'assunzione di un valore medio unico utilizzabile nel fingerprinting.





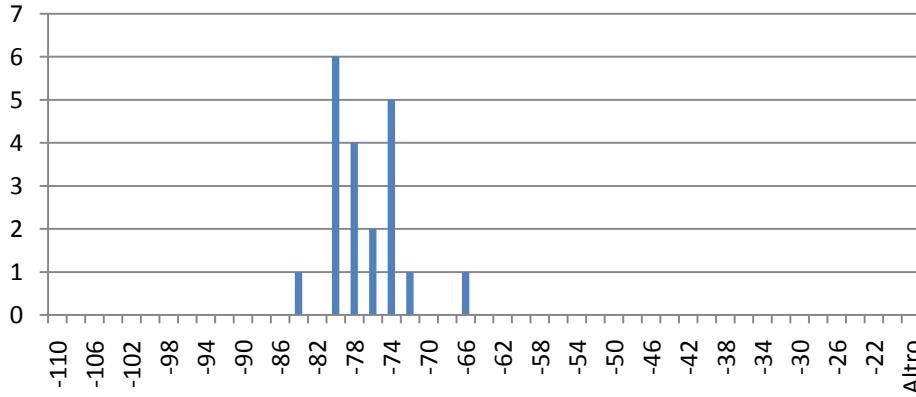
3-7 Frequenza RSSI



3-8 Frequenza RSSI

Frequenza Dev. Std.=4.00 RSSI=-77.4 ALX6

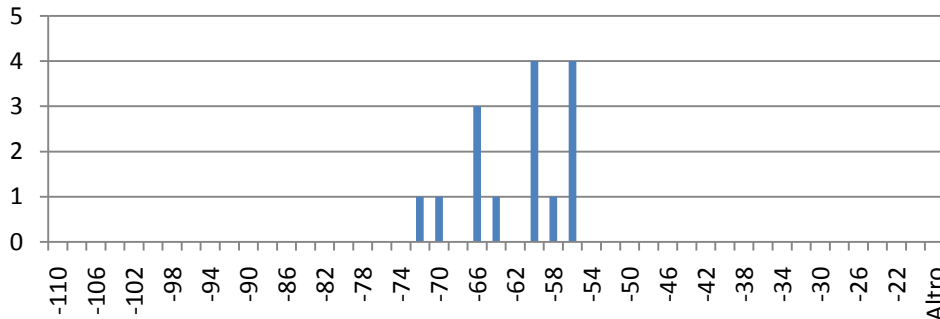
■ Frequenza Dev. Std.=4.00 RSSI=-77.4 ALX6



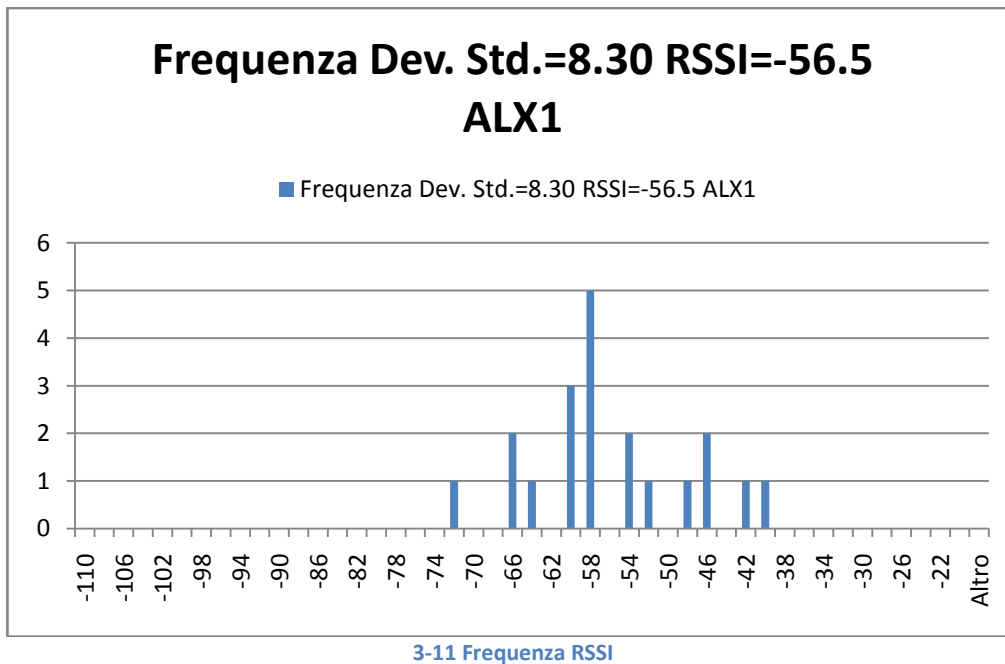
3-9 Frequenza RSSI

Frequenza Dev. Std.=5.00 RSSI=-62.3 ALX1

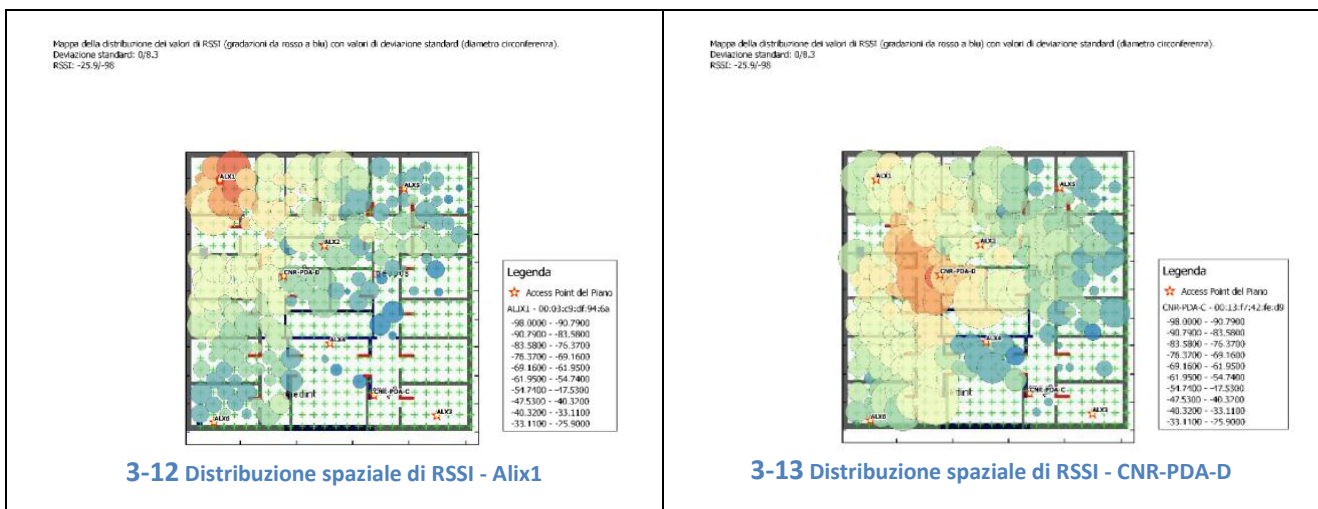
■ Frequenza Dev. Std.=5.00 RSSI=-62.3 ALX1



3-10 Frequenza RSSI



Le figure 3-12 e 3-13 mostrano rispettivamente i valori di campo di un AP Alinx (bassa varianza) e di un AP SMC (alta varianza) . Si evidenzia la concentrazione dei punti ad alta varianza nelle vicinanze della sorgente.



Valutazione: La condizione di segnale non rilevato va considerata come indicazione di basso campo e va assegnato ad essa un valore di saturazione al minimo effettivamente rilevato. La alta variabilità del segnale appare correlata a condizioni geometriche quali vicinanza alla sorgente, probabilmente anche a riflessioni in sua prossimità. In mancanza di modellazione maggiormente probatoria l'affidabilità dei punti di riferimento ad alta variabilità è limitata.

3.3.3 Test C: effetto della fluttuazione interna delle componenti del pattern di fingerprint

Problema: Il fingerprint di un punto di riferimento sulla radio mappa è costituito dalla ennupla ordinata (vettore n dimensionale) dei *valori medi* di campo nel punto, uno per ogni AP.

Il fingerprint del punto di test è costituito dalla ennupla dei valori *istantanei* del campo nel punto con lo stesso ordine.

La dissimilitudine fra il pattern in un punto di riferimento e quello nel punto di test è espressa dalla distanza euclidea fra i rispettivi vettori.

La stima della posizione del cliente mobile avviene, infine, selezionando dalla radiomappa i punti di riferimento con pattern di fingerprint più simile a quello del punto di test ed attribuendogli il baricentro delle loro posizioni note .

Nell'algoritmo in uso, ogni componente del fingerprint contribuisce alla distanza euclidea proporzionalmente alla differenza fra il valore istantaneo nel punto di test ed valore medio atteso nella radio mappa, indipendentemente dalla posizione geografica del punto di test.

In presenza di campi decrescenti monotonamente a simmetria circolare, i metodi di triangolazione mostrano come siano sufficienti i valori di campo di tre AP a localizzare univocamente una area con il pattern cercato.

Nel caso di campi a geometria complessa, più aree possono casualmente vantare distanze simili fra i rispettivi fingerprint e quello del punto di test. L'aumentare del numero di componenti nel fingerprint (AP) aumenta la probabilità di esistenza di un solo minimo locale, mentre, invece, la variabilità del valore atteso sulla radio mappa, opera come rumore ed aumenta la probabilità di esistenza di più minimi locali nella funzione distanza complessiva.

Ipotesi1: Se nel comparare la distanza fra il fingerprint del punto di test e quello dei punti di riferimento si filtrano le componenti relative ad AP il cui segnale è soggetto ad alta fluttuazione, si si aumenta la probabilità di individuare punti a pattern simile principalmente fra i punti vicini spazialmente (recall) e quindi, indirettamente, si migliora la accuratezza di localizzazione.

Ipotesi2: Filtrando le componenti relative ad AP il cui segnale è soggetto ad alta fluttuazione, si aumenta la probabilità di individuare punti a pattern simile principalmente fra i punti vicini spazialmente (recall) e quindi, indirettamente, si migliora la accuratezza di localizzazione

Metodo di stima:

Dato l'insieme $\{V_k\}$ dei punti di riferimento ordinato per distanza di pattern dal punto di test tramite l'algoritmo di clustering KNN , e l'insieme $\{V_d\}$ dei punti di riferimento ordinati per distanza geometrica dal punto di test, si assume come indice di recall in funzione di k , il numero di vicini compresi fra i primi k in $\{V_d\}$ presenti fra i primi k in $\{V_k\}$, normalizzato su k .

Data sets:

AreaP2_AR1 (riferimento) vs AreaP2_AT

Evidenze:

La figura 4.x mostra l'errore medio di localizzazione ottenuto calcolando la posizione del punto di test come centroide dei primi k vicini trovati con l'algoritmo KNN. Il minimo si raggiunge velocemente per $k=4$ e l'errore ricresce leggermente all'aumentare di k .

La figura 4.y mostra come la varianza dell'errore decresca monotonamente all'aumentare di K .

La figura 5 mostra come la capacità di recall dell'algoritmo KNN, si mantenga al di sotto del 50%, anche se migliorato dall'assegnazione del valore di minimo al segnale non rilevato.

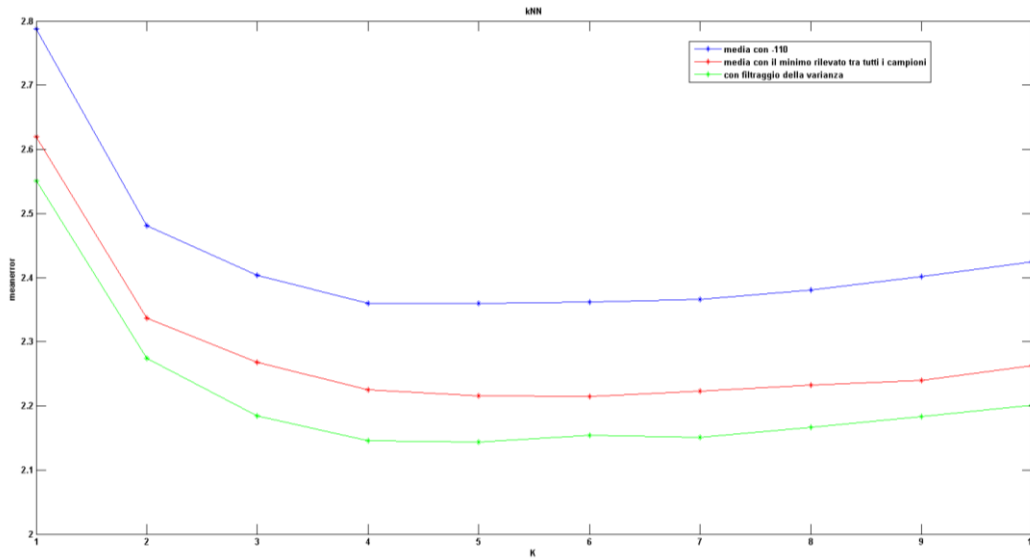


Figura 4.x AreaP2 Media dell'errore al variare di k

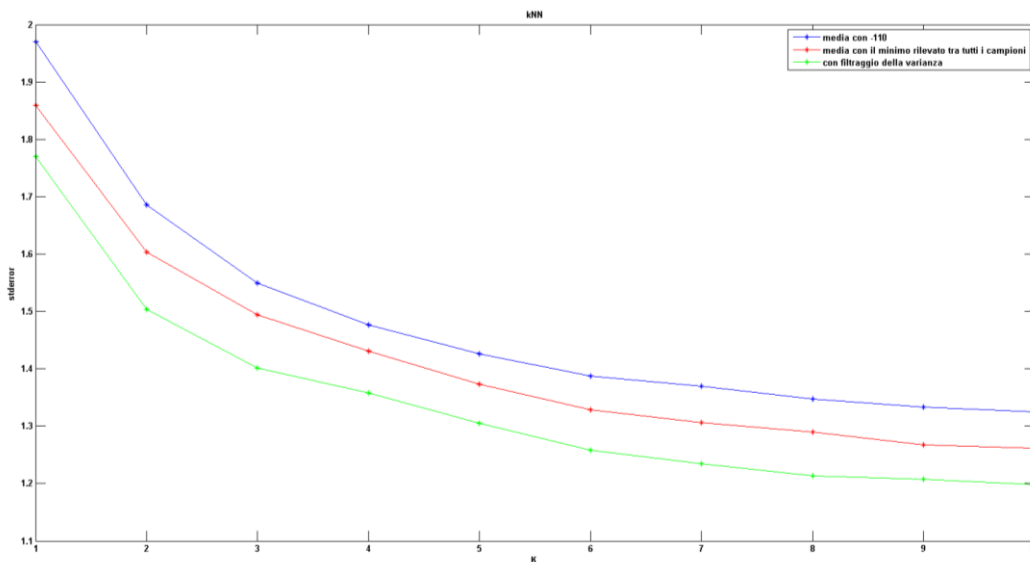


Figura 4.y AreaP2 Deviazione standard dell'errore al variare di k

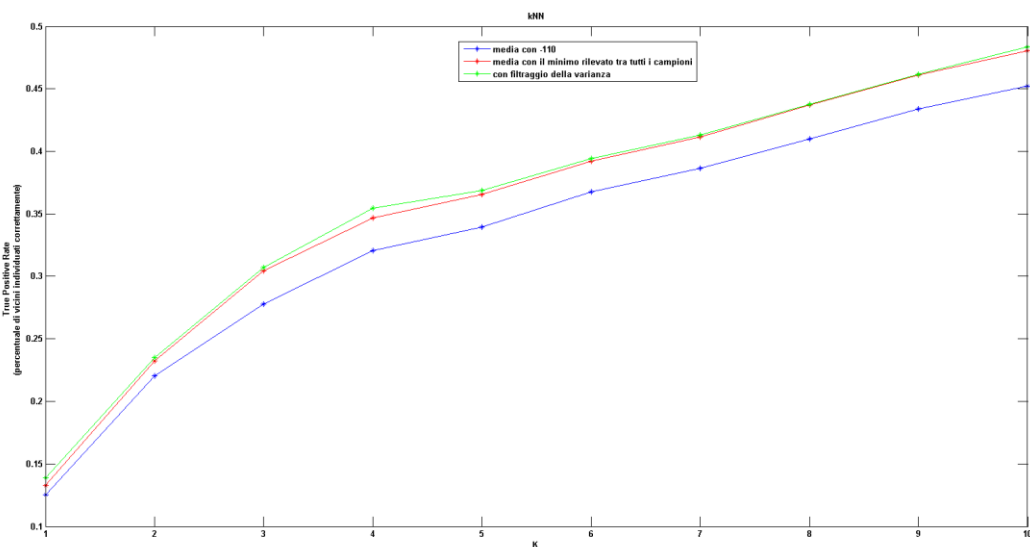


Figura 5 AreaP2 Percentuale di vicini individuati correttamente al variare di k

Valutazioni:

L'algoritmo di KNN ritrova, nel proprio intorno, una frazione di vicini simili corretti superiore a quella degli algoritmi di autoapprendimento utilizzati. Solo pochi però entrano nella rosa dei primi k selezionati, quindi l'indice di recall varia solo leggermente. L'effetto non si evidenzia come aumento significativo della percentuale di vicini correttamente ritrovati fra i primi k, ma come miglioramento della differenza fra i rank assegnati in $\{V_d\}$ e $\{V_k\}$

3.3.4 Test E: Influenza della densità di campionamento dei punti di riferimento

Problema: Il fingerprint di un punto di riferimento del sito rappresenta il valore medio dei campi nel punto, estensibili per interpolazione ad una zona limitrofa in cui il campo mantiene caratteristiche di monotonicità e regolarità (derivata seconda monotona). Come si può evincere dalle mappe di campo riportate in Appendice B, tale condizione non è rispettata nelle vicinanze degli AP, nelle zone limitrofe alla lineOfSight al di là di aperture, in presenza di deviazioni nei corridoi e dovunque la geometria del sito produca interferenza significativa del segnale diretto (line of site) e di quello riflesso, parzialmente assorbito, e scatterato (shadow).

Ipotesi: All'aumentare della densità del campionamento decresce l'area di copertura del singolo punto di riferimento ed aumenta la rappresentatività della sua signature.

Supponendo corretta la individuazione dei vicini da parte dell'algoritmo di clustering, l'indeterminazione introdotta dalla interpolazione fra vicini (baricentro) è comparabile o inferiore alla distanza media fra i punti di riferimento.

Metodo di stima:

Analisi della distribuzione della accuratezza di localizzazione dei punti di test in funzione della distanza media dei punti di riferimento (punti di campionamento della radio mappa), utilizzando l'algoritmo pretrattamento degli outliers ed analisi con l'algoritmo più performante (KNN con $k=3$)
Esclusione dalla funzione distanza delle componenti con alta varianza

La distanza media è calcolata come $\sqrt{\text{areaSitoCampionata}/n\text{PuntiRiferimento}}$, l'area del CNR, AreaP2 è stimata in circa 580 m², l'area utile del Polo didattico, PoloP1, è stimata in circa 2250 m².

Data sets:

AreaP2_AR1 (riferimento) vs AreaP2_AT (test)

PoloP1_VR (riferimento) vs PoloP1_VT

Evidenze:

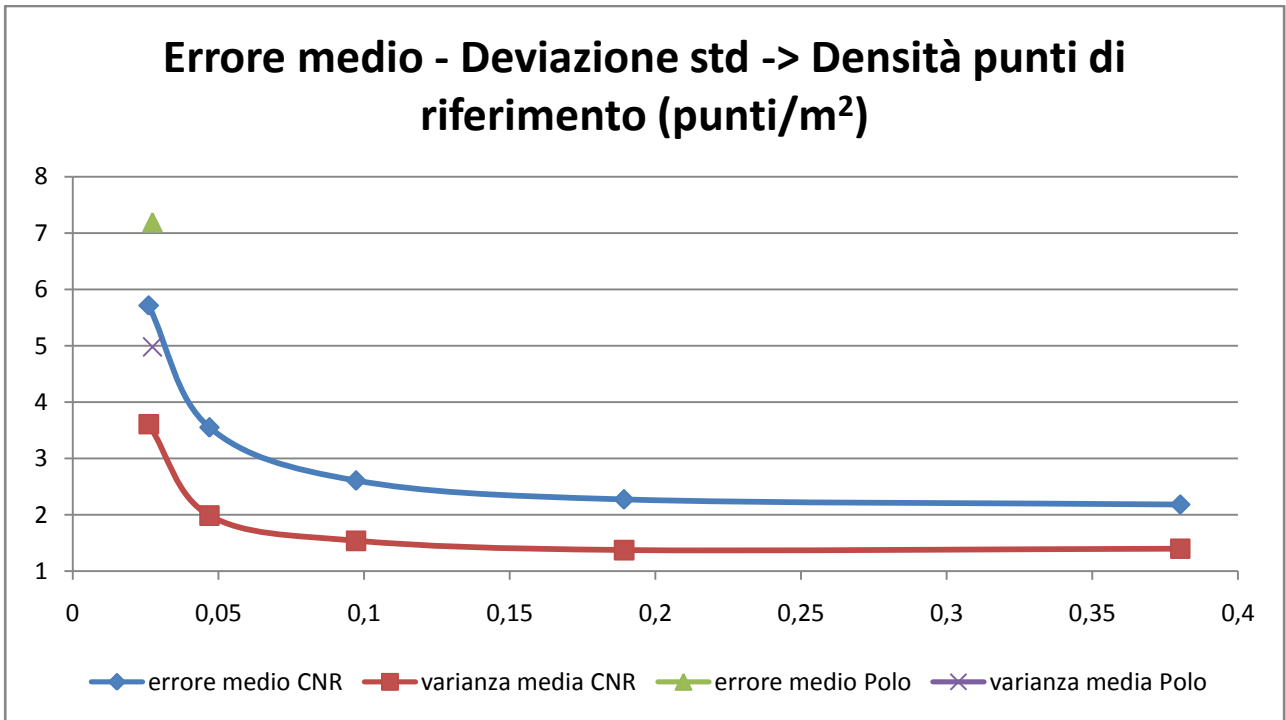
Per il sito AreaP2 L'accuratezza varia da 2.2 per distanza media dei punti 2m a 5.7 per distanza media dei punti pari a m 6.6, con una precisione che varia tra 1.4 m e 3.6 m.

Per il sito PoloP1 l'accuratezza è pari a 7.2 metri circa per una distanza media dei punti di 7 m, con una precisione di circa 5 m.

Valutazioni:

L'indeterminazione introdotta dalla interpolazione fra vicini (baricentro) è comparabile o inferiore alla distanza media fra i punti di riferimento.

Oltre un certo valore di densità di punti di rilevazione, il miglioramento nell'errore di localizzazione che si può ottenere potenzialmente, non è proporzionale all'incremento del numero di punti, e di conseguenza ai costi supportati, per cui esiste un valore ottimale di densità, oltre cui non è conveniente andare, dal momento che le variazioni dei segnali sono tali da non riuscire ad ottenerne vantaggio.



3-14 Grafico andamento errore di localizzazione/Densità

3.3.5 Test F: scelta dell'algoritmo di localizzazione

Problema:

k-nearest-neighbor (kNN)

In questo metodo vengono cercati i k fingerprint di riferimento più vicini (dal punto di vista della distanza euclidea) ai fingerprint di test. Successivamente vengono mediate le k posizioni più vicine, ottenendo una stima della posizione del punto di test.

clustering con k-means

I fingerprint di riferimento vengono divisi in due cluster con l'algoritmo kmeans. L'obiettivo dell'algoritmo kmeans è di minimizzare la varianza totale intra-cluster. Ogni cluster viene identificato mediante un centroide. L'algoritmo segue una procedura iterativa:

- inizialmente crea K partizioni e assegna ad ogni partizione i punti di riferimento casualmente
- calcola il centroide di ogni partizione
- costruisce una nuova partizione associando ogni punto di riferimento al cluster il cui centroide è più vicino ad esso
- vengono ricalcolati i centroidi per i nuovi cluster e così via, finché l'algoritmo non converge (i centroidi non cambiano più).

Per ogni punto di test si cerca il fingerprint di riferimento più vicino (dal punto di vista della distanza euclidea). I "vicini" del punto di test saranno i punti di riferimento appartenenti allo stesso cluster del punto di riferimento trovato più vicino al punto di test. La procedura continua iterativamente con la riclusterizzazione dei "vicini" del punto di test in due cluster. I nuovi "vicini" del punto di test saranno i punti di riferimento appartenenti allo stesso cluster del punto di riferimento trovato più vicino al punto di test. La procedura converge quando il numero di "vicini" è minore di k (k=10).

Apprendimento del rank

Per ogni fingerprint istantaneo di test, viene creato un vettore composto dal fingerprint istantaneo di test e dal fingerprint medio di ogni punto di riferimento. In totale, per ogni fingerprint istantaneo di test vengono creati tanti vettori quanti punti di riferimento abbiamo. Ad ogni vettore di questo tipo viene associato un rango (l'ordine in cui si trova la distanza euclidea, nello spazio delle coordinate tra il punto di test e quello di riferimento, tra le distanze dal punto di test a tutti i punti di riferimento).

L'algoritmo kNN supervisionato viene usato per costruire un modello di apprendimento del rango da un insieme di training.

Data sets:

AreaP2_AR1 (riferimento) vs AreaP2_AT (test)

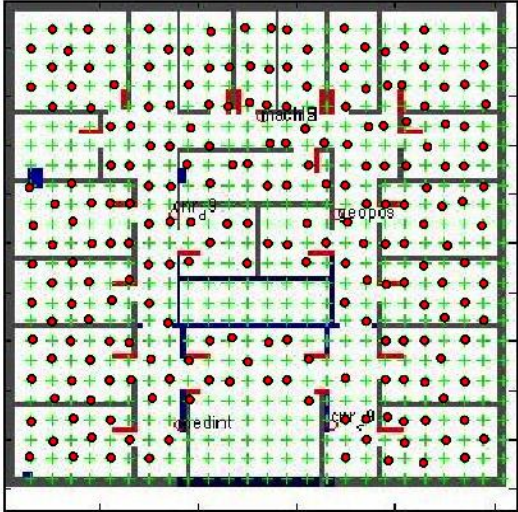
PoloP1_VR (riferimento) vs PoloP1_VT

Evidenze:

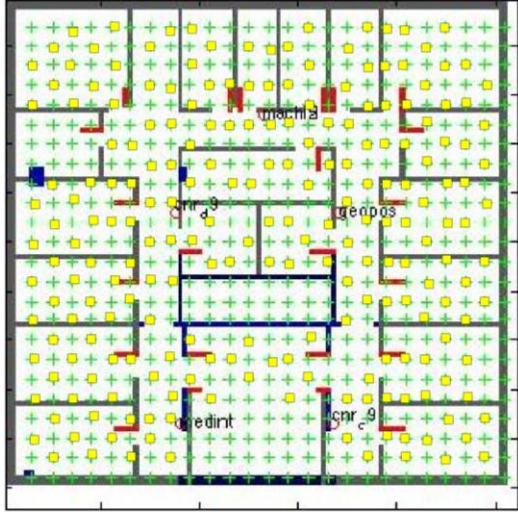
k-nearest-neighbor (kNN)

La media e la deviazione standard dell'errore nel testbed AreaP2 con 219 punti di riferimento, 220 punti di test e 8 AP

Posizione dei punti presi come riferimento



Posizionamento punti di test



Posizionamento Access Points

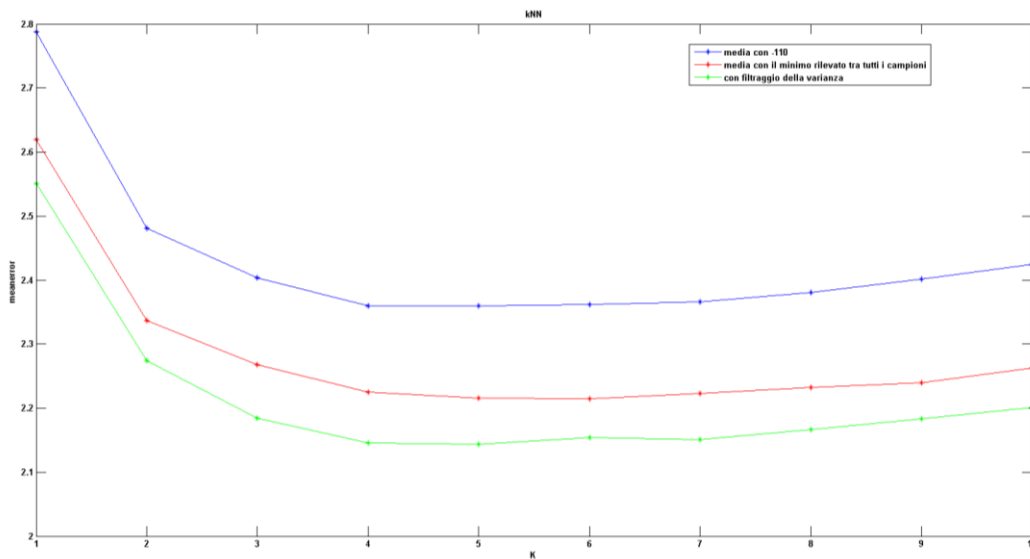
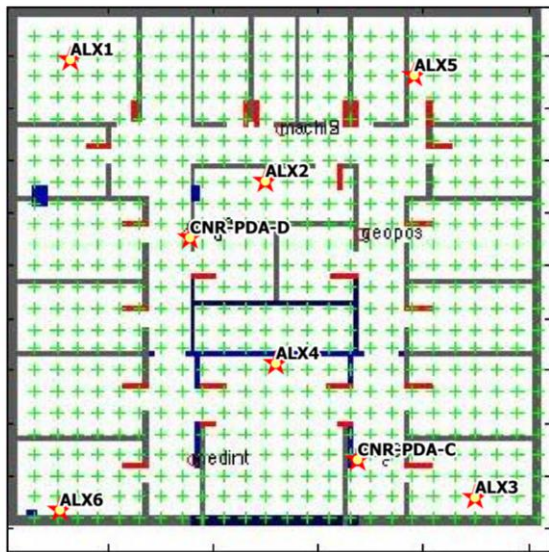


Figura 3-1 AreaP2 Media dell'errore al variare di k

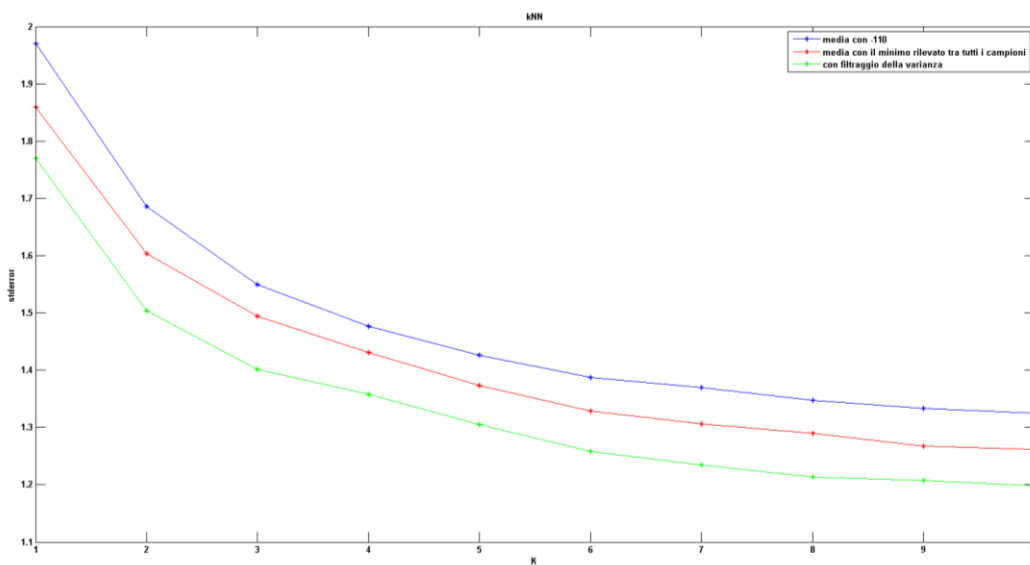


Figura 3-2 AreaP2 Deviazione standard dell'errore al variare di k

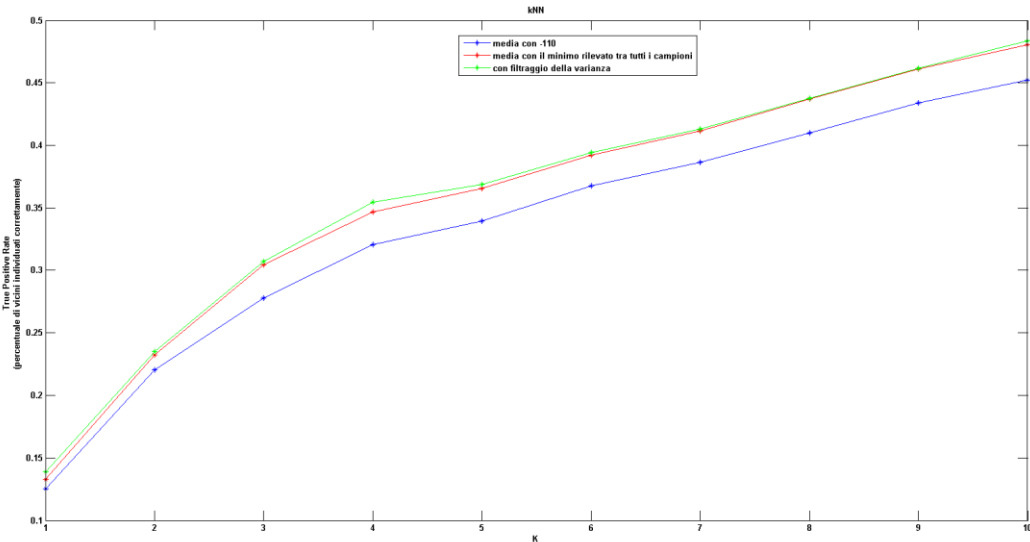


Figura 3-5 AreaP2 Percentuale di vicini individuati correttamente al variare di k

Nella Figura 3-5 si può notare la percentuale di vicini individuati “correttamente” tra i primi k (non necessariamente nella posizione giusta), al variare di k.

Per esempio con il filtraggio della varianza per un punto di test vengono individuati i primi 10 vicini di riferimento che sono: in prima posizione quello che doveva essere (considerata la distanza euclidea nello spazio delle coordinate) in quarta posizione; in seconda posizione quello che doveva essere in 20esima posizione e così via

4 20 12 1 6 35 36 14 28 9.

Quindi per k=1 vengono individuati 0 vicini correttamente, per k=2 0, per k=3 ancora 0, ma per k=4 2 vicini vengono individuati correttamente su 4 (anche se non nella posizione giusta). In un altro caso, in cui tra i primi 4 vicini vengono individuati correttamente sempre 2, ma i vicini individuati invece di essere, nello spazio delle coordinate, il primo e il quarto sono il primo e il secondo, la stima del punto di test migliora.

Si può dimostrare quanto detto sopra usando il coefficiente di rank correlation di Spearman. Il coefficiente di Spearman si calcola

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

dove d_i è la differenza tra ranghi

$$d_i = x_i - y_i$$

e n è la lunghezza di x e di y.

Per esempio, nel nostro caso $n=4$, $x=[2\ 5\ 1\ 3]$ è l'ordine stimato e y è l'ordine corretto, cioè $y=[1\ 2\ 3\ 4]$. In questo caso il coefficiente di Spearman è di 0.9000. In un altro caso, sempre con $n=4$, $x=[3\ 1\ 2\ 7]$ e $y=[1\ 2\ 3\ 4]$, con un coefficiente di Spearman di 0.7000. Nel primo caso, l'errore di stima è di 0.5080, mentre nel secondo caso l'errore è di 0.8703.

Nella Figura 3-6 si può notare la variazione del coefficiente medio di Spearman al variare di k.

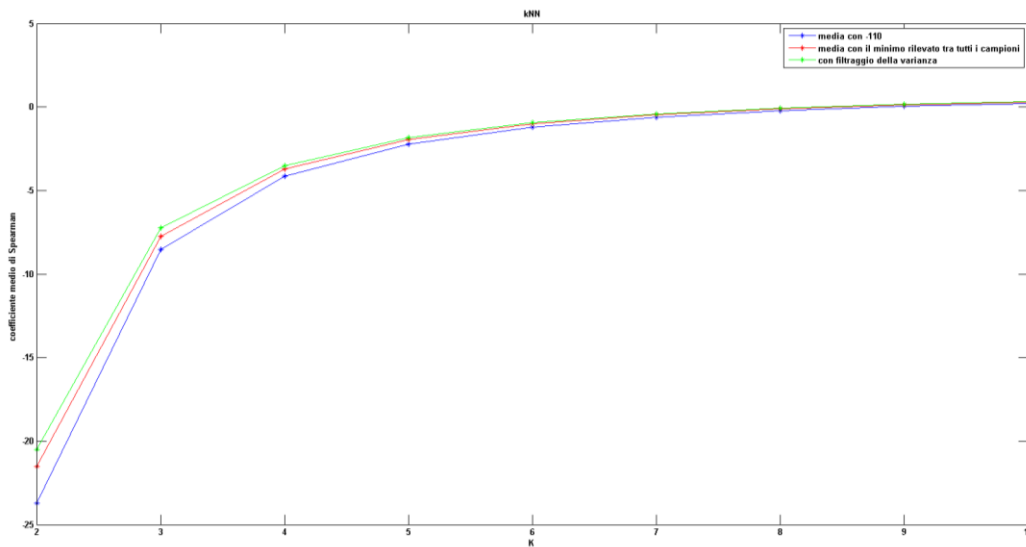
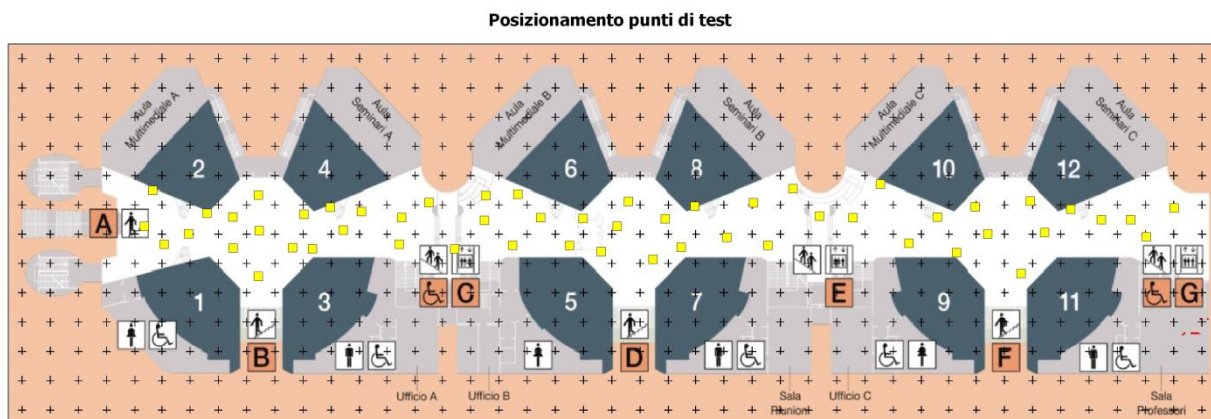
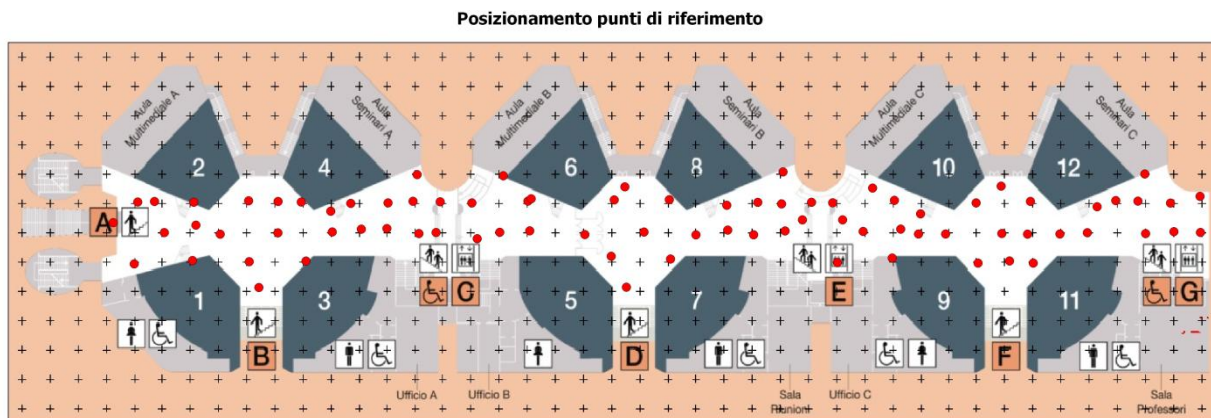


Figura 3-6 AreaP2 Coefficiente medio di Spearman al variare di k

La media e la deviazione standard dell'errore nel testbed PoloP1 con 82 punti di riferimento, 54 punti di test e 24 AP varia al variare di k come mostrato nelle Figure 3-7 e 3-8.



Posizionamento Access Points

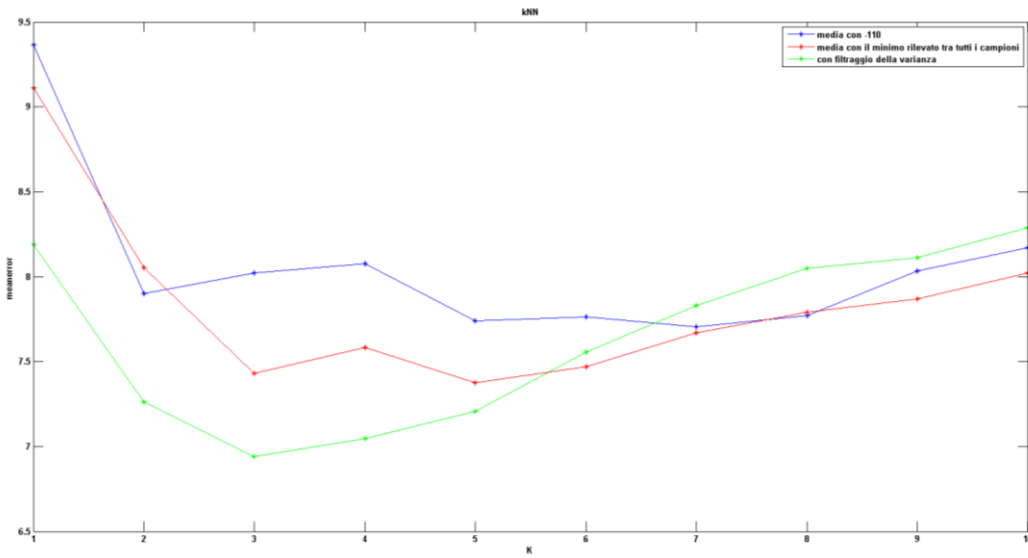
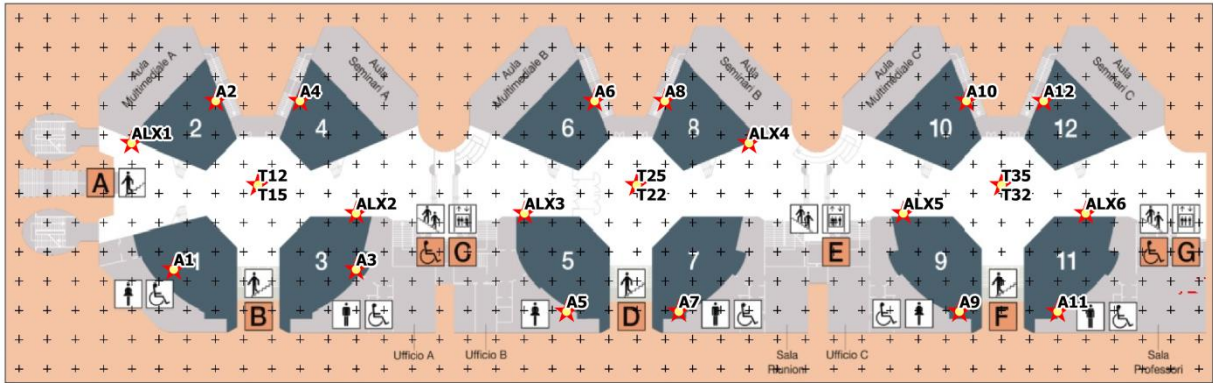


Figura 3-7 PoloP1 Media dell'errore al variare di k

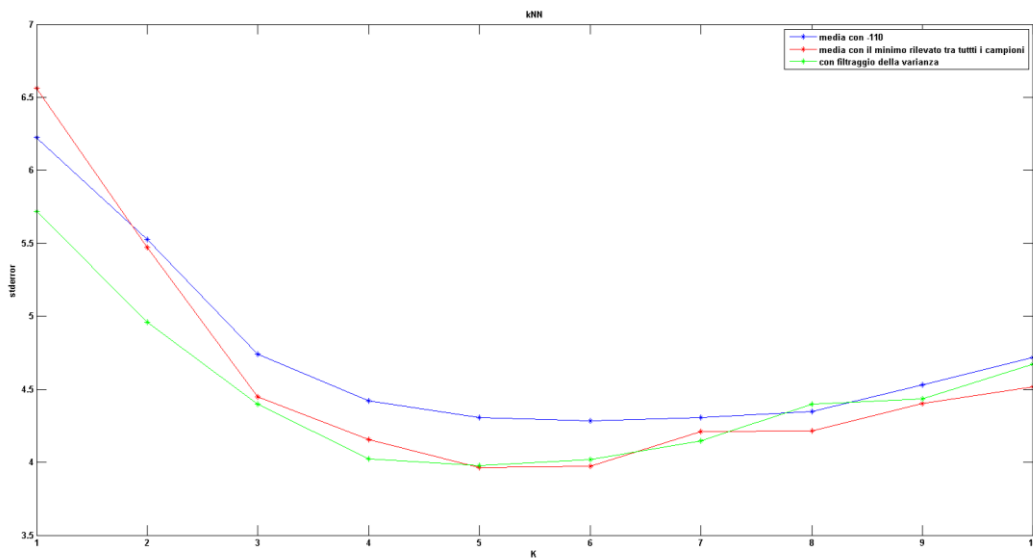


Figura 3-8 PoloP1 Deviazione standard dell'errore al variare di k

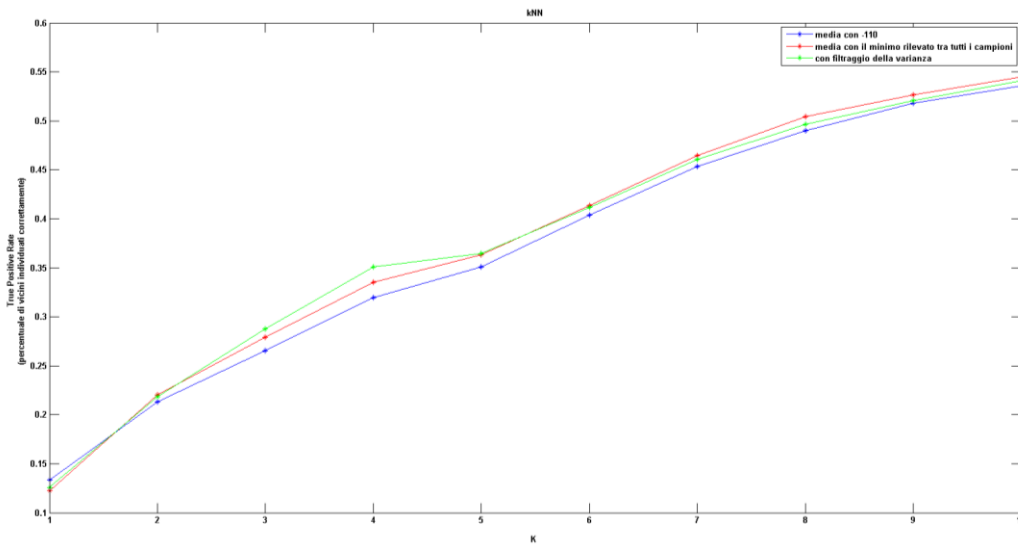


Figura 3-9 PoloP1 Percentuale di vicini individuati correttamente al variare di k

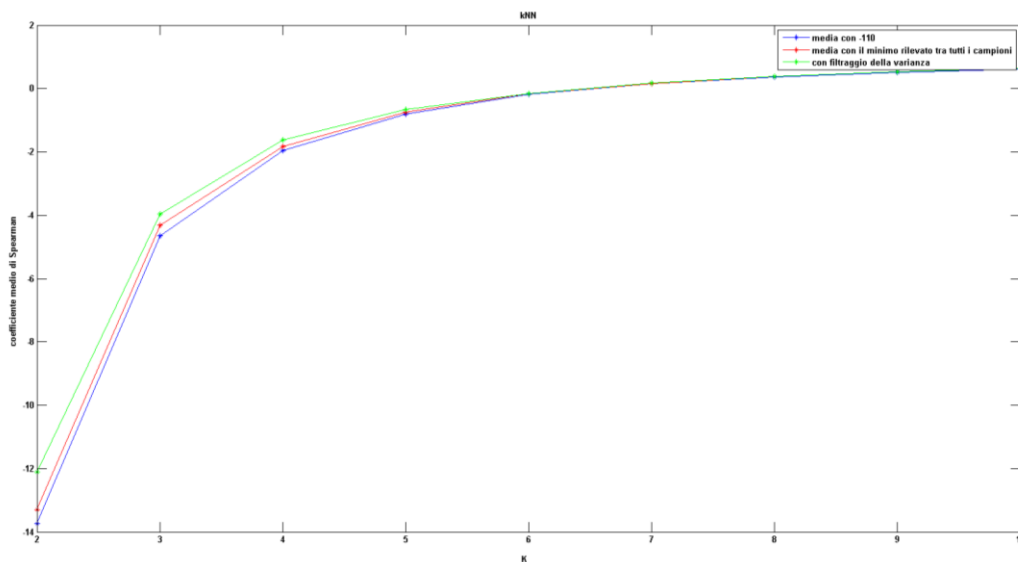


Figura 3-10 PoloP1 Coefficiente medio di Spearman al variare di k

clustering con k-means

Essendo kmeans un algoritmo non stabile (ad ogni run si ottengono risultati diversi), per ottenere i risultati esposti nelle tabelle in basso si fa una media su 10 run.

area	mean	std
media con -110	2.8125	1.7825
con filtraggio della varianza	2.5739	1.5812

polo	mean	std
media con -110	8.8003	6.0190
con filtraggio della varianza	8.1623	5.7022

apprendimento del rank

Nella Figura 3-11 si può vedere come è stato diviso l'insieme dei punti di test del testbed area, in punti di training e punti di test.

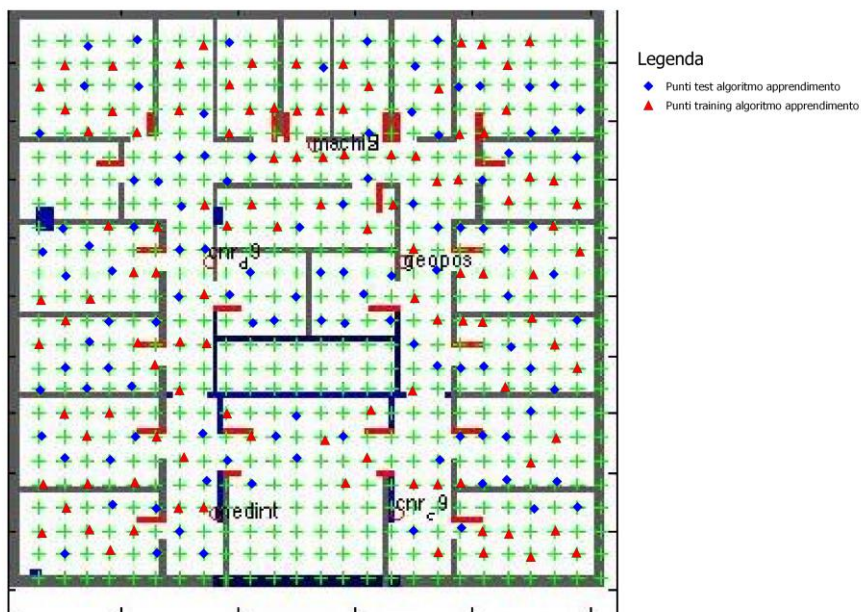


Figura 3-11 AreaP2 Punti di training e punti di test

La media e la deviazione standard dell'errore nel testbed AreaP2 varia al variare di k come mostrato nelle Figure 3-12 e 3-13.

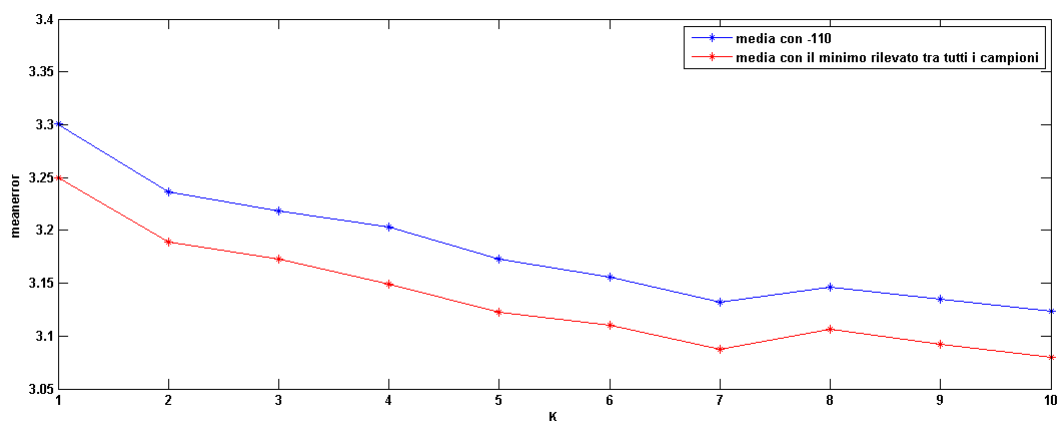


Figura 3-12 AreaP2 Media dell'errore al variare di k

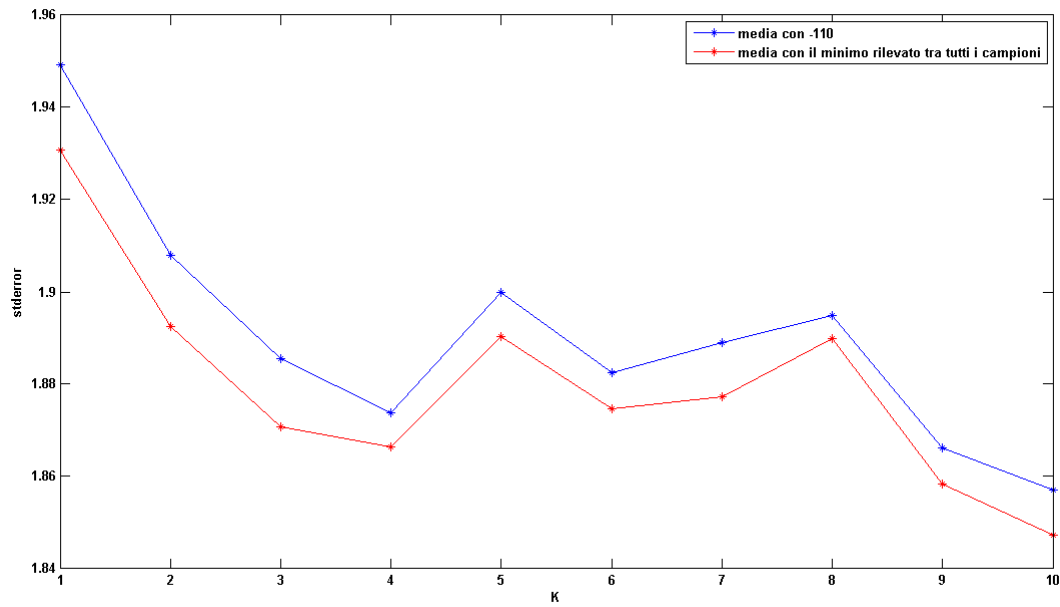


Figura 3-13 AreaP2 Deviazione standard dell'errore al variare di k

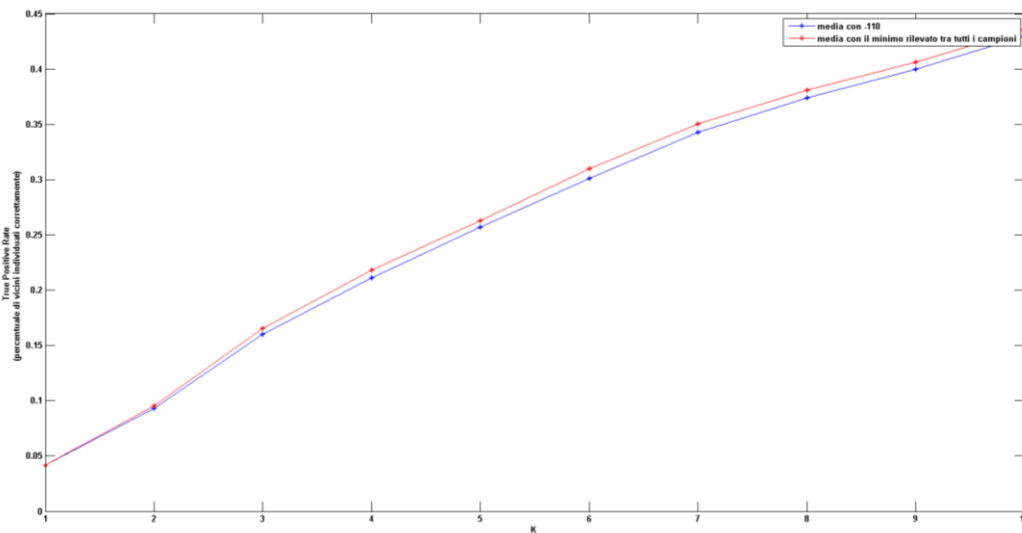


Figura 3-14 AreaP2 Percentuale di vicini individuati correttamente al variare di k

Usando una divisione diversa in punti di training e punti di test la media e la deviazione standard dell'errore migliorano. (Figura 3-16 e Figura 3-17)

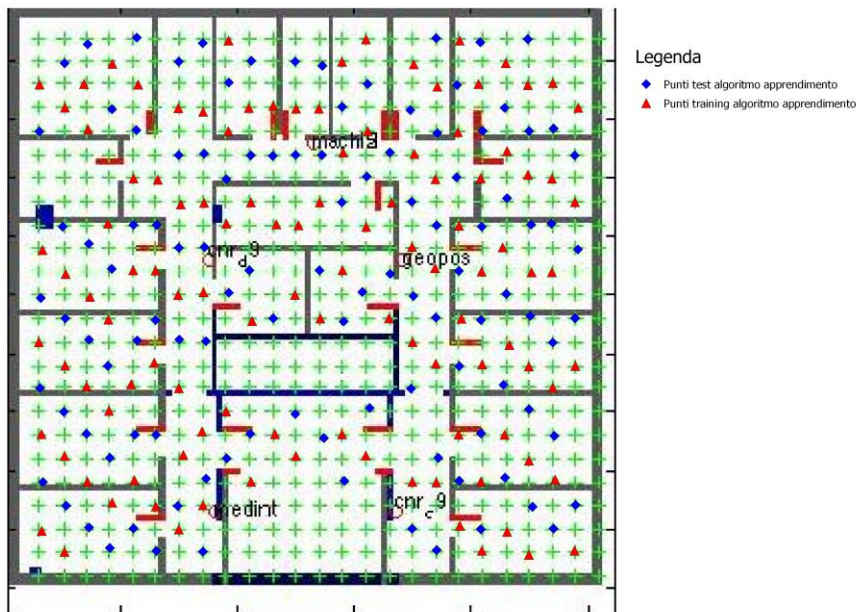


Figura 3-15 AreaP2 Nuova divisione in punti di training e punti di test

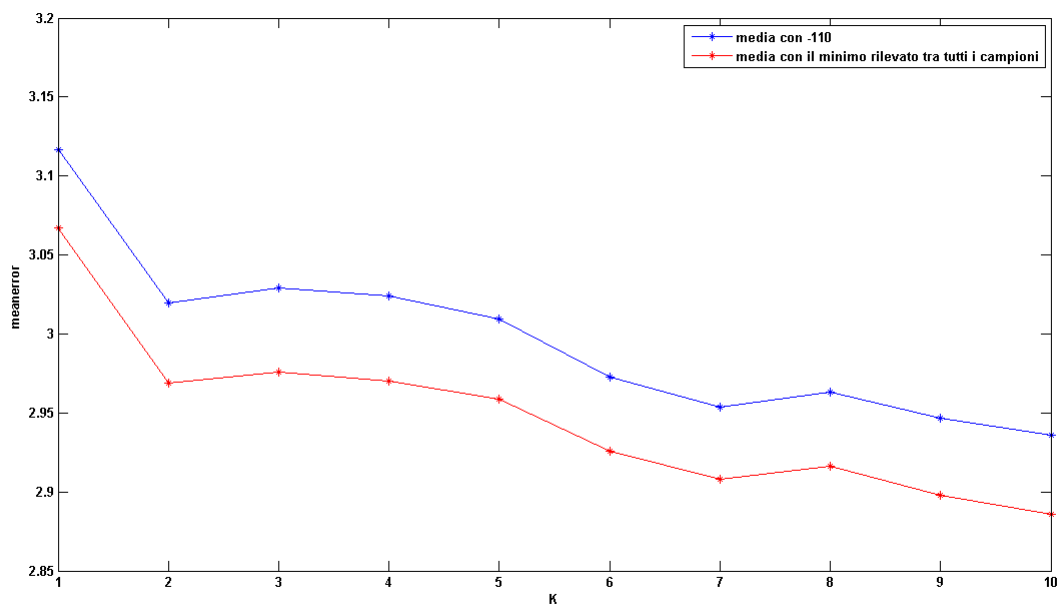


Figura 3-16 AreaP2 Media dell'errore al variare di k (nuova divisione in training e test)

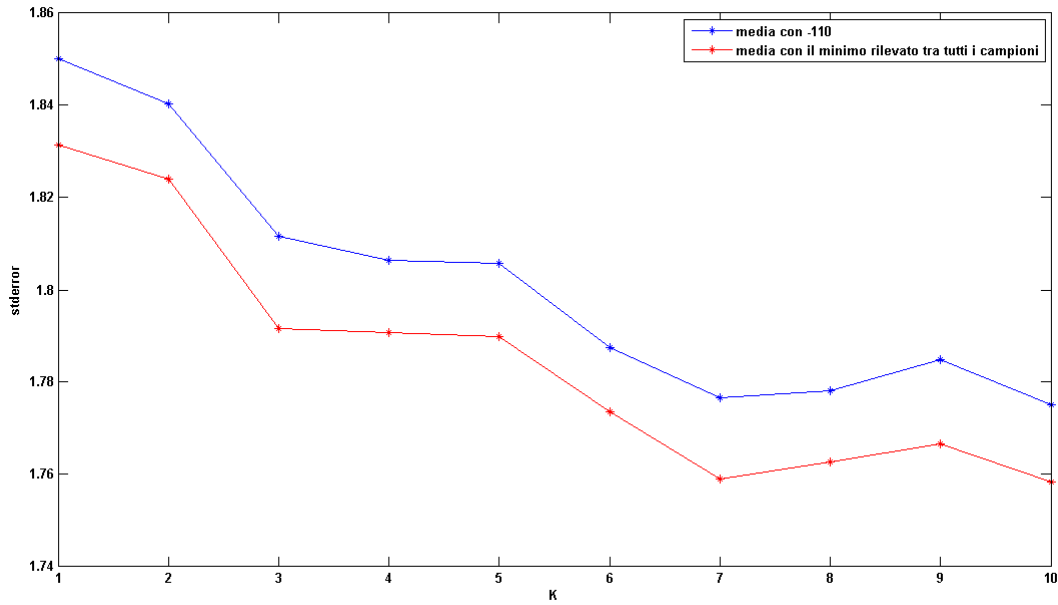


Figura 3-17 AreaP2 Deviazione standard dell'errore al variare di k (nuova divisione in training e test)

Nella Figura 3-18 si può vedere come è stato diviso l'insieme dei punti di test del testbed PoloP1, in punti di training e punti di test.

Legenda

- ▲ Punti training algoritmo apprendimento
- ◆ Punti test algoritmo apprendimento

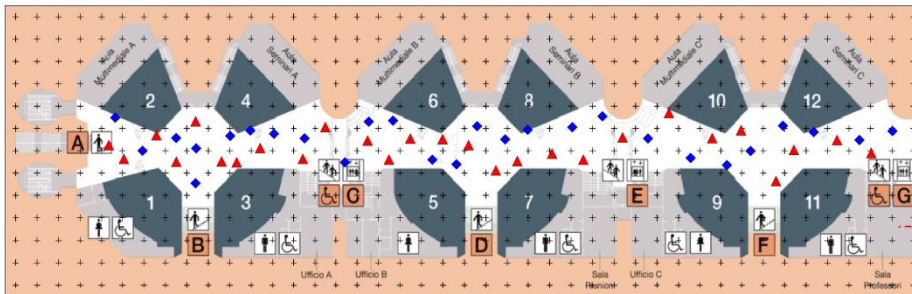


Figura 3-18 AreaP2 Punti di training e punti di test

La media e la deviazione standard dell'errore nel testbed PoloP1 varia al variare di k come mostrato nelle Figure 3-19 e 3-20.

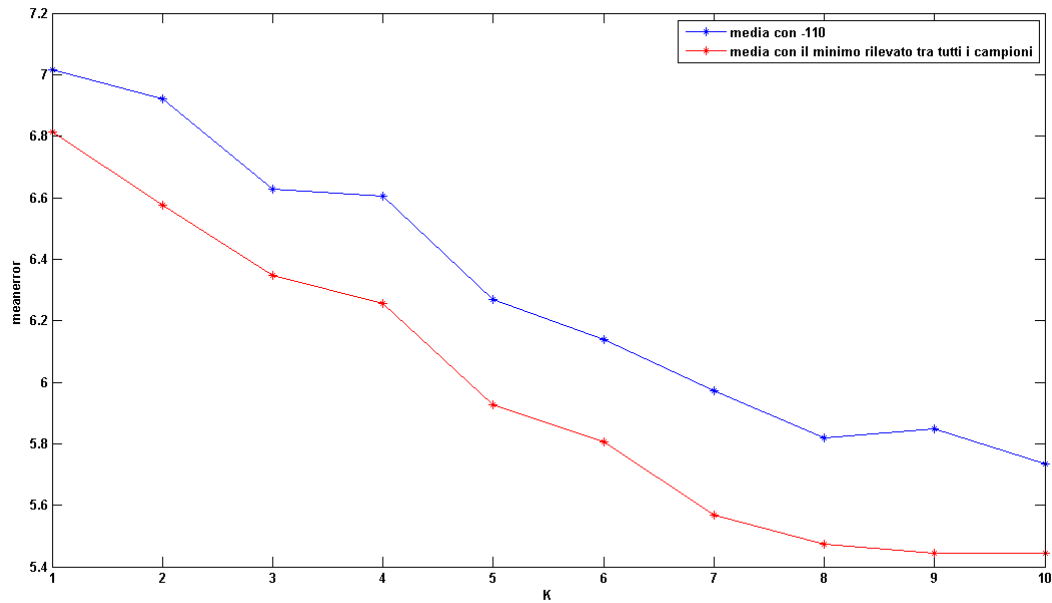


Figura 3-19 PoloP1 Media dell'errore al variare di k

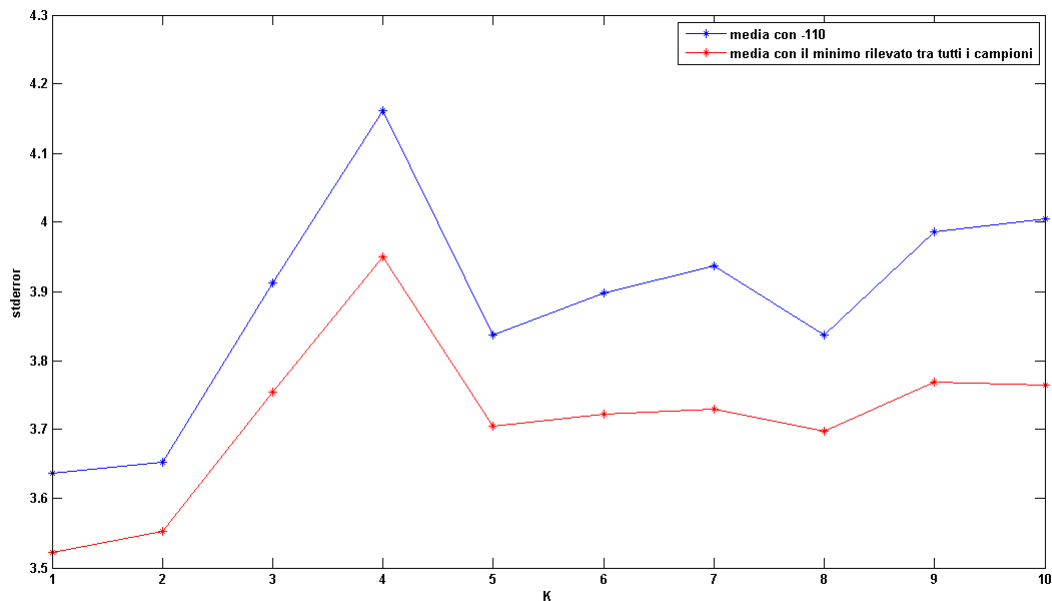


Figura 3-20 PoloP1 Deviazione standard dell'errore al variare di k

Se consideriamo nel testbed AreaP2 la stessa densità di punti di riferimento che abbiamo nel testbed PoloP1, cioè considerando solo i 15 punti di riferimento che si possono vedere nella Figura 3-21, l'algoritmo di apprendimento del rank funziona meglio dell'algoritmo knn così come si può notare dalle Figure 3-22 e 3-23 con i risultati. L'errore medio migliora del 11,4310%.

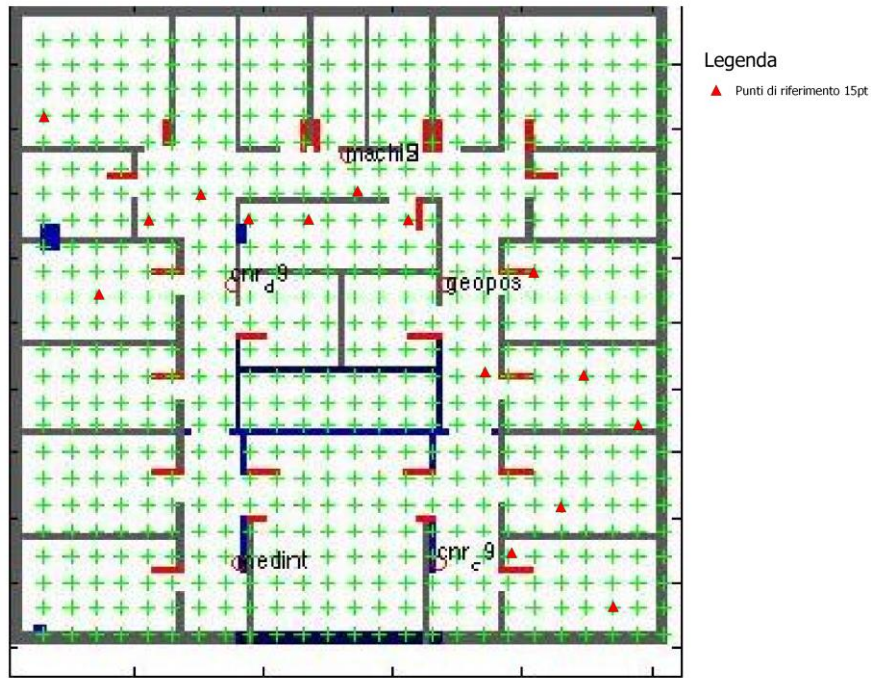


Figura 3-21 AreaP2 15 punti di riferimento

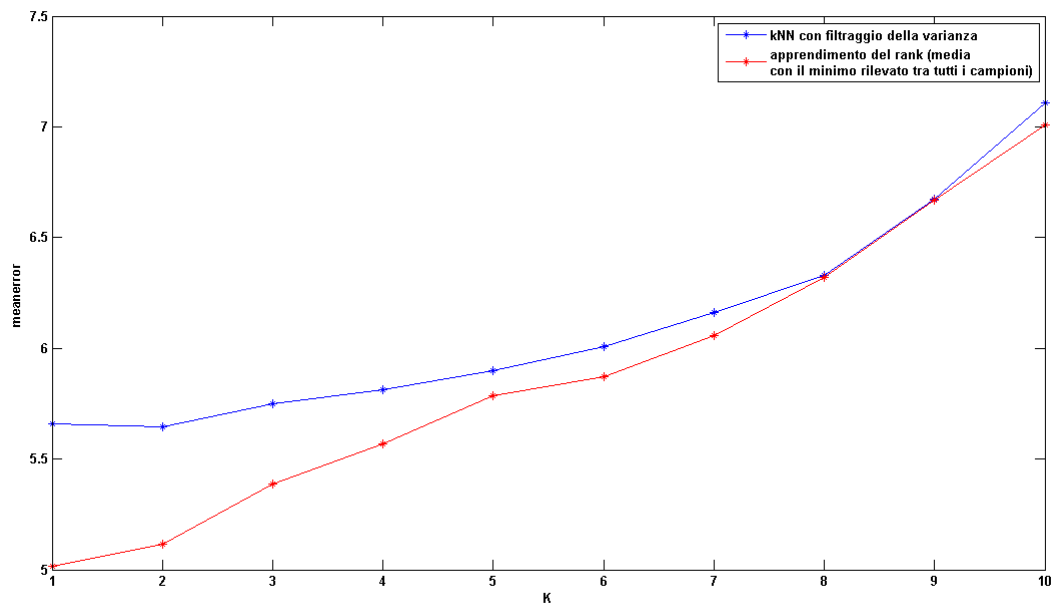


Figura 3-22 AreaP2 Media dell'errore al variare di k (kNN al confronto con apprendimento del rank)

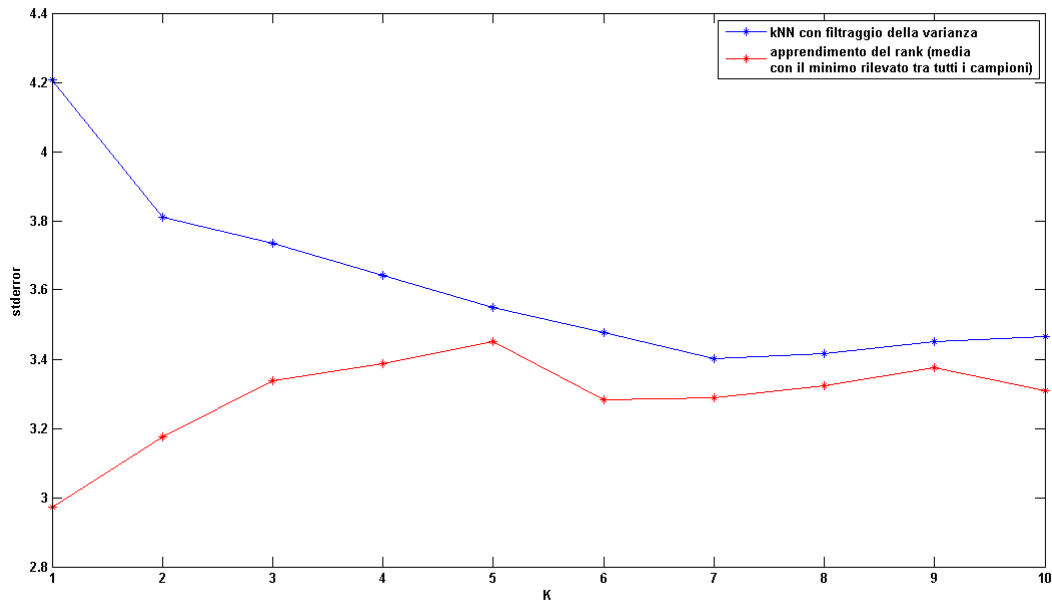


Figura 3-23 AreaP2 Deviazione standard dell'errore al variare di k (kNN al confronto con apprendimento del rank)

Il miglioramento introdotto dall'algorithmo di apprendimento del rank rispetto all'algorithmo kNN con filtraggio della varianza è evidenziato anche nella Figura 3-24 dove sono messe a confronto le distribuzioni cumulative dell'errore dei due algoritmi.

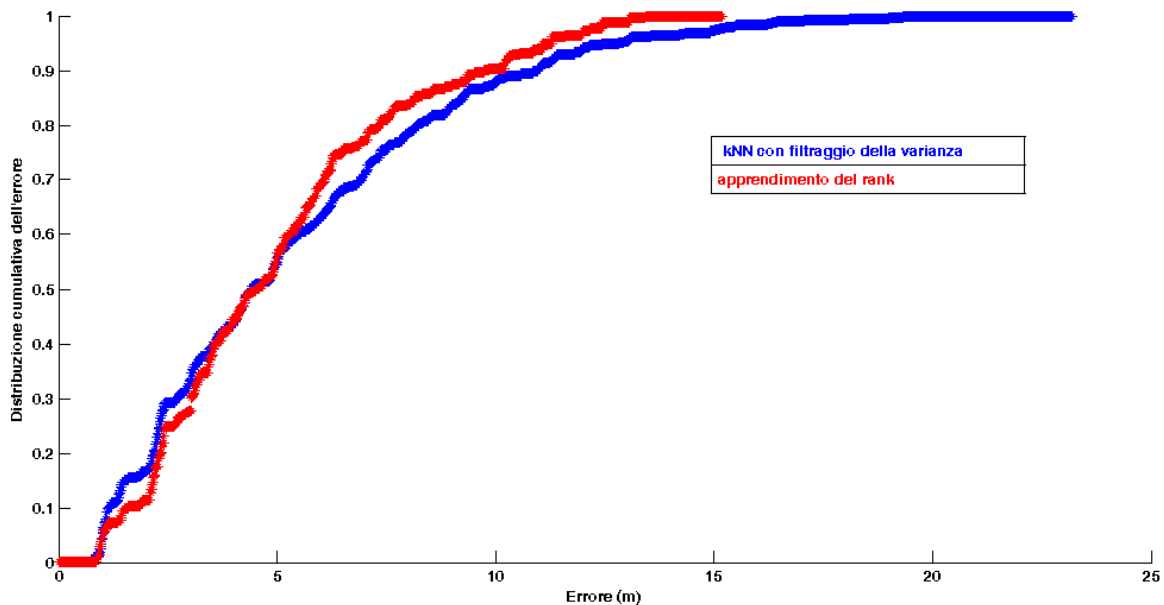


Figura 3-24 AreaP2 Distribuzione cumulativa dell'errore (kNN al confronto con apprendimento del rank)

Valutazioni:

Il test F suggerisce che l'algoritmo kNN è più performante in testbed con densità di fingerprints di riferimento alta. Le prestazioni degli algoritmi di autoapprendimento del rank sono migliori rispetto a quelle dell'algoritmo kNN con filtraggio della varianza in testbed con densità di fingerprints di riferimento bassa. Il test F mostra come la variabilità del segnale campionato induca difformità del pattern di test dai fingerprint dei punti geograficamente vicini, mantenendo l'indice di recall del KNN al di sotto del 50%, anche per alti valori di k.

3.3.6 Test G: incertezza sulla stima della posizione dell'operatore durante l'acquisizione della radiomappa.

Problema: La radio mappa è costruita campionando l'area del sito in analisi in un certo numero di punti identificabili su una griglia che copre uniformemente l'area di interesse.

Per l'identificazione della posizione l'operatore ha a disposizione un riferimento fisico premarcato (es. sul pavimento) o un marker grafico posizionato sulla mappa mostrata sul display dalla applicazione di logging. La posizione del punto di riferimento/test attuale viene comunicata all'applicazione tramite puntamento di un'area sulla mappa mostrata su touch-screen. La scala della mappa mostrata sul display ha valore di zoom selezionabile dall'operatore.

Ci sono due componenti dell'errore: la prima è l'incertezza nel posizionamento sul marker fisico o nella stima della sua posizione in riferimento agli elementi strutturali presenti nella mappa (muri, porte etc); la seconda, è costituita dall'errore di indicazione sul touch-screen del punto di riferimento supposto.

Nel caso del sito AreaP2, di dimensione 25x25 mq circa, altamente strutturato, è stata utilizzata una griglia di campionamento uniforme con maglia di 1 metro e sono stati apposti markers sul pavimento sui nodi della griglia a distanza 1 o di 2 metri in funzione della accessibilità dei locali. Nel caso del sito PoloP1, di dimensione 170 x52mq circa, confinato, è stata utilizzata una griglia di campionamento uniforme con maglia di 4 metri.

In entrambi i casi la griglia era mostrata sul display.

Ipotesi: l'errore di autoposizionamento, usando i marker grafici sulla mappa, corrisponde alla capacità dell'operatore di segmentare percettivamente in un numero limitato di parti uguali lo spazio fra riferimenti strutturali dell'ambiente. L'errore migliora con l'esperienza, peggiora all'aumentare del numero di parti in cui frazionare lo spazio. Il numero di parti ottimale è stimato fra 4 a 6.

Metodo di stima 1: empirico basato sulla esperienza di mappatura riportata dagli operatori

Metodo di stima 2: sperimentale indiretto: confronto sul miglioramento nell'accuratezza media ottenuto utilizzando come set di test il set di riferimento già utilizzato per la costruzione delle radio mappe. Ponendo nell'algorithmo KNN $K=1$, l'errore di autovalutazione della posizione è nullo se il punto è correttamente individuato come primo vicino di se stesso.

Evidenze:

	test	riferimento		
PoloP1				
AreaP2				

La tabella mostra l'errore medio di accuratezza per

Valutazione empirica: L'errore di posizionamento dei marker fisici è pari ad alcuni centimetri e quello sul marker è dell'ordine del decimetro.

Autovalutazione della capacità di stima: valore massimo fra 1/6 della distanza fra gli elementi strutturali presenti, e metà del passo della griglia di riferimento sulla mappa.

Valutazione: L'errore di posizionamento dei marker fisici è pari ad alcuni centimetri e quello sul marker è dell'ordine del decimetro, quindi entrambi sono trascurabili rispetto all'errore totale dell'ordine dei metri.

L'errore di autoposizionamento tramite marker sul display è valutabile in 1m per il sito AreaP2 (con griglia da 1 o 2 m e distanza media dei muri 4/6 m) ed in 2m per il sito PoloP1 (griglia 4 m, distanza con media muri 8/12 m).

L'errore di indicazione sul touchscreen è da considerarsi trascurabile se la mappa è mostrata sul display con risoluzione intorno a pochi cm per pixel (zoom elevato).

3.3.7 Controllo network-side della variabilità lenta del campo a radiofrequenza

Problema: La ricerca di un sistema che possa percepire delle variazioni delle configurazioni di campo degli AP risulta essere giustificata dal fatto, che nella localizzazione con il metodo del fingerprinting una nuova operazione di calibrazione è necessaria ogni qual volta vengono inseriti AP, in presenza di cambi topologici come presenza di muri, armadi, variazioni della densità degli ambienti, o in generale variazioni delle condizioni elettromagnetiche generate da effetti di shadowing, tempovarianza del canale, ecc...

Modelli di calibrazione del campo differenti dovrebbero essere utilizzati a secondo le condizioni di tempovarianza del canale radio, in particolare dovrebbero essere resi adattativi (usando dei checkpoint fissi che effettuino delle misure) o statici, in funzione delle ore del giorno considerando ad esempio locali altamente frequentati, locali semi vuoti.

Tecniche predittive del segnale elettromagnetico che facciano uso di modelli di propagazione applicati all'ambiente specifico, accompagnati da informazioni quali la presenza di ostacoli in muratura, porte o di altri oggetti ingombranti combinate ad informazioni relative a misure prese a campione in posizioni selezionate può condurre ad un miglioramento delle prestazioni della localizzazione.

Nel capitolo si discutono esperimenti realizzati per verificare se sia possibile rilevare una variabilità temporale lenta del campo a radiofrequenza generato da AP 802.11 in determinate zone della radio mappa attraverso misure incrociate di intensità di RSSI effettuate dagli stessi AP (modalità di acquisizione network side).

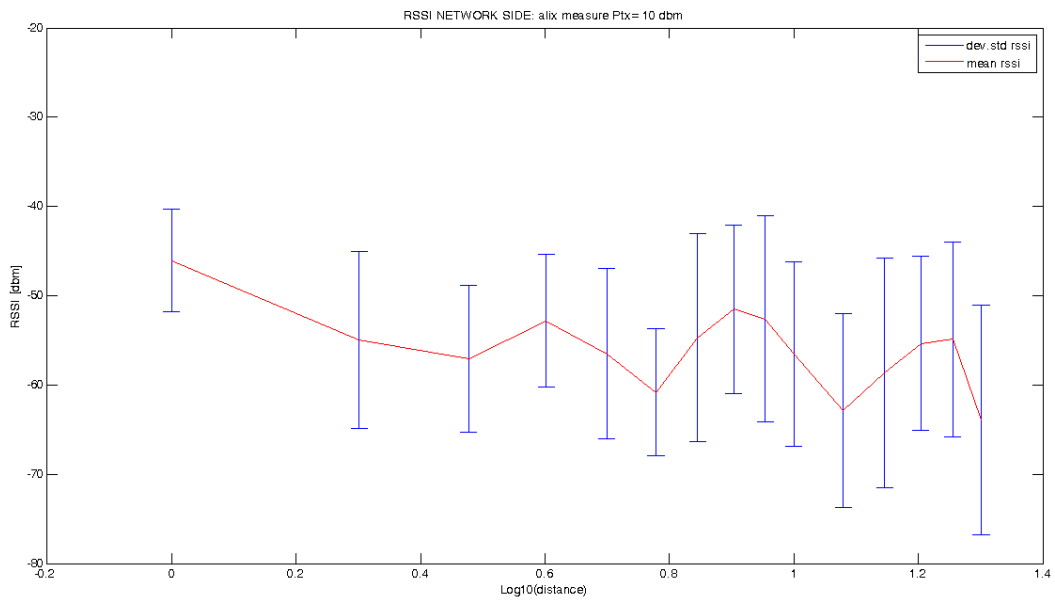
Il fine di tali misure è osservare la possibilità di selezionare fra due radio mappe precalcolate in due differenti configurazioni ambientali, e scegliere quella più adeguata a rappresentare la configurazione di campo attuale che tenga conto appunto delle variazioni.

Ipotesi 1: E' possibile rilevare il valore dell'intensità di campo di un AP emettitore, da un device mobile ricevitore (client-side) come segnale stabile con intensità media dipendente dalla posizione nel sito con la caratteristica di avere una bassa varianza

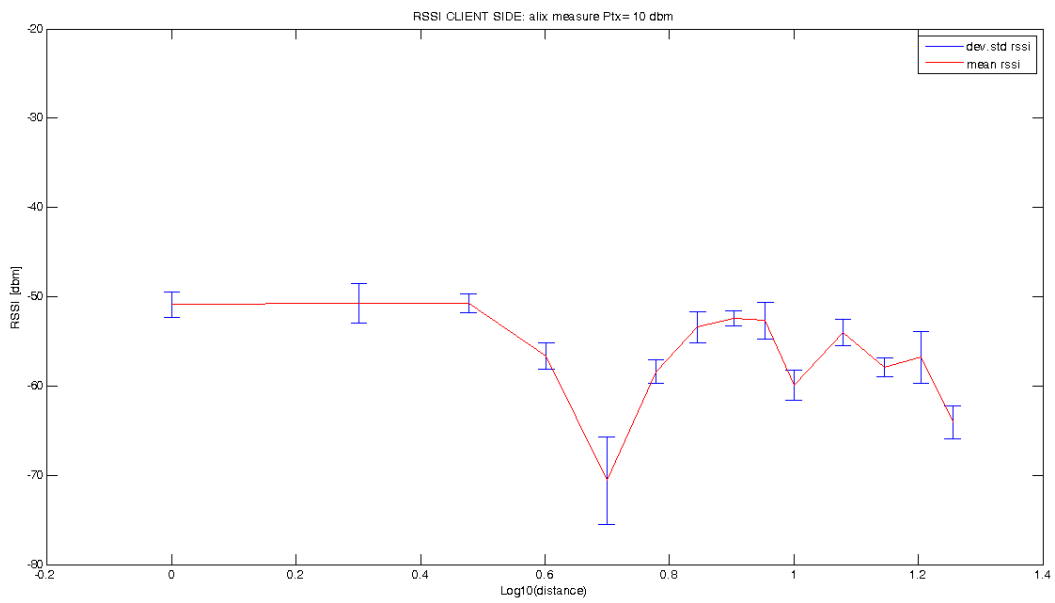
Ipotesi 2: Il valore dell'intensità di campo di un AP emettitore, rilevata da un device mobile ricevitore (client-side) e da un secondo AP ricevitore (network-side) è correlata e proporzionale (quindi è possibile effettuare una calibrazione)

Ipotesi 3: I valori dell'intensità di campo di un AP emettitore, rilevati da un insieme di AP device mobile (client-side) e da un secondo AP (network-side) è correlata e proporzionale (quindi è possibile effettuare una equalizzazione dei rispettivi valori)

Metodo di stima 1: Comparazione di valori medi

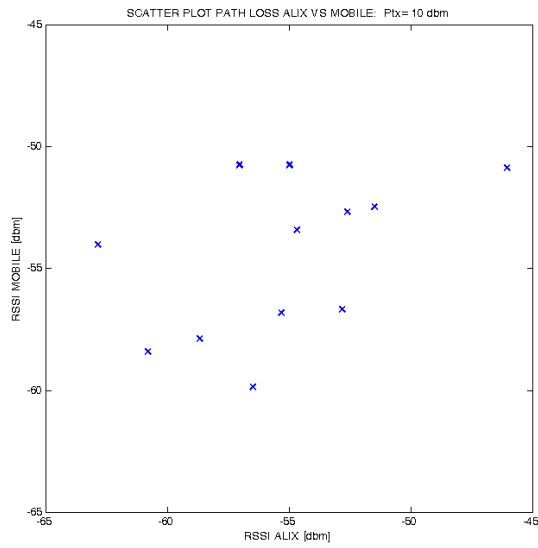


3-15 Misure di pah loss network side



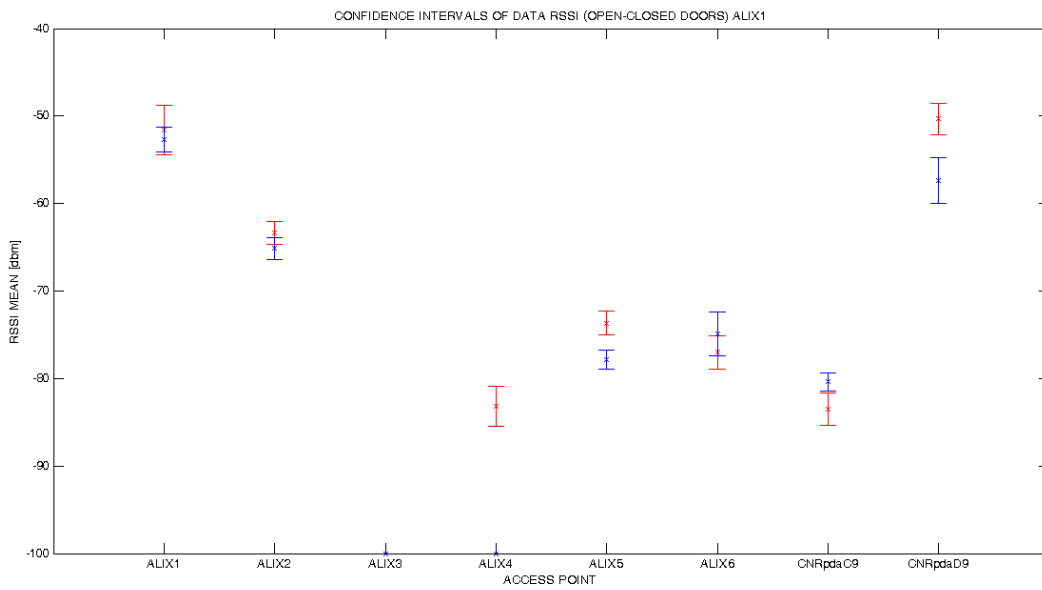
3-16 Misure di pah loss client side

Metodo di stima 2: Comparazione di valori medi

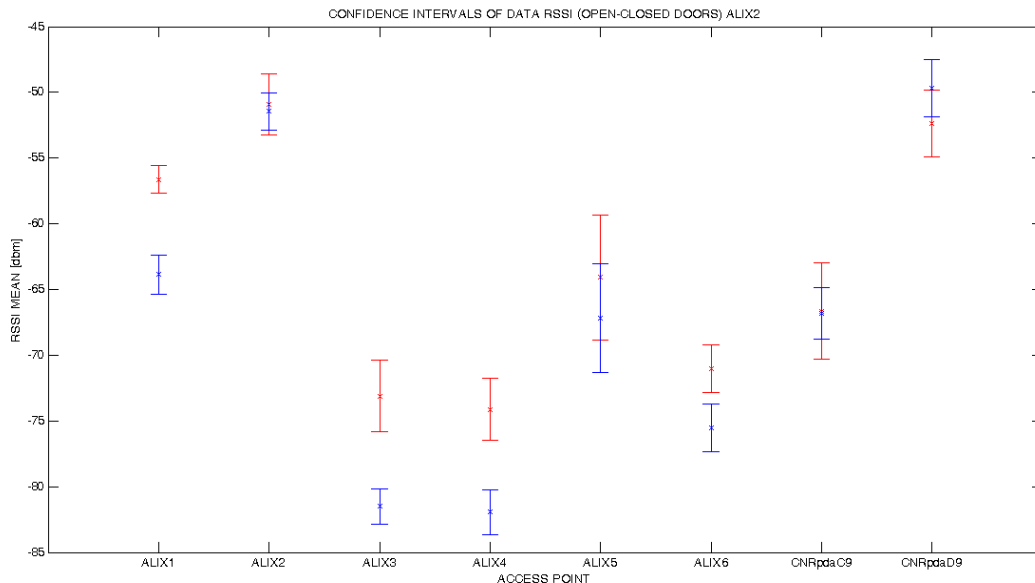


3-17 Scatter plot del path loss network vs client side

Metodo di stima 2: Comparazione di valori medi



3-18 Variabilità dell'rss misurato a porte aperte e chiuse dall'alix 1



3-19 Variabilità dell'rssI misurato a porte aperte e chiuse dall'alix 2

Evidenze:

Le letture del segnale network-side effettuate in locale aperto confinato hanno un range dinamico che spazia tra -40 a -75 dBm (con segnale alla sorgente di -10 dBm) e livelli di deviazione standard variabili e in alcuni casi pari a 15 dBm). In locale aperto confinato si evince scarsa evidenza di path-loss entro i 15 metri.

Lo stesso segnale è rilevato client-side dal terminale mobile con dinamica inferiore (da -50 a -70 dBm e con deviazione standard più contenuta (3-6 dBm).

In locale aperto confinato si evince parimenti scarsa evidenza di path-loss entro i 15 metri.

Il diagramma di scatter mostrato precedentemente come il valore dell'intensità di campo di un AP emettitore, rilevata da un device mobile ricevitore (client-side) e da un secondo AP ricevitore (network-side) denota una scarsa correlazione tra le misure.

In locale confinato, i livelli medi mostrano una variabilità minore, sia nelle condizione di porte aperte che chiuse.

I valori del segnale in numerosi punti mostrano significative differenze nelle due condizioni (porte aperte o chiuse)

Valutazioni:

Il segnale network-side presenta elevata variabilità nel sito AREAPT. Tale variabilità si riduce ai valori attesi per misure di device mobile nelle misure nel sito strutturato AREAPT.

I valori dell'intensità di campo rilevati dall'Ap e dal mobile sono scarsamente correlati e un'eventuale equalizzazione sarebbe imprecisa.

Il pattern dei valori di intensità rilevati da ogni AP risulta in molti casi differente per le due condizioni Porte Aperte/Porte Chiuse e suggerisce la possibilità di estendere lo studio specializzando ogni AP al controllo di variazioni significative in una differente zona del sito. L'esperienza empirica acquisita durante gli esperimenti di test tende a suggerire che un ampio margine di variabilità nel sito AREAPT sia da attribuirsi al mancato controllo della direzione di orientamento relativo delle antenne direzionali dei due AP ed ai livelli di potenza in trasmissione degli AP stessi.

3.4 Setting:

L'approccio di misure che consente l'acquisizione dei valori di potenza dei segnali trasmessi tra i vari AP è realizzato in modalità network side.

In particolare ogni AP (alix2D2) è dotato di due schede mini pci WLAN *bcm4318kfbg*, di cui una (wlan0) è fatta lavorare in modalità master mediante l'utilizzo del daemon *hostapd* sotto ambiente linux; nello stesso istante l'altra interfaccia (wlan1) effettua la misura dei valori di potenza (RSSI) dei beacon ricevuti dagli altri AP.

Nello specifico è stato sviluppato un applicativo java compatibile con jdk v1.7, che effettua una chiamata di sistema "iwlist" in grado di ritornare informazioni sull'interfaccia di rete wireless, come le letture di RSSI relative ad uno scanning passivo effettuato sulla seconda interfaccia.

L'operazione di scanning così come previsto dallo standard è utile per trovare un dispositivo a cui associarsi. Nel caso passivo il nodo wireless si mette in ascolto per catturare il beacon contenente il SSID della rete e fare il join ad essa.

Le misure di potenza sono condotte sui segnali di beacon, che è uno dei frame di management previsti dallo standard IEEE 802.11, frame il cui compito è di contenere tutte le informazioni sulla rete e di sincronizzare le stazioni all'interno della wlan.

Lo stesso applicativo è in grado di interagire con un database al fine di effettuare lo storage dei dati in real time.

Il tempo minimo tra una scansione e la successiva è di circa 4 sec ed è il tempo che impiega l'interfaccia per fare un hopping tra i canali a 2.4 Ghz.

Il valore medio e la deviazione standard di RSSI percepita da ogni alix è fatta risalire ogni 10 secondi e in particolare tali valori sono generati a partire da una media mobile con una finestra di 5 sample.

Il fatto di utilizzare il valore medio piuttosto che i valori istantanei è giustificato dal fatto di dover percepire le variazioni lente del segnale in seguito a variazioni topologico/ambientali dell'ambiente di test e che sono quindi apprezzabili analizzando un certo intervallo di tempo.

3.5 Esperimenti effettuati:

- a) **Setting per comparazione letture client-side e network-side (descrizione dell'ambiente AreaPT e delle condizioni di misura)**

Il setting per le misure è volto ad evidenziare il cambiamento della configurazione del campo generato dai vari AP in funzione di condizioni variabili; per questo motivo misure di rssi sono condotte a porte aperte e chiuse dai vari AP.

Contestualmente allo stesso setting sono realizzate misure client side realizzate nelle medesime condizioni in punti specifici della radiomappa.

L'ambiente di test è il secondo piano dell'ICAR CNR, un ambiente indoor molto partizionato dalla forma quadrata con un area di circa 676 m². Le porte divisorie dei corridoi e delle stanze hanno uno spessore di circa 5cm e presentano un'intelaiatura metallica così come osservabile in una delle figure seguenti.

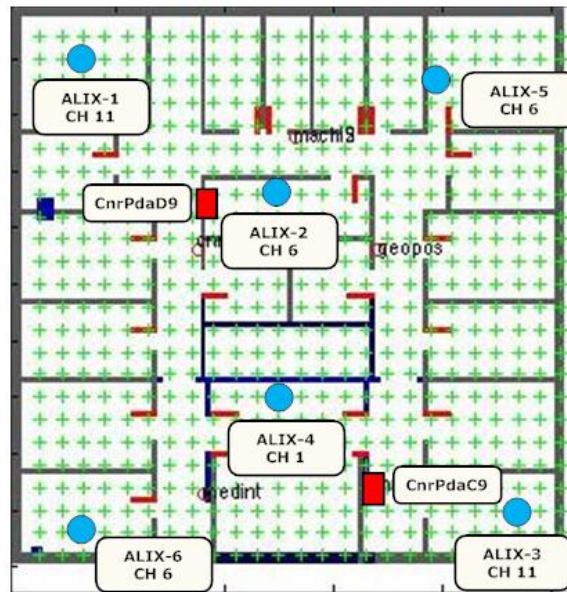
Il planning degli AP valido nella configurazione delle misure network e client side è stato effettuato in modo da ottimizzare la copertura del piano in termini di copertura del segnale, cercando di evitare l'interferenza co-channel; per questo si è privilegiata una disposizione perimetrale degli stessi così come osservabile in figura.

Nell'area sono presente altri AP di interesse quali il CnrPdaD9 e C9 (smc2555w-ag2), in particolare sono state disposte altre 6 alix-2d2 capaci di operare in modalità network side così come esposto precedentemente.

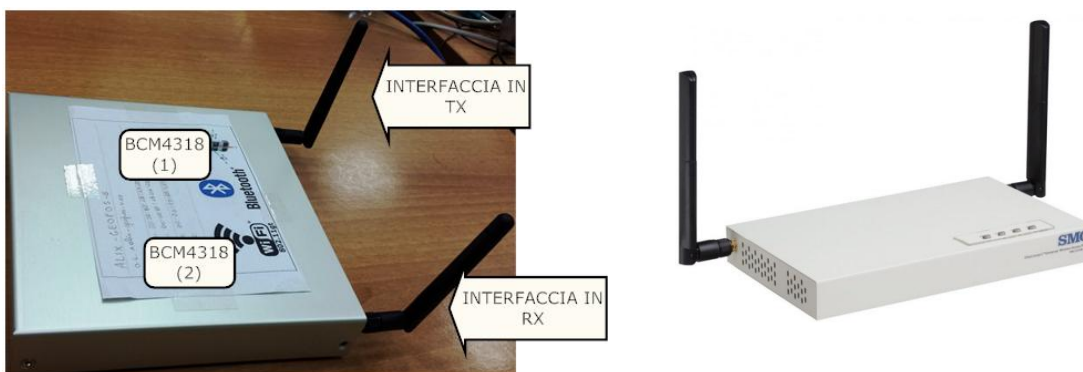
Le misure client side sono condotte mappando l'edificio con una griglia ad 1 m, e rilevando con smartphone samsung galaxy s 3-4 (s.o. android 4.4.2), GT-I907P (android 4.0) in ogni punto il valore di rssi percepito dagli AP di interesse.

Le misure network successive a quelle client side sono state effettuate a porte aperte e chiuse per una durata temporale di 10 minuti senza alcuna variazione nella disposizione degli ap.

(Setting e relativi AP utilizzati)



Posizionamento degli AP nella campagna di misure network e client side



AP utilizzati nella campagna di misure



Tipologia di porte (corridoio e stanze)

a) **Analisi della corrispondenza delle misure network e client side attraverso studi sul path-loss**

Per poter opportunamente calibrare la mappa una volta rilevate delle variazioni di campo dagli ap, è necessario riferire le variazioni di RSSI intercettate dagli AP, a quelle osservate dallo smartphone. Per questo motivo è necessario osservare la dinamica di acquisizione dell’RSSI al variare della distanza, sia per l’alix, che per lo smartphone di test.

Questa operazione prende il nome di equalizzazione ed è necessaria in quanto schede diverse gestiscono la lettura dell’RSSI in modo differente a secondo il chipset e i driver.

Bisogna osservare che nella modalità network side la rete effettua le misure che sono riferite allo stesso hardware, nella modalità client side le misure sono condotte con dispositivi di brand differenti e quindi con configurazioni hardware e driver differenti; per tale motivo l’operazione di creazione della radiomappa dovrebbe essere eseguita con lo stesso smartphone.

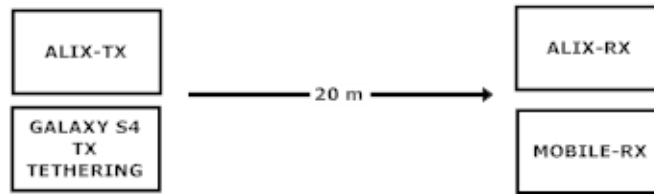
Al fine di correlare le misure le misure realizzate network side con lo smartphone e quelle realizzate client side sono state effettuate delle misure di decadimento del segnale.

Il setup prevede un AP alix in trasmissione, e due unità in ricezione per la misurazione del valore di potenza del segnale, quali uno smartphone GT-I907P (android 4.0) e un alix.

Le Alix sono equipaggiate di antenne omni 3dbi RP-SMA per funzionamento a 2.4 Ghz.

Le misure sono state condotte nel corridoio esterno adiacente alla zona relax dell’icar cnr, analizzando le distanze da 0-20 metri, con step di 1 metro tra 1-10m e step di 2m tra 12-20 metri.

In particolare le misure sono state condotte al variare della potenza in trasmissione dell’alix settata rispettivamente a 20, 10 e 5 dbm.



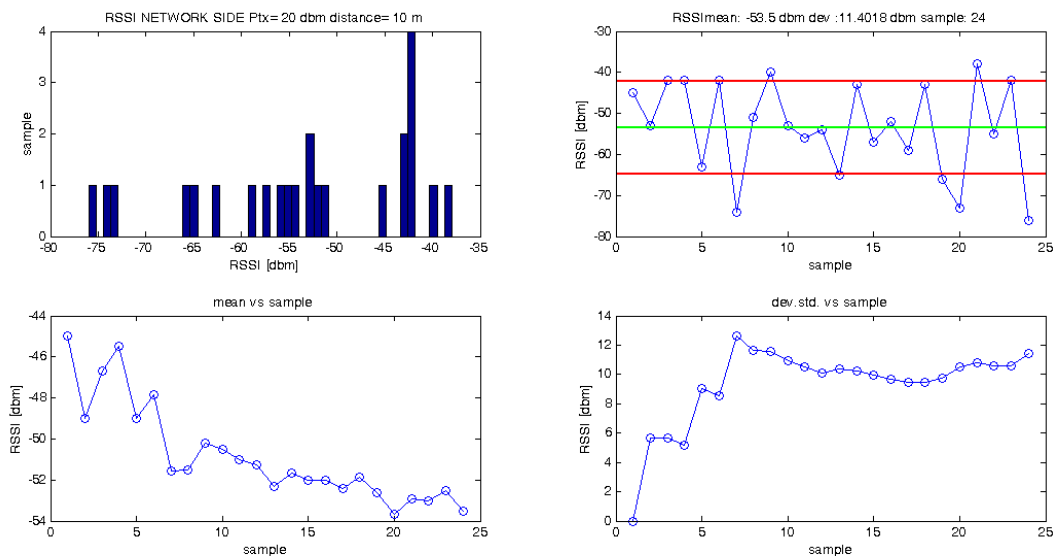
Setup misure path loss

In tutti i casi in cui l'alix effettua la misura considerando diversi livelli di potenza in trasmissione non si osserva un comportamento perfettamente monotono decrescente, ma vi sono continue oscillazioni del segnale imputabili a condizioni di riflessione, multipath generato dai muri adiacenti alla zona di misura.

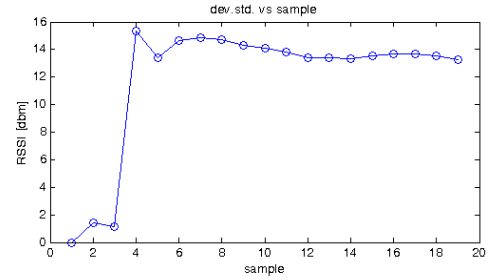
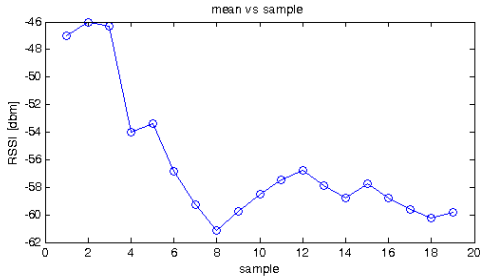
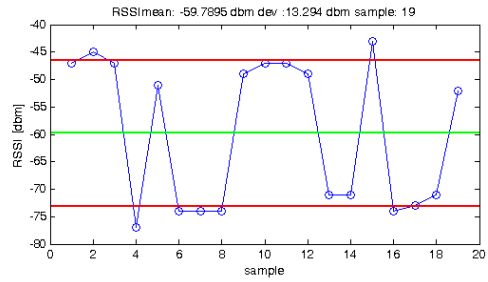
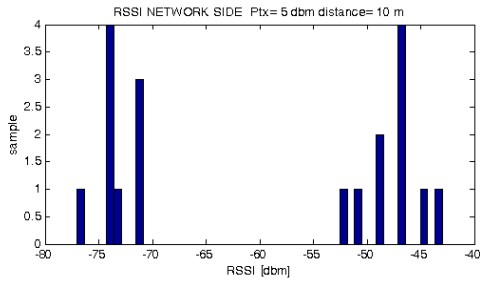
In particolare risalta l'elevato valore di varianza del segnale misurato a diverse distanze nella condizione in cui venga misurato il segnale proveniente dall'alix, varianza che risulta notevolmente ridotta quando viene misurato il segnale di uno smartphone in tethering.

Tale aspetto potrebbe essere giustificato dall'effetto prodotto dalle antenne in dotazione nelle alix, e dalla relativa prossimità del device in ricezione.

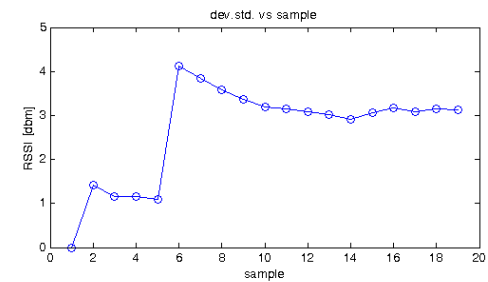
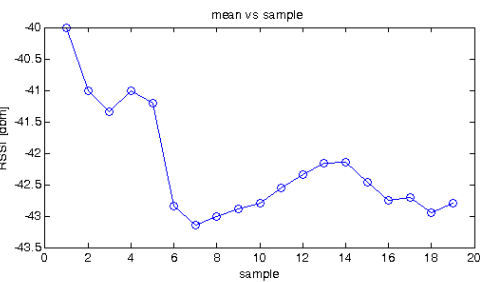
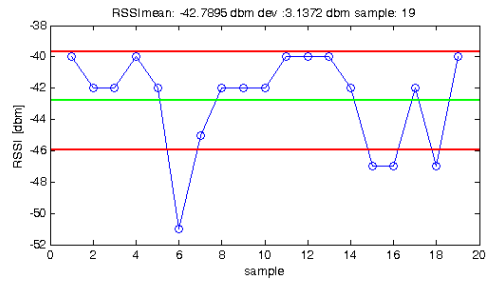
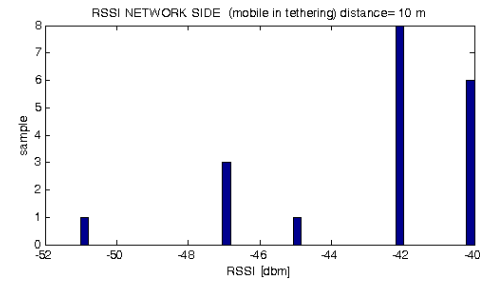
In molti casi si osserva una distribuzione bimodale della potenza misurata dall'alix, che in alcuni punti risulta avere una grande range dinamico di valori di RSSI.



misura a 10m effettuata dall'alix, quando a trasmettere e l'alix con Ptx=20dbm

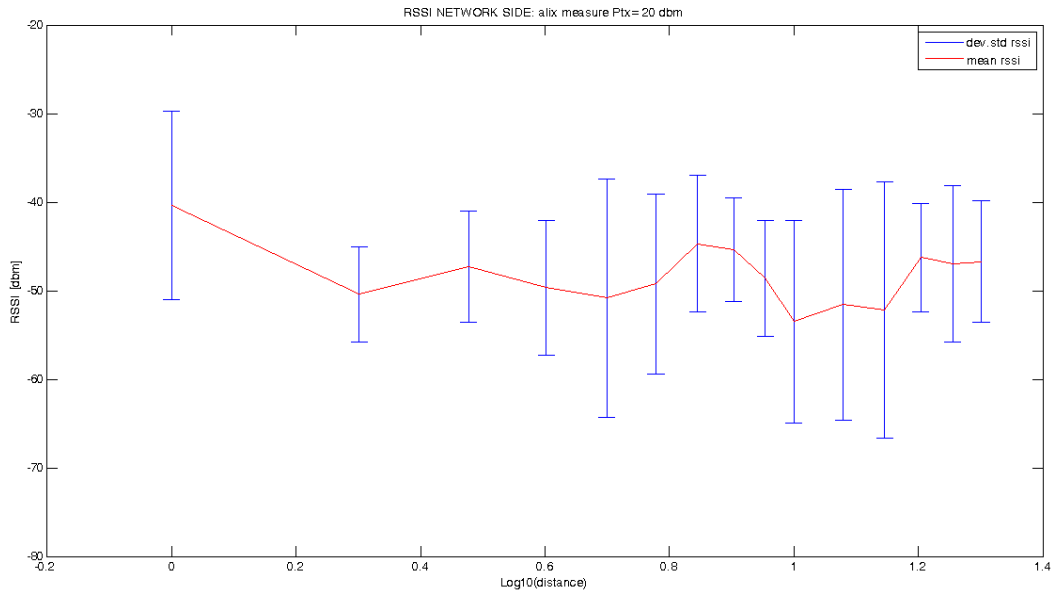


misura a 10m effettuata dall'alix, quando a trasmettere e l'alix con Ptx=5dbm

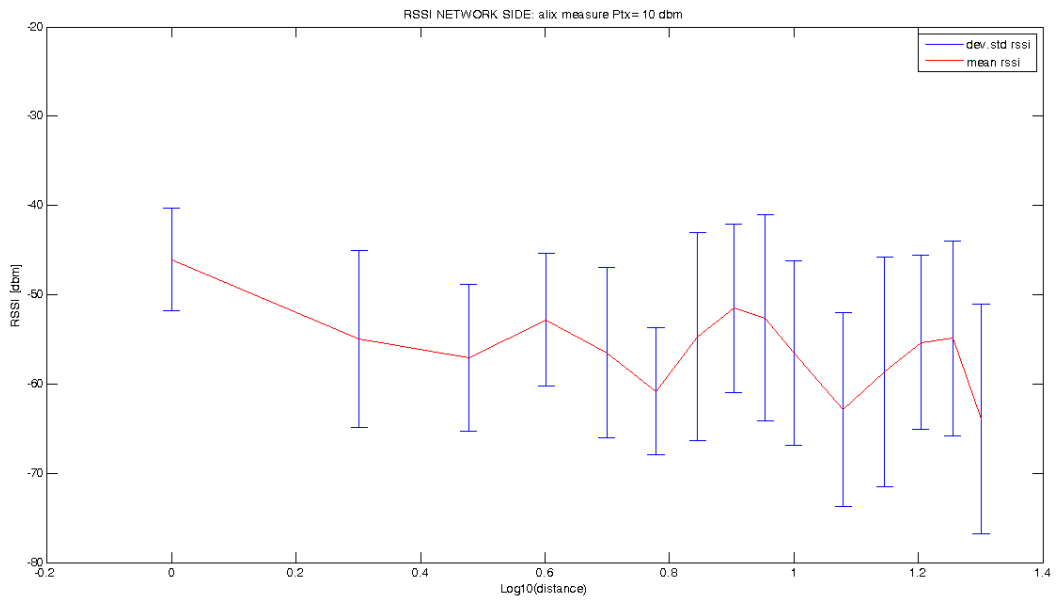


misura a 10m effettuata dall'alix, quando a trasmettere è il mobile in tethering

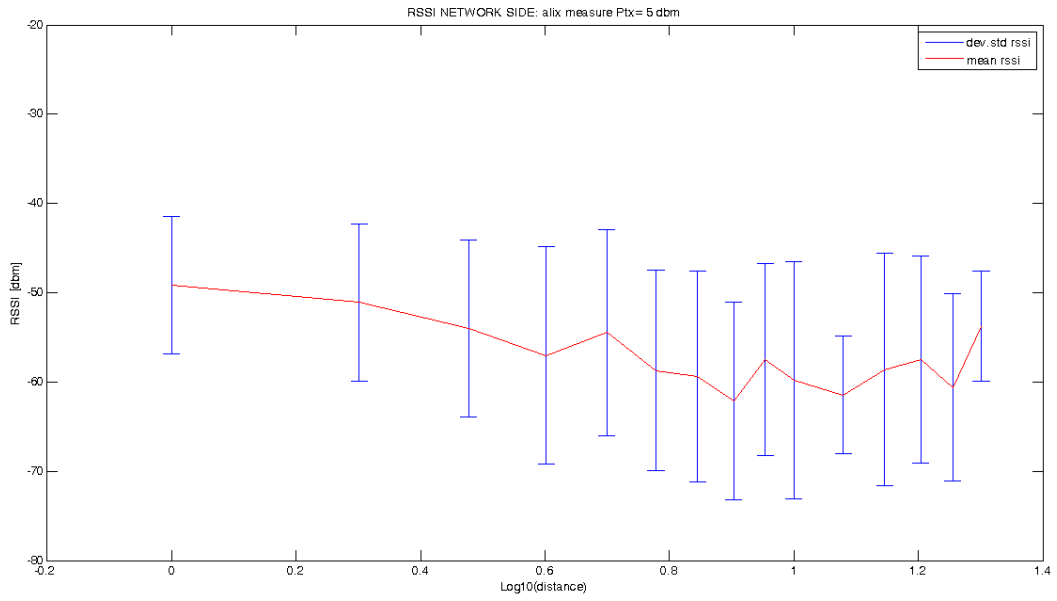
Dai grafici precedenti osservando gli stessi istanti in cui l'alix misura il mobile in tethering e l'alix in trasmissione, la percezione dei valori di potenza nonché le dinamiche di valori sono notevolmente differenti, sintomo che oltre a questioni driver, hardware dei chipset delle schede wlan una variazione così significativa dei livelli di potenza è dovuta anche all'antenna e al suo orientamento.



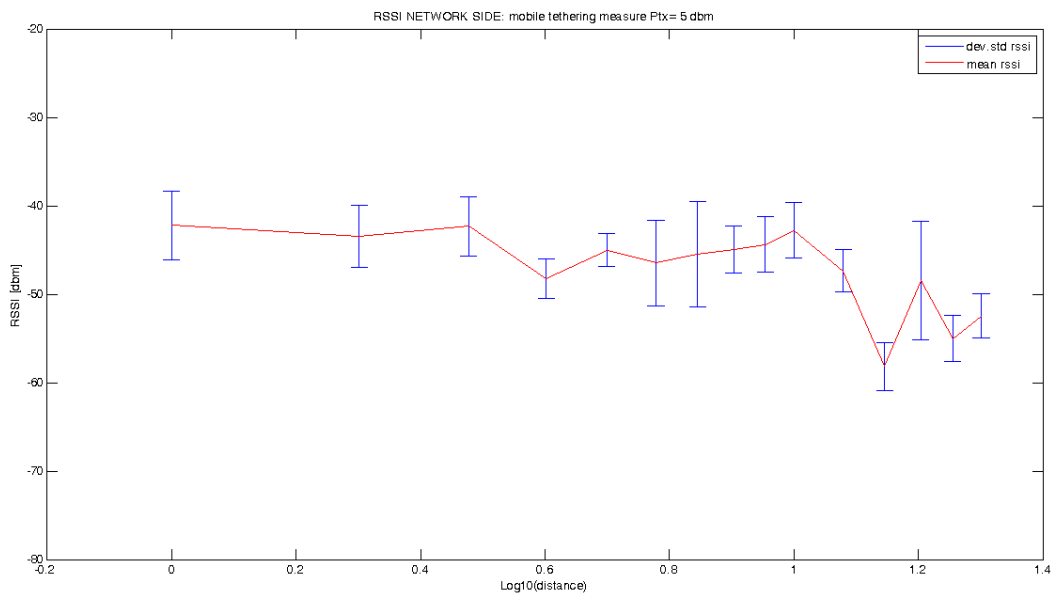
misura di path loss effettuata dall'alix a 20 dbm



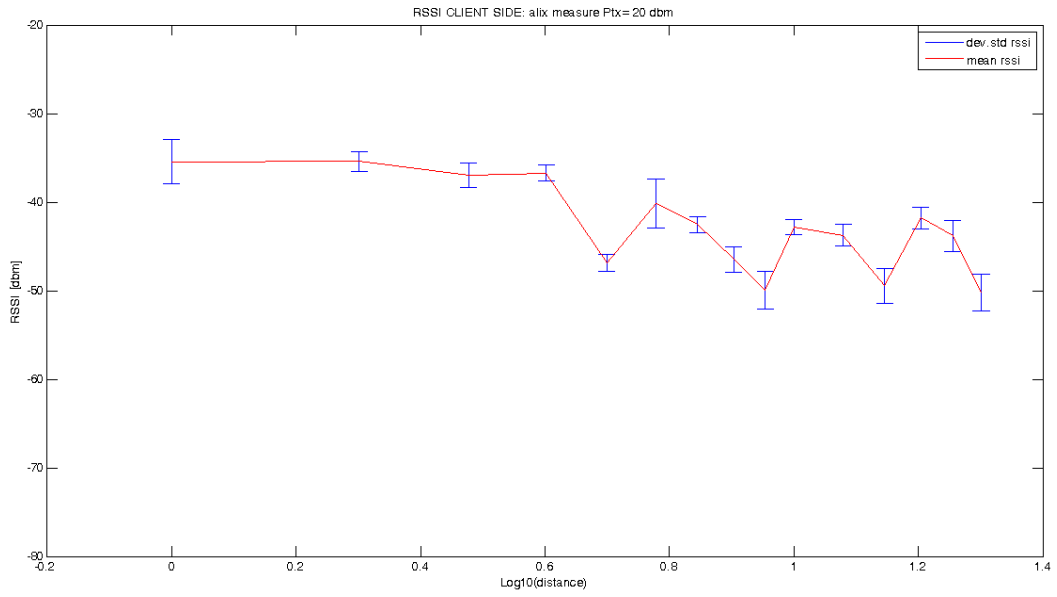
misura di path loss effettuata dall'alix a 10 dbm



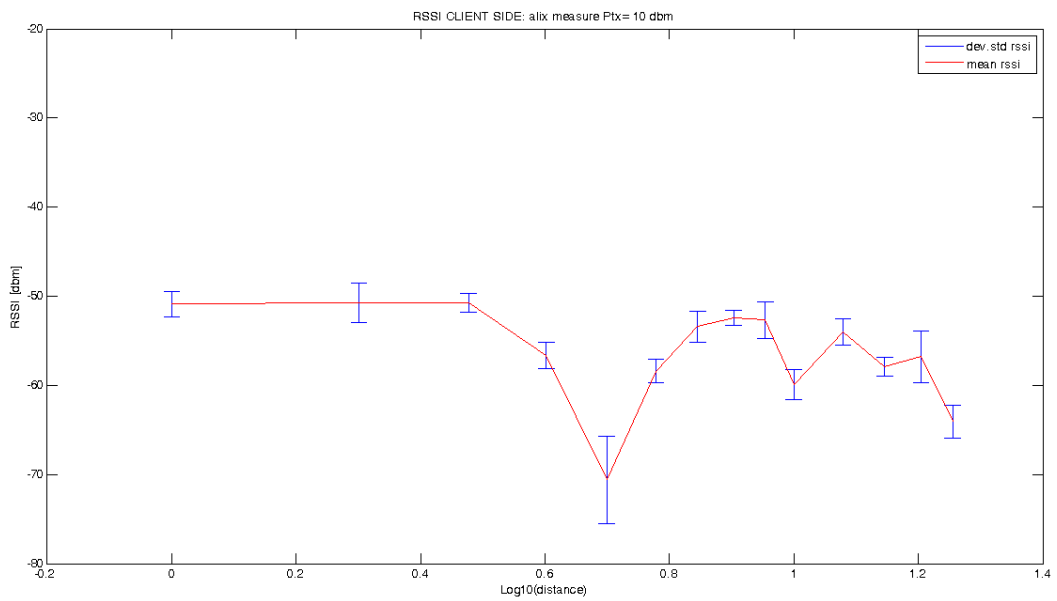
misura di path loss effettuata dall'alix a 5 dbm



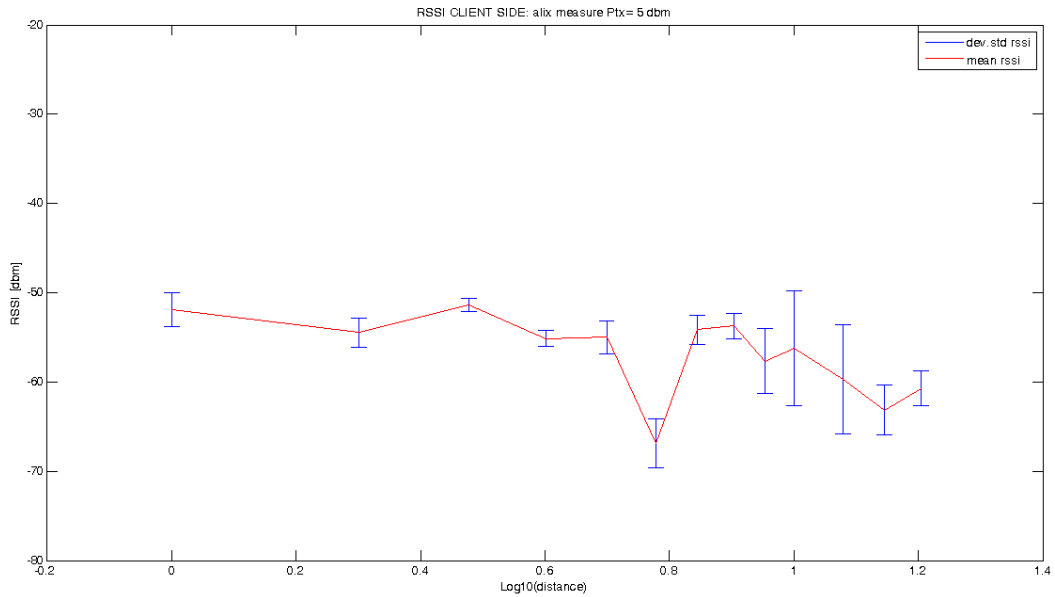
misura di path loss effettuata dall'alix con mobile in tethering



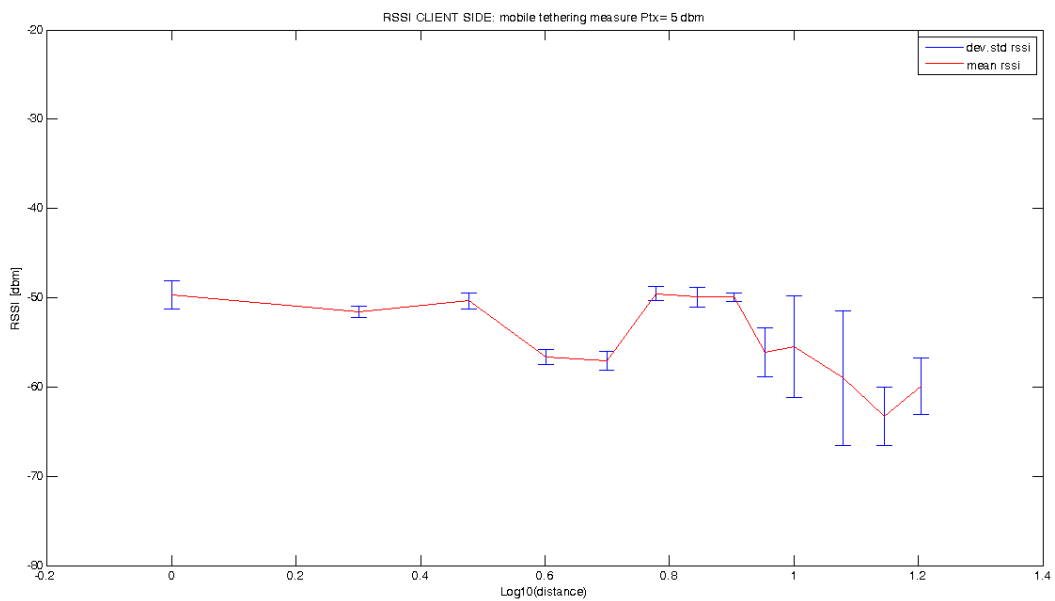
misura di path loss effettuata dal mobile a 20 dbm



misura di path loss effettuata dal mobile a 10 dbm



misura di path loss effettuata dal mobile a 5 dbm

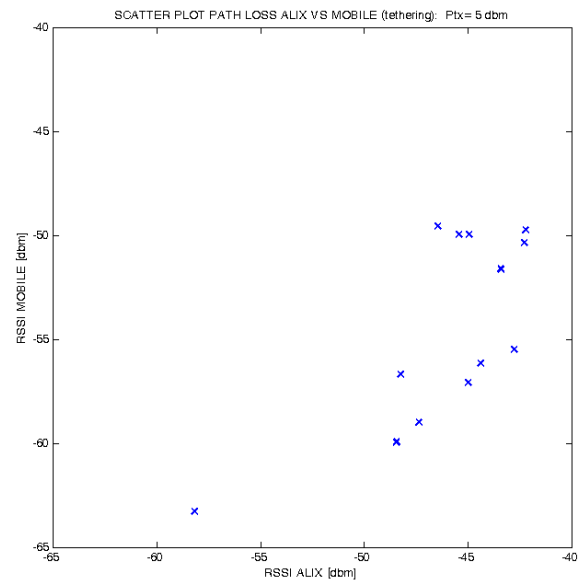
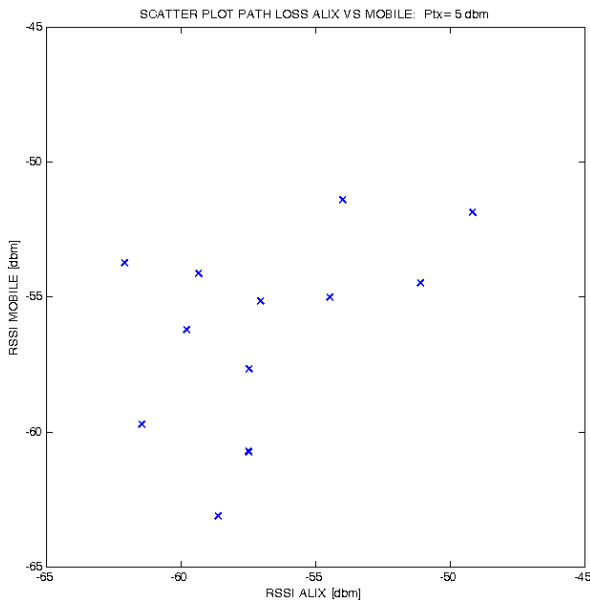
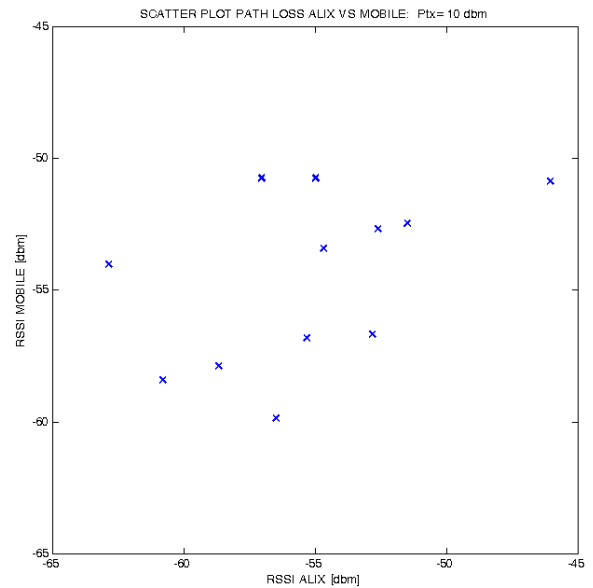
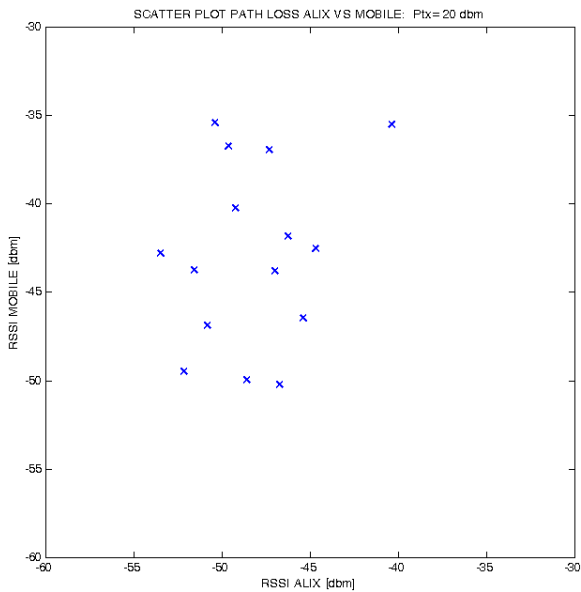


misura di path loss effettuata dal mobile con mobile in trasmissione in tethering

MISURE MOBILE vs ALIX

Nel seguito sono riportati gli scatter plot relativamente alle misure condotte con il mobile e con l'alix.

In generale non si osserva un grado di correlazione tra le due misure se non in pochi punti, questo è dovuto ai fattori esposti precedentemente.



Scatter plot delle misure eseguite dall'alix e dal mobile con diversi settaggi della potenza in trasmissione

b) variazione dei valori di RSSI network side al variare delle condizioni ambientali (porte aperte/chiusure)

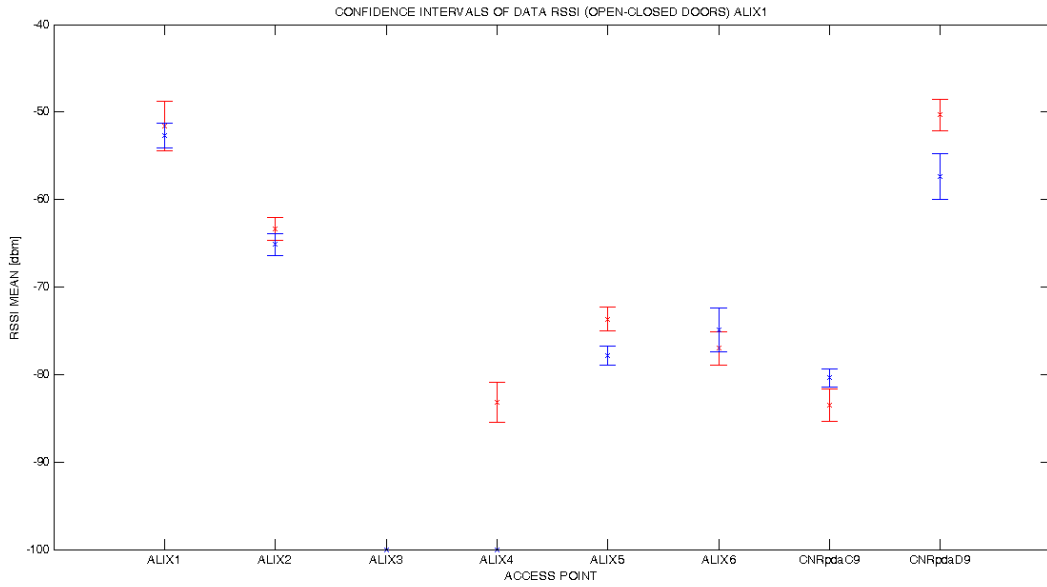
In questa parte vengono analizzati i risultati inerenti alle misure effettuate nell'ambiente di test nelle condizioni di apertura e chiusura delle porte.

Le condizioni di porta aperta-chiusa simuleranno un cambiamento delle condizioni di propagazione nell'ambiente, variazioni che potrebbero avvenire normalmente in seguito a variazioni strutturali, topologiche dell'ambiente di test.

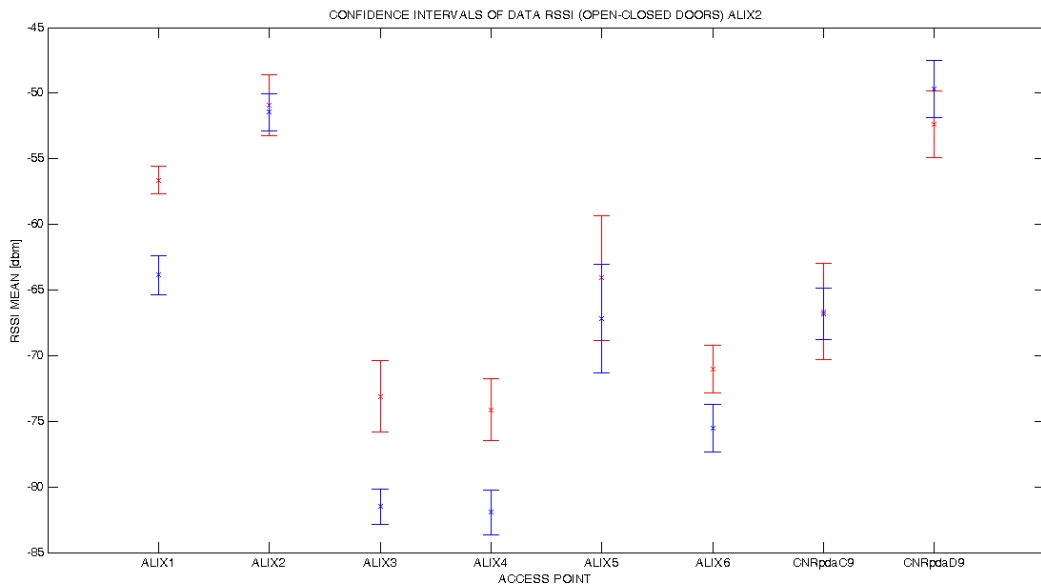
Il set di misure sono realizzate a distanza temporale di un'ora circa e le rilevazioni sono condotte per la durata di 10 minuti circa.

MISURE NETWORK SIDE

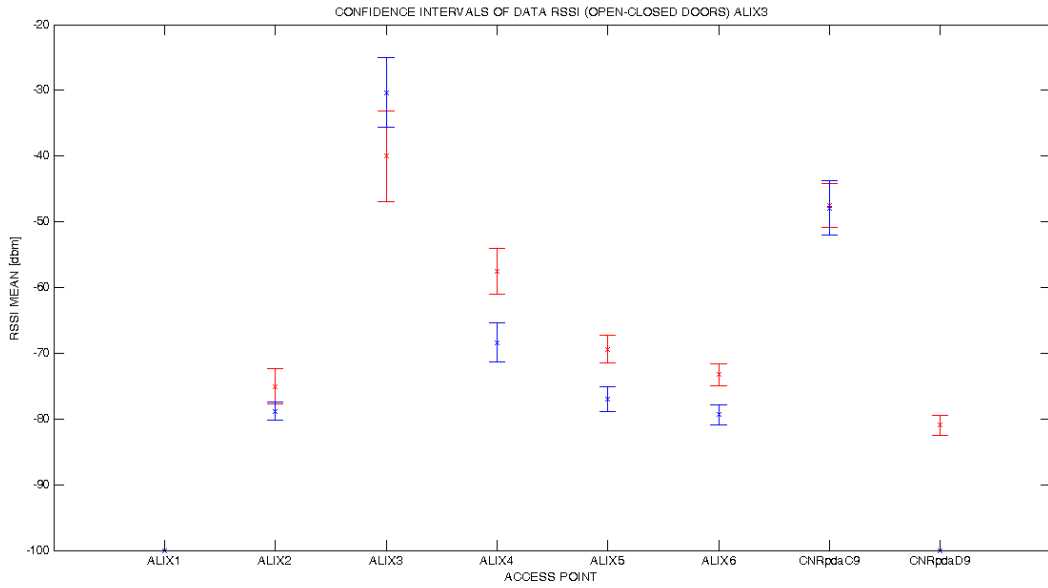
LEGGENDA



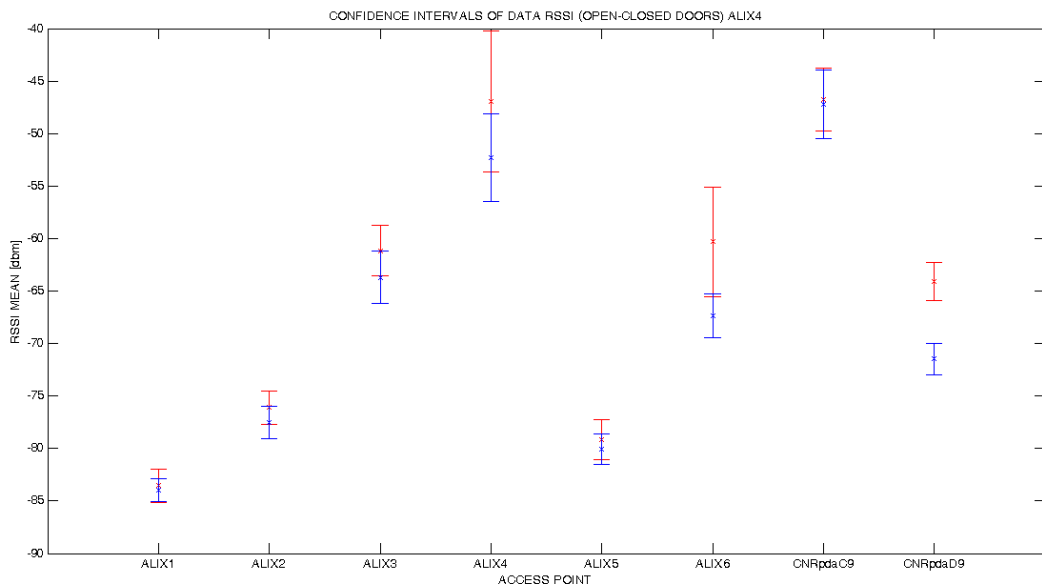
Misura network side relativamente all' Alix-1



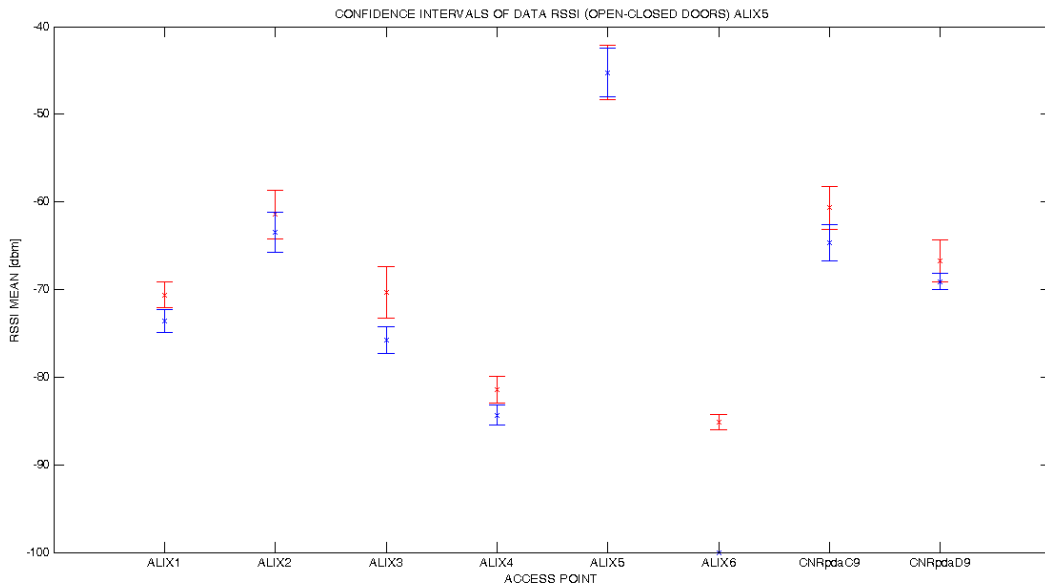
Misura network side relativamente all' Alix-2



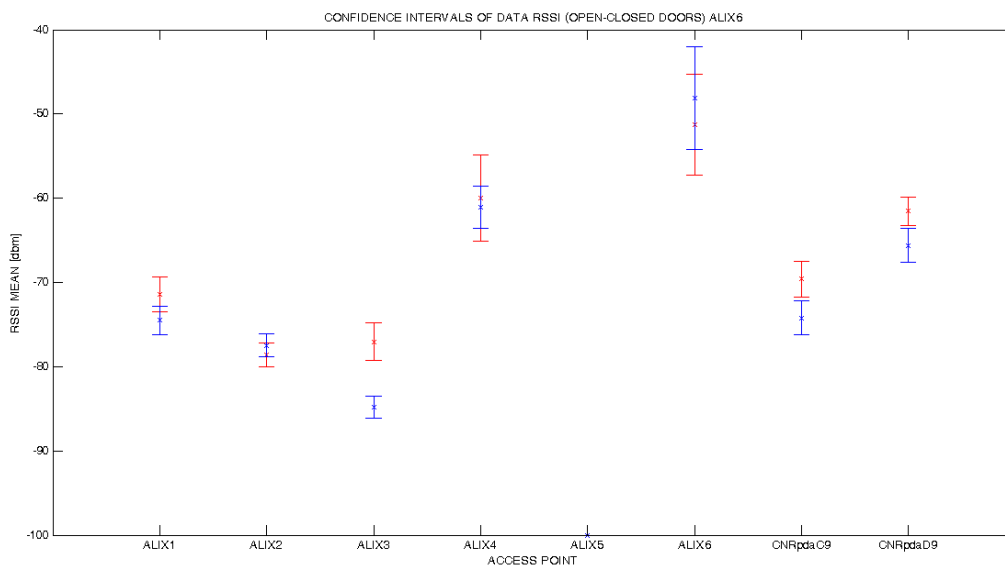
Misura network side relativamente all' Alix-3



Misura network side relativamente all' Alix-4



Misura network side relativamente all' Alix-5



Misura network side relativamente all' Alix-6

- c) comparazione dei valori di distanza complessiva del fingerprinting per i punti di test AreaP2 dalle due radio mappe.

La variazione del campo misurato in condizioni di porte aperte e chiuse sono osservabili per più delle volte, in alcuni casi il valore di potenza del segnale è oltre il livello minimo di sensibilità del ricevitore per cui punti in cui si riceveva il segnale nella configurazione a porte aperte, non sono più osservabili nella configurazione a porte chiuse.

Si osserva che in alcuni casi la chiusura delle porte genera delle microzone dell'ambiente di test in cui si ha un'interferenza costruttiva per cui la chiusura delle porte genera un aumento del segnale, casi in cui il segnale subisce un'interferenza distruttiva e si assiste quindi ad una diminuzione.

In ogni caso è osservabile che le variazioni di segnale non sono molto nette ed osservabili, questo perché rientrano negli intervalli di confidenza della deviazione standard e perché la chiusura delle porte non genera un così cospicuo abbattimento del segnale, così come si avrebbe se al posto delle porte si avesse considerato la presenza di nuovi muri per la valutazione dell'analisi delle variazioni del campo.

Nel seguito sono riportati in tabella i valori di potenza percepiti dalle alix nella configurazione di porte aperte e chiuse.

(*) il valore di -100 è un valore fittizio volto ad indicare che il ricevitore non ha percepito un segnale utile che si assesti alle soglie minima della sensibilità del ricevitore.

OPEN-CLOSE (indica le misurazioni effettuate dalle alix in presenza rispettivamente di porte aperte-chiuse)

ALIX-1				
	OPEN		CLOSE	
	RSSI	DEV.STD	RSSI	DEV.STD
ALIX-1	-51.61	2.83	-52.64	1.42
ALIX-2	-63.31	1.31	-65.14	1.22
ALIX-3	-100	0	-100	0
ALIX-4	-83.18	2.26	-100	0
ALIX-5	-73.67	1.34	-77.81	1.09
ALIX-6	-77.02	1.92	-74.89	2.51
CnrPdaC9	-83.50	1.83	-80.39	1.02
CnrPdaD9	-50.30	1.79	-57.40	2.60

ALIX-2				
	OPEN		CLOSE	
	RSSI	DEV.STD	RSSI	DEV.STD
ALIX-1	-56.63	1.07	-63.87	1.49
ALIX-2	-50.90	2.33	-51.45	1.42
ALIX-3	-73.10	2.70	-81.50	1.35
ALIX-4	-74.12	2.37	-81.95	1.73
ALIX-5	-64.09	4.78	-67.18	4.11
ALIX-6	-71.03	1.82	-75.53	1.82
CnrPdaC9	-66.64	3.67	-66.82	1.94
CnrPdaD9	-52.36	2.56	-49.68	2.19

ALIX-3				
	OPEN		CLOSE	
	RSSI	DEV.STD	RSSI	DEV.STD
ALIX-1	-100	0	-100	0
ALIX-2	-75.04	2.74	-78.86	1.37
ALIX-3	-40.03	6.91	-30.33	5.31
ALIX-4	-57.51	3.50	-68.39	3.00
ALIX-5	-69.37	2.08	-77.01	1.84
ALIX-6	-73.27	1.65	-79.36	1.48
CnrPdaC9	-47.56	3.34	-47.90	4.13
CnrPdaD9	-80.97	1.56	-100	0

ALIX-4				
	OPEN		CLOSE	
	RSSI	DEV.STD	RSSI	DEV.STD
ALIX-1	-83.55	1.58	-83.98	1.06
ALIX-2	-76.11	1.60	-77.50	1.54
ALIX-3	-61.17	2.40	-63.71	2.49
ALIX-4	-46.96	6.73	-52.31	4.17
ALIX-5	-79.12	1.90	-80.05	1.44
ALIX-6	-60.30	5.23	-67.34	2.11
CnrPdaC9	-46.78	2.98	-47.18	3.27
CnrPdaD9	-64.11	1.80	-71.48	1.53

ALIX-5				
	OPEN		CLOSE	
	RSSI	DEV.STD	RSSI	DEV.STD
ALIX-1	-70.60	1.47	-73.61	1.34
ALIX-2	-61.44	2.76	-63.45	2.25
ALIX-3	-70.33	2.99	-75.76	1.51
ALIX-4	-81.42	1.55	-84.33	1.13
ALIX-5	-45.25	3.08	-45.23	2.79
ALIX-6	-85.13	0.84	-100	0
CnrPdaC9	-60.66	2.46	-64.67	2.04
CnrPdaD9	-66.77	2.39	-69.11	0.94

ALIX-6				
	OPEN		CLOSE	
	RSSI	DEV.STD	RSSI	DEV.STD
ALIX-1	-71.44	2.05	-74.51	1.73
ALIX-2	-78.58	1.41	-77.49	1.36
ALIX-3	-77.03	2.22	-84.77	1.31
ALIX-4	-60.02	5.12	-61.03	2.51
ALIX-5	-100	0	-100	0
ALIX-6	-51.27	6.01	-48.13	6.08
CnrPdaC9	-69.58	2.11	-74.22	1.99
CnrPdaD9	-61.55	1.73	-65.59	1.99

3.6 Valutazione dei risultati degli esperimenti e problemi ancora aperti

Le misure condotte hanno evidenziato la possibilità di osservare con un sistema che utilizzi alix2d2, la possibilità di osservare le variazioni di segnale in seguito a variazioni topologiche, simulate nel caso in questione dall'aver condotto delle misure in condizioni di porte aperte e chiuse.

Uno dei problemi ancora aperti e che dovrebbero essere investigati è la modalità di correzione della radiomappa una volta registrate le variazioni del campo a radiofrequenza generato dagli ap disposti nell'ambiente. Si potrebbe pensare di cambiare i valori dei pattern della radiomappa coerentemente con il fatto che la variazione interessi una microzona o l'intera area.

Andrebbe sicuramente approfondito l'aspetto legato agli effetti di antenna relativamente alle condizioni di trasmissione, che di percezione dell'RSSI, aspetto delicato in un sistema di localizzazione basato sull'analisi di potenza dei segnali trasmessi da dispositivi 802.11.

Rimane da approfondire la modalità di riferire la variazione intercettata dallo smartphone con quella percepita dall'alice in relazione ad una possibile fase di equalizzazione

4 Conclusioni

4.1 Tecniche di ottimizzazione e regole di best practice

Gli esperimenti sopra riportati e l'esperienza empirica accumulata portano alla seguente valutazione riassuntiva degli effetti sulla accuratezza di localizzazione ed alla assunzione delle tecniche di ottimizzazione sotto indicate:

Trattamento della assenza di segnale: Il Test B indica che la mancata rilevazione istantanea di segnale è da considerarsi come segnale debole e che l'assegnazione forzata del valore minimo rilevato nel sito è preferibile all'assegnazione di un valore fuori scala o al non trattamento della rilevazione.

Trattamento dei punti di riferimento ad alta fluttuazione: Il test C indica che laddove il segnale mostra variabilità molto superiore alla media, la sua distribuzione tende ad allontanarsi da quella normale ed appare correlata a condizioni geometriche quali vicinanza alla sorgente, probabilmente anche a riflessioni in sua prossimità. In mancanza di modellazione maggiormente probatoria l'affidabilità dei punti di riferimento ad alta variabilità è limitata ed è conveniente escludere tali punti dal calcolo della radio mappa, limitatamente alla relativa componente del fingerprinting.

L'esclusione di tali outliers non diminuisce direttamente la presenza di falsi positivi fra i primi K vicini individuati dall'algoritmo di clustering KNN, ma migliora il ranking dei vicini corretti geograficamente.

Numero minimo di componenti da garantire nel pattern: Il test D indica che, nel caso di ambiente strutturato un numero di 6 componenti valide, pari ad una densità di 1 AP per 100 mq è sufficiente a garantire un pattern discriminante.

Il valore di riferimento indicato da CISCO nelle sue regole di best practice è pari ad 1 AP per 50 mq circa (valore doppio).

Densità spaziale di campionamento: Il test E mostra come l'indeterminazione introdotta dalla interpolazione fra vicini (baricentro) sia comparabile o inferiore alla distanza media fra i punti di riferimento (grossolanamente approssimata dalla misura della diagonale della maglia di campionamento). Il costo della costruzione della radio mappa cresce però almeno linearmente col numero di punti campionamento.

Selezione dell'algoritmo di clustering: Il test F suggerisce che l'algoritmo più performante è anche il più semplice: il KNN con valore di k fra 3 e 5. Le prestazioni sono superiori a quelle del KMeans e degli algoritmi di autoapprendimento (posizione e rank). Il test F mostra come la variabilità del segnale campionato induca difformità del pattern di test dai fingerprint dei punti geograficamente vicini, mantenendo l'indice di recall del KNN al di sotto del 50%, anche per alti valori di k.

Incertezza sulla stima della posizione dell'operatore: il test G mostra, indirettamente, come l'incertezza nella capacità di stima dell'operatore della propria posizione contribuisca alla incertezza totale. Tale incertezza è stimata pari ad un ennesimo (fra $\frac{1}{4}$ ed $\frac{1}{6}$) della distanza media fra i punti di riferimento disponibili (circa 2 m nel sito PoloP1 e 20 cm nel sito AreaP2) e trova conferma nell'aumento di 3.5 metri nell'errore medio di localizzazione nel sito PoloP1 e di 20 cm nel sito AreaP2 in caso di utilizzo degli stessi punti di campionamento come punti di test.

Valutazione dei contributi all'errore di localizzazione : Gli esperimenti sopra riportati e l'esperienza empirica accumulata portano alla seguente valutazione riassuntiva dei contributi all'errore di localizzazione :

Variabile\Sito	PoloP1	AreaP2
Errore totale	7.6	2.4
Autostima Posizione operatore	<2 metri	<0.5 metri
Modello terminale	N/A	N/A
Correzione segnale N/A	< 0.5 m	0.2 m
Filtraggio outliers in radiomappa	< 0.5 m	0.1 m

4.2 Varianti alle regole CISCO per la pianificazione della posizione AP

Gli esperimenti sopra riportati e l'esperienza empirica accumulata portano alla formulazione delle seguenti modifiche alle regole CISCO per la pianificazione della posizione AP:

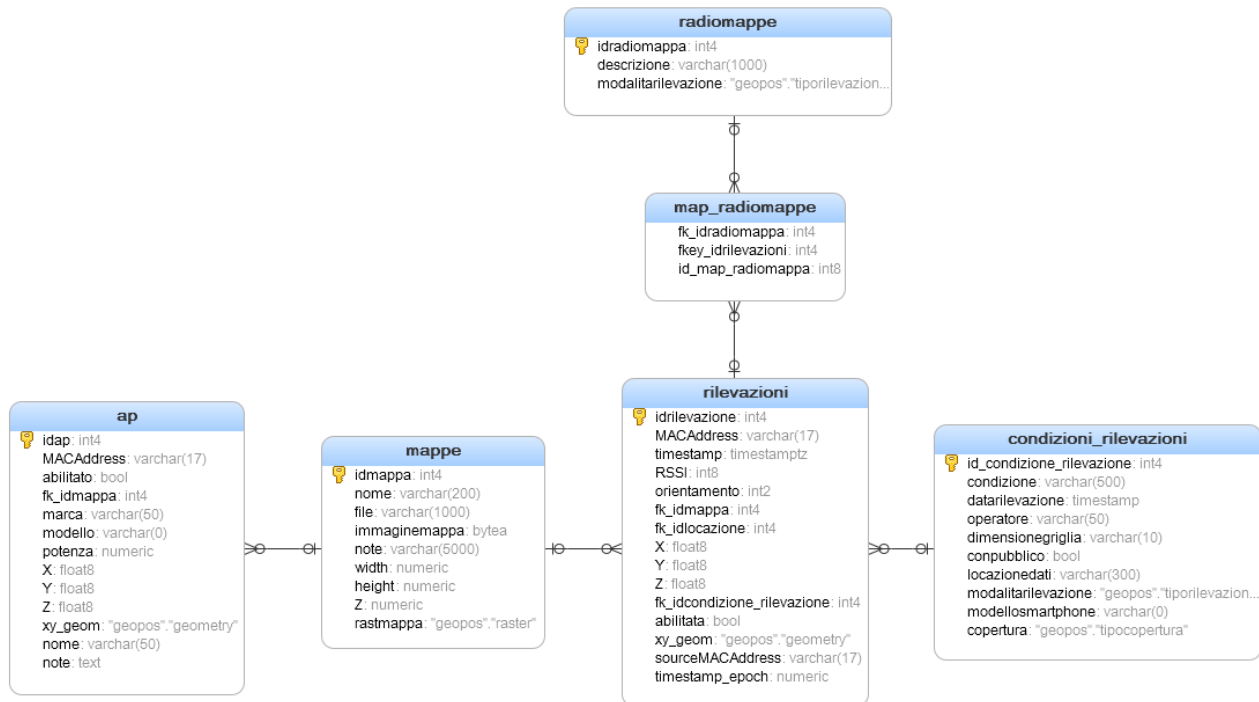
- piazzare AP con linea di vista relativa che attraversi le aree dove ci si aspettano variazioni (es. chiusura porte metalliche o zone di affollamento)
- ove permesso dalla accessibilità e dalla disponibilità di tempo operatore, campionare lo spazio ad alta risoluzione (1 metro) , assegnando ad operatori diversi piccole zone intervallate uniformemente.
- acquisire in ogni punto un numero di samples elevato (10-20 samples)

4.3 Ulteriori test da sviluppare

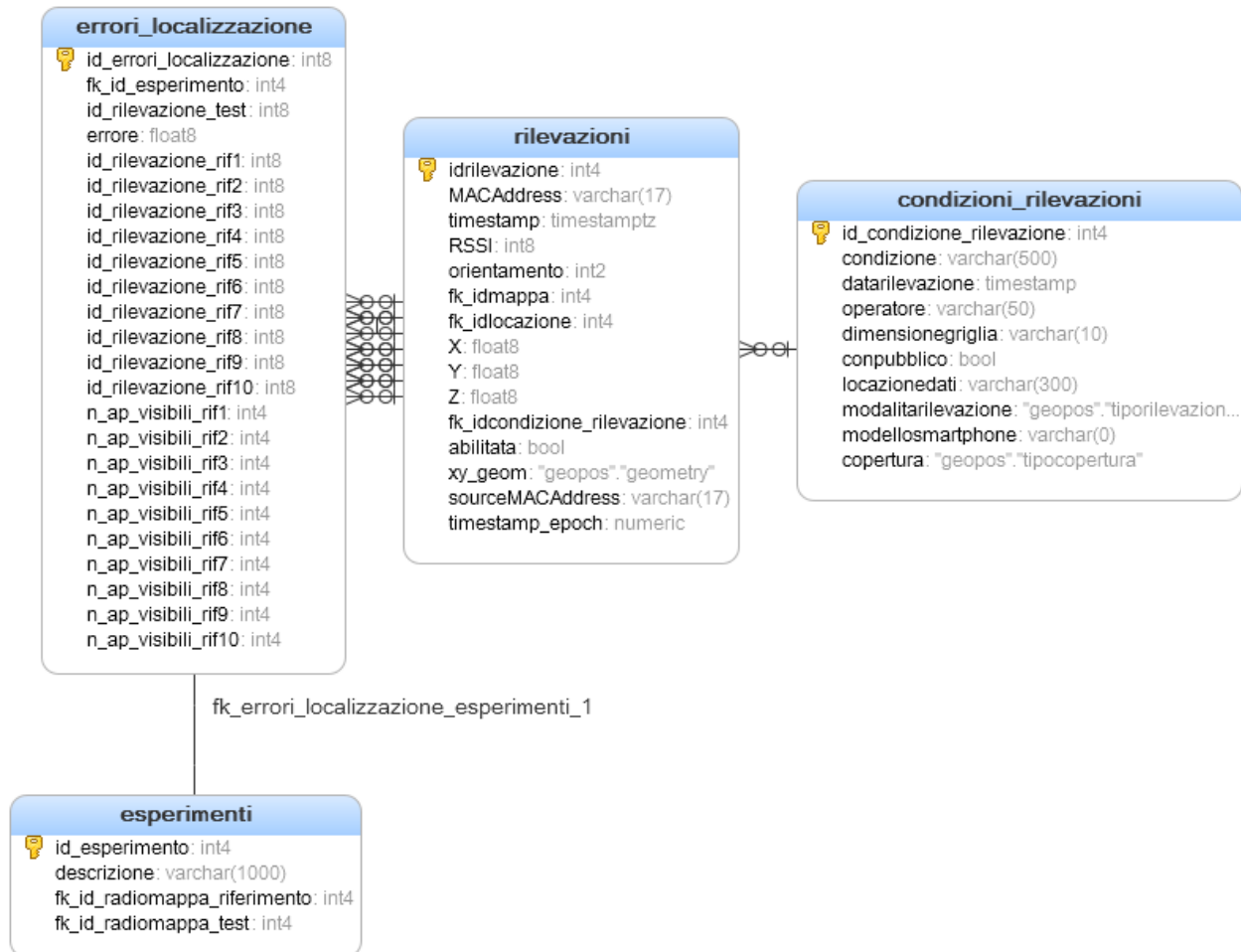
- Comparazione valori e distribuzioni segnali acquisiti da mobile di categoria diversa (equalizzazione, test A).
- attribuzione di pesi alle componenti del fingerprint in base alla variabilità del segnale nel punto (attualmente si usa peso 1 per componenti a varianza sotto soglia e 0 sopra soglia), trasformando il KNN in un WKNN (weighted KNN).
- Sviluppo di algoritmo di localizzazione iterativo che, individuata nel passo precedente la posizione approssimativa con l'algoritmo WKNN sopra descritto, ricalcoli i vicini attribuendo pesi anche alle componenti del segnale istantaneo in base alla variabilità del segnale nelle vicinanze punto stimato.

5 Appendice A: struttura della Base di Dati ed esempi di query in PostgreSQL

Il diagramma di seguito descrive la struttura tabellare, i campi e le relazioni del dataBase relazionale realizzato su PostgreSQL, al fine di gestire, filtrare, analizzare le rilevazioni di RSSI ottenute sul campo, nelle diverse locazioni di test (Polo Didattico, Area CNR...).



In aggiunta a tale struttura dati, ne è stata realizzata un'altra, ad essa correlata, per analizzare e memorizzare in maniera sistematica, i risultati dei test e delle simulazioni. Tale struttura prevede la memorizzazione degli errori di localizzazione ottenuti con le informazioni riguardanti le condizioni sperimentali di esecuzione.



5.1 Query di esempio

A titolo di esempio delle potenzialità della struttura dati implementata si riporta una query per estrarre il numero di AP visibili in ogni punto di rilevazione preso come riferimento nelle misure effettuate all'Area del CNR:

```

SELECT count(*) AS n_ap_visibili,
       b."X",
       b."Y"
FROM ( SELECT max(h.rssi_media) AS rssi_media,
              max(h.rssi_dev) AS rssi_dev,
              avg(h."RSSI") AS rssi_media_integrata,
              stddev_pop(h."RSSI") AS rssi_dev_integrata,
              h."X",
              h."Y",
              h.xy_geom,
    
```

```

        h."MACAddress",
        (min(min(h.idrilevazione)) OVER (PARTITION BY h."X", h."Y")) AS
idrilevazione,
        (min(min(h."RSSI")) OVER ()):character varying(24) AS min_rssi
FROM ( SELECT d.idrilevazione,
        d.rssi_media,
        d.rssi_dev,
        COALESCE(d."RSSI", d."RSSI", ((-110))::bigint) AS "RSSI",
        max(d.min_rssi) OVER ( ) AS minrssi,
        c."X",
        c."Y",
        c."MACAddress",
        c.xy_geom
FROM (( SELECT a."MACAddress",
        b_1."X",
        b_1."Y",
        b_1.xy_geom,
        b_1."timestamp"
FROM ( SELECT DISTINCT
punti_riferimento_area_8ap."MACAddress"
        FROM punti_riferimento_area_8ap
        ORDER BY punti_riferimento_area_8ap."MACAddress") a,
( SELECT count(*) OVER (PARTITION BY
punti_riferimento_area_8ap."X", punti_riferimento_area_8ap."Y") AS samples,
        punti_riferimento_area_8ap."timestamp",
        punti_riferimento_area_8ap.xy_geom,
        punti_riferimento_area_8ap."X",
        punti_riferimento_area_8ap."Y"
FROM punti_riferimento_area_8ap
        GROUP BY punti_riferimento_area_8ap.xy_geom,
punti_riferimento_area_8ap."X", punti_riferimento_area_8ap."Y",
punti_riferimento_area_8ap."timestamp"
        ORDER BY punti_riferimento_area_8ap."X",
punti_riferimento_area_8ap."Y", punti_riferimento_area_8ap."timestamp") b_1) c
LEFT JOIN ( SELECT punti_riferimento_area_8ap.idrilevazione,
stddev_pop(punti_riferimento_area_8ap."RSSI") OVER w AS
rssi_dev,
        avg(punti_riferimento_area_8ap."RSSI") OVER w AS
rssi_media,
        row_number() OVER w AS indice,
        (min(punti_riferimento_area_8ap."RSSI") OVER ( ) - 1) AS
min_rssi,
        punti_riferimento_area_8ap."RSSI",
        punti_riferimento_area_8ap."X",
        punti_riferimento_area_8ap."Y",
        punti_riferimento_area_8ap."MACAddress",
        punti_riferimento_area_8ap."timestamp"
FROM punti_riferimento_area_8ap
WINDOW w AS (PARTITION BY punti_riferimento_area_8ap."X",
punti_riferimento_area_8ap."Y", punti_riferimento_area_8ap."MACAddress" ORDER BY
punti_riferimento_area_8ap."X", punti_riferimento_area_8ap."Y",
punti_riferimento_area_8ap."MACAddress")) d ON (((((c."X" = d."X") AND (c."Y" = d."Y"))
AND ((c."MACAddress")::text = (d."MACAddress")::text)) AND (c."timestamp" =
d."timestamp")))) h
        GROUP BY h.xy_geom, h."X", h."Y", h."MACAddress"
        HAVING (max(h."RSSI") > max(h.minrssi))
        ORDER BY (h."X")::real, (h."Y")::real, h."MACAddress") b
GROUP BY b.xy_geom, b."X", b."Y"
ORDER BY count(*);

```

Il risultato di questa interrogazione è il seguente:

N. AP Visibili	X	Y
5	5.35688161849976	2.40168285369873
5	9.24020385742188	8.29541206359863
5	2.41875553131104	1.38653826713562
5	1.33258509635925	2.3778247833252
6	2.39630246162415	14.342604637146
6	3.39280295372009	4.31631231307983
6	7.32176113128662	1.35264337062836
6	22.2768115997314	20.2483654022217
6	22.2122974395752	14.3355865478516
6	1.244961977005	9.4776496887207
6	23.344554901123	21.3608875274658
6	21.3599033355713	21.3817291259766
6	1.32934284210205	4.3429069519043
6	21.3081836700439	17.2880668640137
6	21.3017864227295	15.4696254730225
6	5.36778879165649	21.3562812805176
6	21.2471408843994	23.5459136962891
6	23.2376365661621	10.1793031692505
6	22.4060211181641	22.3934764862061
6	2.42199778556824	5.35442686080933
6	24.4214382171631	5.21859455108643
6	20.2515449523926	21.4296684265137
6	19.4779758453369	23.4294815063477
6	5.55060720443726	4.21908140182495
6	22.3702774047852	8.39457511901855
6	2.53406977653503	20.2506980895996
6	19.4273242950439	21.1989803314209
6	10.3879117965698	3.36776781082153
6	19.3222103118896	2.54527187347412
6	17.231632232666	1.3412446975708
6	22.3285331726074	16.1253890991211
6	1.37149262428284	15.349310874939
6	10.464394569397	1.50427794456482
6	24.3059234619141	3.20849084854126
6	24.222827911377	11.4228115081787
6	11.4017820358276	3.31819629669189
6	15.4250526428223	8.32624340057373
6	4.27301120758057	3.36055731773376
6	14.2956237792969	5.34009218215942
6	23.3498001098633	4.20310544967651
6	12.189003944397	1.31287825107574
6	24.3295288085938	22.3139953613281
6	4.37010908126831	5.41080570220947
6	24.3182182312012	16.249626159668
7	8.37900161743164	13.3975791931152
7	8.38904476165771	11.2114925384521
7	8.43686866760254	9.42491245269775
7	9.33855152130127	4.20840454101562

7	9.34511756896973	13.2530689239502
7	9.34621238708496	2.48187947273254
7	9.4395580291748	6.42856311798096
7	9.44360256195068	11.2386293411255
7	10.3647203445435	5.23125743865967
7	10.4277429580688	12.2898845672607
7	10.4324655532837	9.25116729736328
7	11.2865343093872	11.3286581039429
7	11.3378925323486	5.33039903640747
7	11.5294208526611	17.1473770141602
7	22.1929244995117	10.4333610534668
7	1.25336873531342	23.2333698272705
7	1.3462108373642	21.4450435638428
7	1.39435088634491	13.3836612701416
7	1.41736102104187	17.2344169616699
7	1.4480082988739	11.4180002212524
7	2.3830246925354	18.2055892944336
7	2.39930176734924	3.32471132278442
7	2.41663670539856	12.4081258773804
7	2.46031641960144	22.4509773254395
7	2.54036545753479	10.3190326690674
7	3.21676135063171	2.49729490280151
7	3.39010143280029	9.26208686828613
7	3.40319871902466	19.3472652435303
7	3.40381407737732	15.4237508773804
7	3.44676446914673	13.2940368652344
7	3.45989561080933	11.2044267654419
7	3.47855925559998	23.2207336425781
7	3.48929572105408	21.2783908843994
7	4.22500038146973	1.38738608360291
7	4.36616182327271	20.3424835205078
7	4.39406538009644	10.2589483261108
7	4.41261720657349	22.2444229125977
7	4.42627811431885	12.3348903656006
7	5.3152117729187	19.3214073181152
7	5.34979438781738	10.275463104248
7	5.37573289871216	12.3146486282349
7	5.39064741134644	23.274169921875
7	5.43857526779175	15.2302980422974
7	6.31326532363892	17.3989582061768
7	6.33545351028442	22.3629875183105
7	6.3358588218689	8.35932159423828
7	6.34518051147461	10.2880897521973
7	6.3905725479126	6.32400798797607
7	6.40610694885254	19.2990531921387
7	7.28810739517212	13.4199342727661
7	7.30360412597656	5.36415529251099
7	7.30749464035034	9.38979434967041
7	7.31552743911743	11.3265466690063
7	7.31722164154053	3.31001472473145
7	7.33452844619751	15.3870916366577
7	7.48819923400879	20.2269325256348
7	8.30596542358398	2.42021799087524

7	8.36710834503174	19.2549591064453
7	13.2869396209717	5.24958610534668
7	13.394305229187	17.3909454345703
7	14.3179187774658	19.3028125762939
7	14.338436126709	17.2458190917969
7	14.3494091033936	12.3972091674805
7	15.2471265792847	6.46741056442261
7	15.3636360168457	2.41855049133301
7	15.4266424179077	18.2441940307617
7	16.3162708282471	9.29227542877197
7	17.232349395752	3.33976793289185
7	17.2509899139404	15.3457107543945
7	17.2822685241699	9.36195373535156
7	17.2892875671387	19.2127628326416
7	17.2918319702148	21.3369789123535
7	17.3018016815186	23.2620086669922
7	17.3109378814697	5.34604167938232
7	17.3948860168457	7.27478456497192
7	17.3983955383301	11.2437715530396
7	18.212043762207	12.3304634094238
7	18.2861328125	2.29741621017456
7	18.3330173492432	4.4289755821228
7	18.3643836975098	10.3050107955933
7	18.3652763366699	22.3540363311768
7	18.4698066711426	6.36623954772949
7	19.325403213501	12.3644819259644
7	19.325403213501	10.3335762023926
7	19.3254833221436	19.2859954833984
7	19.3421001434326	17.3096313476562
7	19.4258193969727	4.25547647476196
7	20.1510734558105	4.23069095611572
7	20.2460517883301	17.3032150268555
7	20.2560234069824	19.279577255249
7	20.2792053222656	12.3356075286865
7	20.2792053222656	10.3592433929443
7	20.2805843353271	2.2478449344635
7	20.2811088562012	14.3308639526367
7	20.3960742950439	6.10385608673096
7	21.2630424499512	11.35142993927
7	21.2858581542969	3.30841064453125
7	21.2858581542969	1.41800117492676
7	21.3164978027344	5.3245677947998
7	21.3275566101074	19.4239807128906
7	21.4180221557617	9.49655914306641
7	21.4599723815918	7.39178943634033
7	22.1412029266357	18.2061367034912
7	22.2750835418701	2.44075036048889
7	22.3101005554199	4.34373474121094
7	22.3702774047852	6.23216009140015
7	22.4641418457031	12.4432420730591
7	23.3080806732178	8.37946510314941
7	23.3275356292725	6.35104322433472
7	23.3566398620605	19.396427154541

7	23.3578491210938	23.2795009613037
7	23.3752326965332	17.2421264648438
7	24.2841739654541	7.37348222732544
7	24.3082485198975	14.3759307861328
7	24.3110523223877	9.43037128448486
7	12.4248208999634	5.12562084197998
7	12.4690027236938	3.27127647399902
8	12.3025827407837	13.3528709411621
8	12.3251581192017	8.25517559051514
8	12.3317623138428	11.2989883422852
8	12.3430681228638	18.3839588165283
8	3.33754205703735	17.3203964233398
8	8.28615951538086	16.39302444445801
8	5.3152117729187	17.3318958282471
8	13.3321046829224	9.38377475738525
8	13.3530874252319	12.3298110961914
8	13.3555593490601	19.3394985198975
8	23.3477993011475	13.3139553070068
8	13.4247159957886	3.42693996429443
8	13.4648513793945	7.21685314178467
8	13.519326210022	1.33086287975311
8	14.2591257095337	7.19811677932739
8	11.5696277618408	8.30900478363037
8	14.3126449584961	9.35574436187744
8	19.3649940490723	14.423996925354
8	14.3218860626221	1.41675925254822
8	4.33333349227905	18.2804260253906
8	14.339394569397	3.32812666893005
8	10.3735752105713	19.3103675842285
8	9.38483619689941	20.3132476806641
8	15.2979793548584	13.2747459411621
8	15.3329963684082	4.31734323501587
8	15.3417510986328	11.3020582199097
8	7.41303682327271	18.2460517883301
8	11.4183025360107	19.3957061767578
8	7.39703178405762	7.34508037567139
8	15.442458152771	20.3773136138916
8	16.2159404754639	7.1861457824707
8	16.3047142028809	12.2974691390991
8	9.35095405578613	18.3192901611328
8	10.4105377197266	17.2957649230957
8	4.25208282470703	16.2150745391846
8	7.33570241928101	23.3111171722412
8	2.3924925327301	16.3192768096924
8	6.31472730636597	14.19606590271
8	6.29654598236084	12.3146486282349
8	17.30153465271	13.3422107696533
8	20.3698120117188	8.36157035827637
8	23.3644161224365	15.3448610305786
8	8.39897918701172	5.25536727905273
8	21.2630424499512	13.4648675918579
8	5.38545942306519	8.33904933929443
8	1.36367321014404	19.2538356781006

8	5.37573289871216	6.3535270690918
8	18.3420124053955	21.304895401001
8	18.3473625183105	8.38122463226318
8	18.3525829315186	16.3094940185547
8	8.37187480926514	7.36743402481079
8	8.33465480804443	22.3790321350098
8	18.3985424041748	14.2094659805298
8	4.26282072067261	14.293288230896
8	19.2263622283936	6.31774520874023
8	19.2855224609375	8.3763427734375
8	10.3879222869873	7.36095857620239

5-1 Ap Visibili per punto

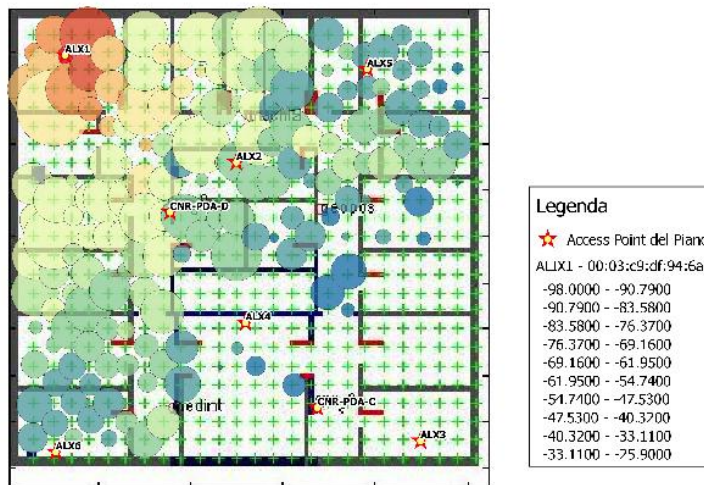
Il risultato permette di conoscere quali siano i punti che presentano una minore copertura in termini di Access Points, soprattutto se unito a strumenti di visualizzazione GIS, così come fatto in molte fasi di analisi. Si può così migliorare la disposizione degli apparati in fase di revisione della pianificazione, mettendo a disposizione uno strumento di supporto alle decisioni.

Questa è soltanto una delle possibilità la base di dati strutturata mette a disposizione, sia in fase di planning che in fase di analisi ed elaborazione.

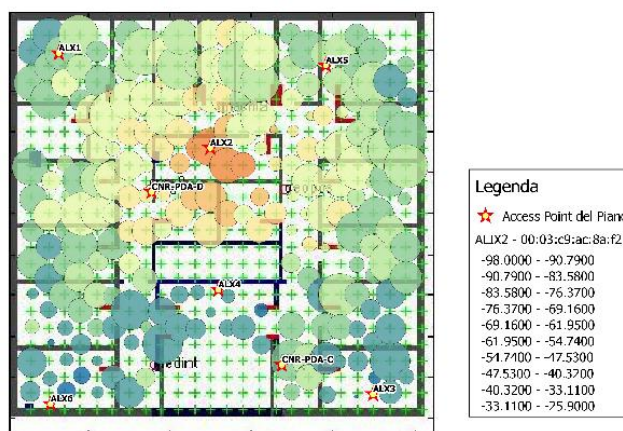
6 Appendice B: mappe di distribuzione spaziale del campo degli AP nel sito Area2P (porte aperte)

la Appendice B riporta le distribuzioni spaziali dei valori medi e delle devizioni standard dell'intensità del campo radio degli 8 AP utilizzati nella survey AREAP2, mediati sul numero di campioni effettivamente rilevati.

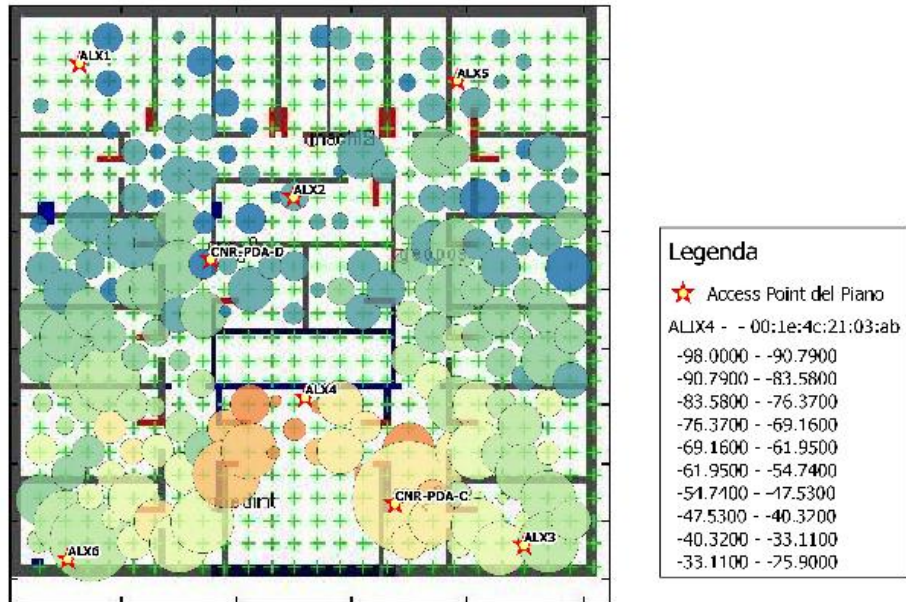
Mappa della distribuzione dei valori di RSSI (gradazioni da rosso a blu) con valori di deviazione standard (diametro circonferenza).
 Deviazione standard: 0/8,3
 RSSI: -25.9/-98



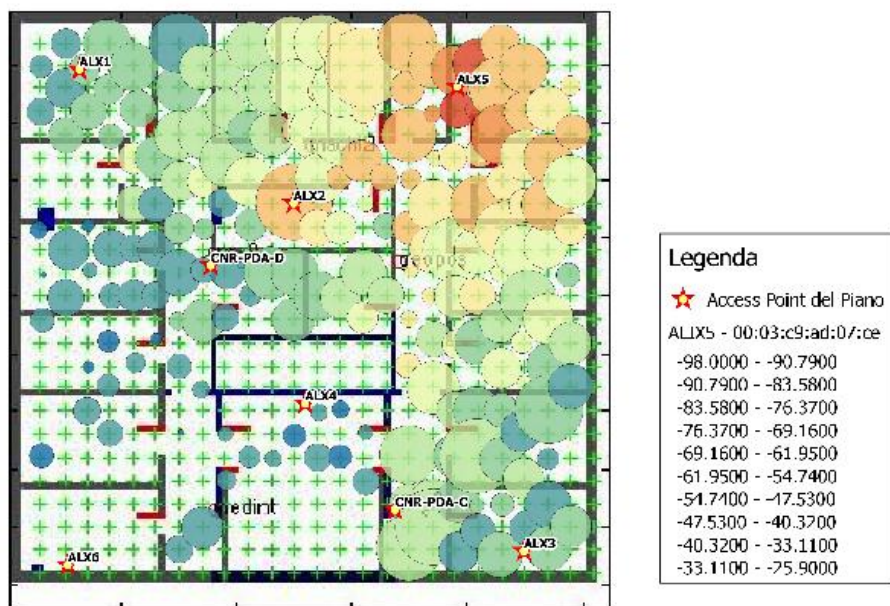
Mappa della distribuzione dei valori di RSSI (gradazioni da rosso a blu) con valori di deviazione standard (diametro circonferenza).
 Deviazione standard: 0/8,3
 RSSI: -25.9/-98



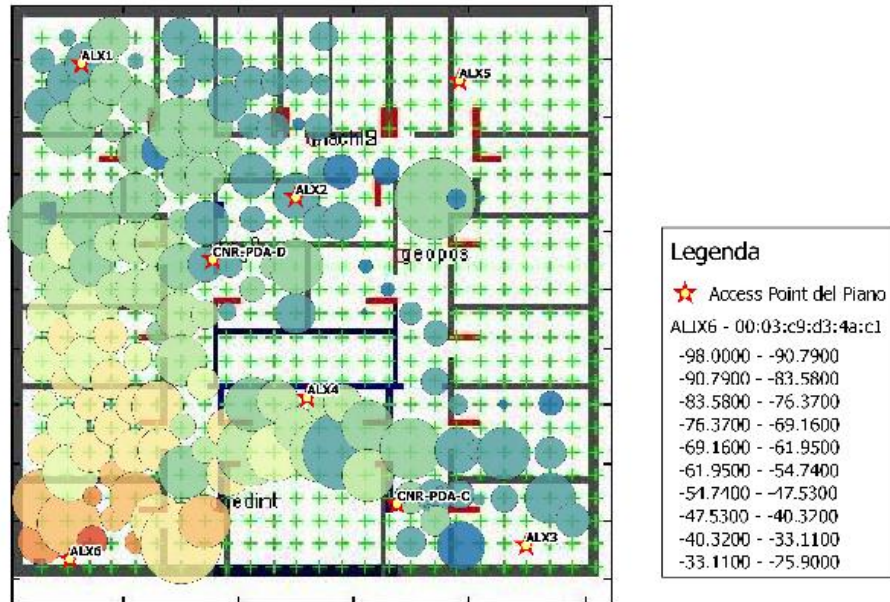
Mappa della distribuzione dei valori di RSSI (gradazioni da rosso a blu) con valori di deviazione standard (diametro circonferenza).
 Deviazione standard: 0/8.3
 RSSI: -25.9/-98



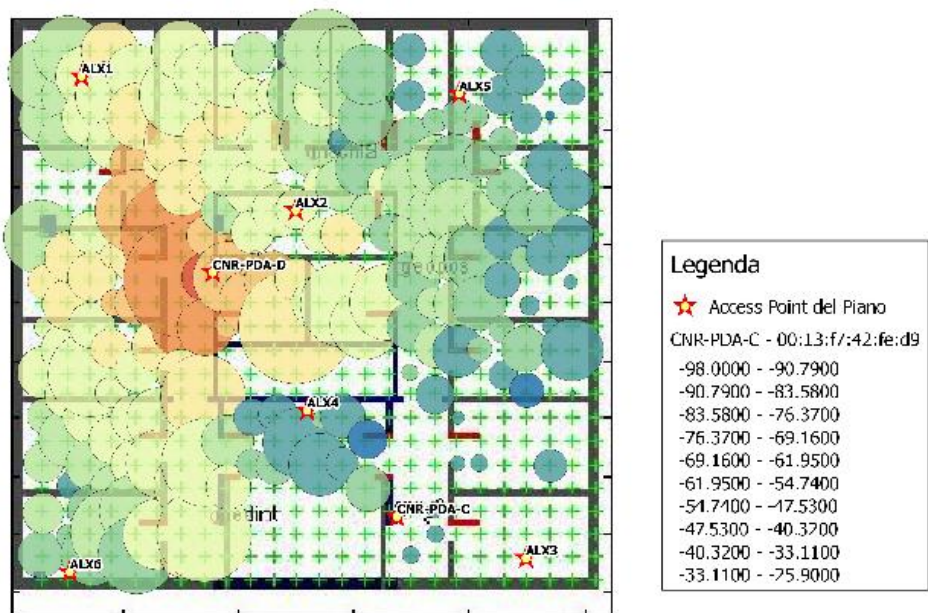
Mappa della distribuzione dei valori di RSSI (gradazioni da rosso a blu) con valori di deviazione standard (diametro circonferenza).
 Deviazione standard: 0/8.3
 RSSI: -25.9/-98



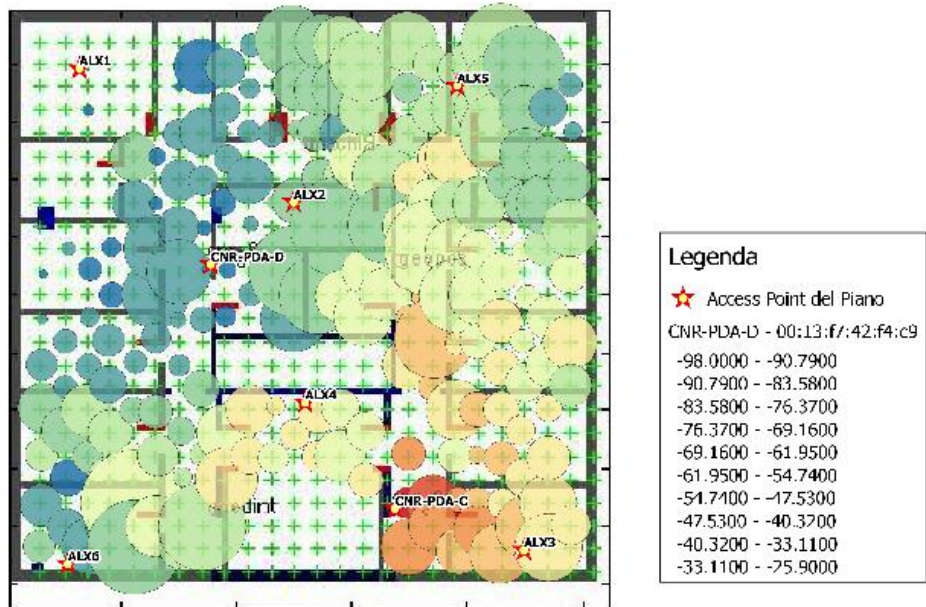
Mapa della distribuzione dei valori di RSSI (gradazioni da rosso a blu) con valori di deviazione standard (diametro circonferenza).
 Deviazione standard: 0/8.3
 RSSI: -25.9/-98



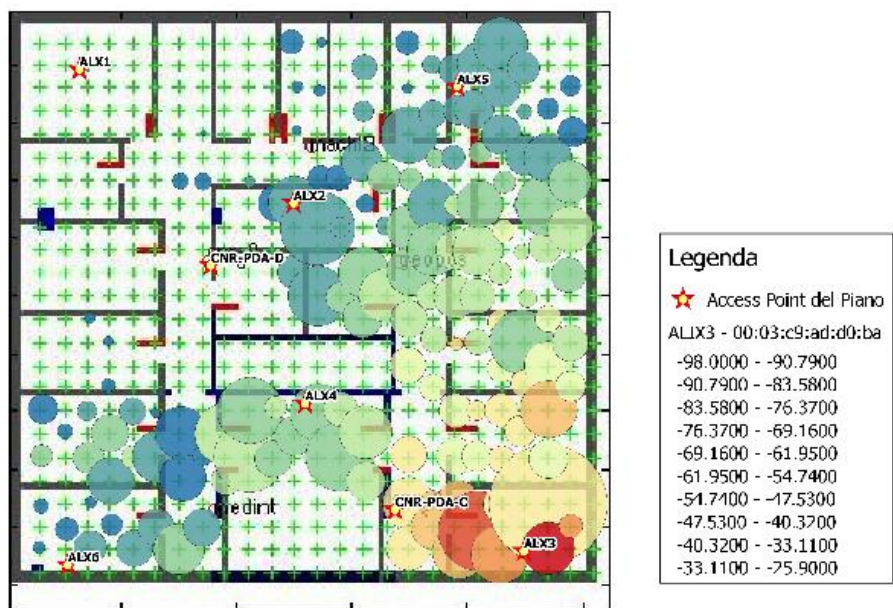
Mapa della distribuzione dei valori di RSSI (gradazioni da rosso a blu) con valori di deviazione standard (diametro circonferenza).
 Deviazione standard: 0/8.3
 RSSI: -25.9/-98



Mappa della distribuzione dei valori di RSSI (gradazioni da rosso a blu) con valori di deviazione standard (diametro circonferenza).
 Deviazione standard: 0/8.3
 RSSI: -25.9/-98



Mappa della distribuzione dei valori di RSSI (gradazioni da rosso a blu) con valori di deviazione standard (diametro circonferenza).
 Deviazione standard: 0/8.3
 RSSI: -25.9/-98



7 Appendice C: sorgente documentato del codice dei moduli mobile logger, radiomap LAN server e mobile localizer

A titolo esemplificativo si riportano le sezioni più significative del codice utilizzato nei dispositivi mobili e nel server, per effettuare applicativamente la localizzazione. La parte applicativa è stata implementata in Java.

7.1 Funzione che implementa l'algoritmo di localizzazione KNN

```
/**
 * Calculates user location based on Weighted/Not Weighted K Nearest
 * Neighbor (KNN) Algorithm
 *
 * @param RM
 *         The radio map structure
 *
 * @param Observed_RSS_Values
 *         RSS values currently observed
 * @param parameter
 *
 * @param isWeighted
 *         To be weighted or not
 *
 * @param K
 *         The number of locations used
 *
 * @return The estimated user location
 */
private static String KNN_WKNN_Algorithm(RadioMapMean RM, ArrayList<String>
Observed_RSS_Values, String parameter, boolean isWeighted) {

    ArrayList<String> RSS_Values;
    float curResult = 0;
    ArrayList<LocDistance> LocDistance_Results_List = new ArrayList<LocDistance>();
    String myLocation = null;
    int K;

    try {
        K = Integer.parseInt(parameter);
    } catch (Exception e) {
        System.err.println("Error parsing K in KNN-WKNN: " + e.getMessage());
        return null;
    }

    // Construct a list with locations-distances pairs for currently
    // observed RSS values
    for (String location : RM.getLocationRSS_HashMap().keySet()) {
        RSS_Values = RM.getLocationRSS_HashMap().get(location);
        curResult = calculateEuclideanDistance(RM, RSS_Values,
Observed_RSS_Values);
```

```

        if (curResult == Float.NEGATIVE_INFINITY) {
            return null;
        }

        LocDistance_Results_List.add(0, new LocDistance(curResult, location));
    }

    // Sort locations-distances pairs based on minimum distances
    Collections.sort(LocDistance_Results_List, new Comparator<LocDistance>() {

        public int compare(LocDistance gd1, LocDistance gd2) {
            return (gd1.getDistance() > gd2.getDistance() ? 1 : (gd1.getDistance()
== gd2.getDistance() ? 0 : -1));
        }
    });

    if (lisWeighted) {
        myLocation = calculateAverageKDistanceLocations(LocDistance_Results_List,
K);
    } else {
        myLocation =
calculateWeightedAverageKDistanceLocations(LocDistance_Results_List, K);
    }

    return myLocation;
}

```

7.2 Funzione che implementa l'algoritmo di localizzazione KMEANS

```

/**
 * Calculates user location based on KMeans Algorithm
 *
 * @param RM
 *         The radio map structure
 *
 * @param Observed_RSS_Values
 *         RSS values currently observed
 * @param parameter
 *
 * @return The estimated user location
 */
private static String KMeans_Algorithm(RadioMapMean RM, ArrayList<String>
Observed_RSS_Values, String parameter) {
    FastVector      atts;
    Instances      data;
    double[]      vals;
    ArrayList<String> RSS_Values;
    ClusterEvaluation eval;
    int n = 2;
    int i,a,k;
    float curResult = 0;

```

```

        ArrayList<LocDistance> LocDistance_Results_List = new
ArrayList<LocDistance>();
        String myLocation = null;
        float sumX = 0.0f;
        float sumY = 0.0f;
        int vicini=50;
        int clustervic=0;
        int[] assignments,vec;
        int cluster0=0;
        int cluster1=0;

        String[] LocationArray = new String[2];
        float x,y,x1,x2,y1,y2;
        int K;

        try {
            K = Integer.parseInt(parameter);
        } catch (Exception e) {
            System.err.println("Error parsing K in KM: " + e.getMessage());
            return null;
        }

atts = new FastVector();
        for (String location : RM.getLocationRSS_HashMap().keySet()) {
            RSS_Values = RM.getLocationRSS_HashMap().get(location);
            for (i = 0; i < RSS_Values.size(); ++i){

                atts.addElement(new
Attribute("att"+String.valueOf(i+1)));
            }
            break;
        }

        // 2. create Instances object
data = new Instances("MyRelation", atts, 0);

        // 3. fill with data
for (String location : RM.getLocationRSS_HashMap().keySet()) {
            RSS_Values = RM.getLocationRSS_HashMap().get(location);

            vals = new double[data.numAttributes()];
            for (i = 0; i < RSS_Values.size(); ++i){
                // - numeric
                vals[i] = Double.parseDouble(RSS_Values.get(i));
            }
            // add
            data.add(new Instance(1.0, vals));
        }

        try {

            String[] options = weka.core.Utils.splitOptions("-I 100");
            SimpleKMeans kmeans = new SimpleKMeans();
            kmeans.setSeed(10);
            kmeans.setOptions(options);
            kmeans.setNumClusters(n);
            kmeans.setPreserveInstancesOrder(true);
            kmeans.buildClusterer(data);

```

```

//create cluster information print result
eval = new ClusterEvaluation();
eval.setClusterer(kmeans);
eval.evaluateClusterer(new Instances(data));

//this array returns the cluster number for each instance
//the array has as many elements as the number of instances
assignments = kmeans.getAssignments();

i = 0;
for(int clusternum : assignments){
    if (clusternum==0){
        cluster0=cluster0+1;
    }else{
        cluster1=cluster1+1;
    }
    i++;
}

// Construct a list with locations-distances pairs for currently
// observed RSS values
for (String location : RM.getLocationRSS_HashMap().keySet()) {
    RSS_Values = RM.getLocationRSS_HashMap().get(location);
    curResult = calculateEuclideanDistance(RSS_Values,
Observed_RSS_Values);

    if (curResult == Float.NEGATIVE_INFINITY)
        return null;

    LocDistance_Results_List.add(0, new LocDistance(curResult,
location));
}

// Sort locations-distances pairs based on minimum distances
Collections.sort(LocDistance_Results_List, new
Comparator<LocDistance>() {

    public int compare(LocDistance gd1, LocDistance gd2) {
        return (gd1.getDistance() > gd2.getDistance() ? 1 :
(gd1.getDistance() == gd2.getDistance() ? 0 : -1));
    }
});

i=0;
for(String s : RM.getLocationRSS_HashMap().keySet()){
    LocationArray =
LocDistance_Results_List.get(0).getLocation().split(" ");
    x1=Float.valueOf(LocationArray[0].trim()).floatValue();
    y1=Float.valueOf(LocationArray[1].trim()).floatValue();
    LocationArray = s.split(" ");
    x2=Float.valueOf(LocationArray[0].trim()).floatValue();
    y2=Float.valueOf(LocationArray[1].trim()).floatValue();
    if(x1==x2&&y1==y2){
        x = x1;
        y = y1;
        clustervic=assignments[i];
        if (assignments[i]==0){
            vicini=cluster0;
        }else{

```

```

        vicini=cluster1;
    }
}
i++;
}

i=0;
for(String s : RM.getLocationRSS_HashMap().keySet()){
    i++;
}
vec = new int[i];
i=0;
for(String s : RM.getLocationRSS_HashMap().keySet()){
    if (assignments[i]==clustervic){
        vec[i]=i;
    }
    else{
        vec[i]=-1;
    }
    i++;
}

while (vicini>=K){

    data = new Instances("MyRelation", atts, 0);
    // 3. fill with new data
    k = 0;
    a = 0;
    for (String location : RM.getLocationRSS_HashMap().keySet()) {
        if (vec[k]!=-1){
            RSS_Values = RM.getLocationRSS_HashMap().get(location);

            vals = new double[data.numAttributes()];
            for (i = 0; i < RSS_Values.size(); ++i){
                // - numeric
                vals[i] = Double.parseDouble(RSS_Values.get(i));
            }
            // add
            data.add(new Instance(1.0, vals));
            a++;
        }
        k++;
    }

    String[] options1 = weka.core.Utils.splitOptions("-I 100");
    SimpleKMeans kmeans1 = new SimpleKMeans();
    kmeans1.setSeed(10);
    kmeans1.setOptions(options1);
    kmeans1.setNumClusters(n);
    kmeans1.setPreserveInstancesOrder(true);
    kmeans1.buildClusterer(data);

    //create cluster information print result
    eval = new ClusterEvaluation();
    eval.setClusterer(kmeans1);
    eval.evaluateClusterer(new Instances(data));

    //this array returns the cluster number for each instance
    //the array has as many elements as the number of instances
    assignments = kmeans1.getAssignments();
}

```

```

        i = 0;
        cluster0=0;
        cluster1=0;
        for(int clusternum : assignments){
            if (clusternum==0){
                cluster0=cluster0+1;
            }else{
                cluster1=cluster1+1;
            }
            i++;
        }

        // Construct a list with locations-distances pairs for currently
        // observed RSS values
        k=0;
        for (String location : RM.getLocationRSS_HashMap().keySet())
    {
        if (vec[k]!=-1){
            RSS_Values = RM.getLocationRSS_HashMap().get(location);
            curResult = calculateEuclideanDistance(RSS_Values,
Observed_RSS_Values);

            if (curResult == Float.NEGATIVE_INFINITY)
                return null;

            LocDistance_Results_List.add(0, new
LocDistance(curResult, location));
        }
        k++;
    }

    // Sort locations-distances pairs based on minimum distances
    Collections.sort(LocDistance_Results_List, new
Comparator<LocDistance>() {

        public int compare(LocDistance gd1, LocDistance
gd2) {
            return (gd1.getDistance() >
gd2.getDistance() ? 1 : (gd1.getDistance() == gd2.getDistance() ? 0 : -1));
        }
    });

    i=0;
    k=0;
    for(String s : RM.getLocationRSS_HashMap().keySet()){
        if (vec[i]!=-1){

            LocationArray =
LocDistance_Results_List.get(0).getLocation().split(" ");
            x1=Float.valueOf(LocationArray[0].trim()).floatValue();
            y1=Float.valueOf(LocationArray[1].trim()).floatValue();
            LocationArray =s.split(" ");
            x2=Float.valueOf(LocationArray[0].trim()).floatValue();
            y2=Float.valueOf(LocationArray[1].trim()).floatValue();
            if(x1==x2&&y1==y2){

```

```

        x = x1;
        y = y1;
        clustervic=assignments[k];
        if (assignments[k]==0){
            vicini=cluster0;
        }else{
            vicini=cluster1;
        }
    }
    k++;
}
i++;
}

i=0;
k=0;
for(String s : RM.getLocationRSS_HashMap().keySet()){
    if (vec[i]!=-1){
        if (assignments[k]!=clustervic){
            vec[i]=-1;
        }
        k++;
    }
    i++;
}

}
i=0;
a=0;
for(String location : RM.getLocationRSS_HashMap().keySet()){
    if (vec[i]!=-1){
        LocationArray =location.split(" ");

x=Float.valueOf(LocationArray[0].trim()).floatValue();

y=Float.valueOf(LocationArray[1].trim()).floatValue();
        RSS_Values = RM.getLocationRSS_HashMap().get(location);
        sumX += x;
        sumY += y;
        a++;
    }
    i++;
}
sumX /= a;
sumY /= a;
myLocation= sumX + " " + sumY;

return myLocation;
} catch (Exception e) {
    return null;
}
}
}

```

8 Appendice D: sorgente documentato degli script di simulazione di algoritmi su piattaforma MATLAB

8.1 Modulo estrazione dati dal RDBMS

Di seguito viene riportato il codice matlab del codice necessario ad estrarre sia il dataset dei riferimenti che quello dei testpoint, effettuando la media dei valori di RSSI con un valore di fondoscala fittizio di -110:

```
conn = database(datasourcename,user,password,driver,databaseurl);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Estrazione Dataset riferimento
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

vistaPuntiRiferimento = 'punti_riferimento_polo_24ap';
%vistaPuntiRiferimento = 'punti_riferimento_area_8ap';

% stringa di estrazione dei dati
sqlStr = sprintf('SELECT MAX (rssi_media) :: VARCHAR (24) rssi_media, MAX
(rssi_dev) :: VARCHAR (24) rssi_dev, AVG ("RSSI") :: VARCHAR (24)
rssi_media_integrata, STDDEV_POP ("RSSI") :: VARCHAR (24) rssi_dev_integrata,
"X" :: VARCHAR (24), "Y" :: VARCHAR (24), "MACAddress", MIN (MIN(idrilevazione))
OVER (PARTITION BY "X", "Y") :: VARCHAR (24) idrilevazione, MIN (MIN("RSSI"))
OVER () :: VARCHAR (24) min_rssi FROM ( SELECT idrilevazione, rssi_media,
rssi_dev, COALESCE ("RSSI", "RSSI", %s) "RSSI", C ."X", C ."Y", C ."MACAddress"
FROM ( SELECT "MACAddress", "X", "Y", "timestamp" FROM ( SELECT DISTINCT
"MACAddress" FROM "%s" ) AS A, ( SELECT COUNT (*) OVER (PARTITION BY "X", "Y")
samples, "timestamp", "X", "Y" FROM "%s" GROUP BY "X", "Y", "timestamp" ORDER BY
"X", "Y", "timestamp" ) AS B ) AS C LEFT JOIN ( SELECT idrilevazione, STDDEV_POP
("RSSI") OVER w rssi_dev, AVG ("RSSI") OVER w rssi_media, ROW_NUMBER () OVER w
indice, MIN ("RSSI") OVER () - 1 min_rssi, "RSSI", "X", "Y", "MACAddress",
"timestamp" FROM "%s" WINDOW w AS ( PARTITION BY "X", "Y", "MACAddress" ORDER BY
"X", "Y", "MACAddress" )) AS D ON C ."X" = D."X" AND C ."Y" = D."Y" AND C
."MACAddress" = D."MACAddress" AND C ."timestamp" = D."timestamp" ) AS H GROUP
BY "X", "Y", "MACAddress" ORDER BY "X" :: REAL, "Y" :: REAL, "MACAddress", '-
110',vistaPuntiRiferimento,vistaPuntiRiferimento,vistaPuntiRiferimento);

curs = exec(conn,sqlStr);
curs = fetch(curs);
result_ap_filtrati = curs.Data;
close(curs);

result = result_ap_filtrati;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

APs=unique(result(:,7));
```



```

RSSIMIN = result(1,9);
[C, ia,ic] = unique(strcat(result(:,5),'|',result(:,6)));
PUNTI = sortrows(cellfun(@str2num,result(ia,[8 5 6])),2);
POINTS=PUNTI(:,2:3);
IDPOINTS=PUNTI(:,1);

RMMEAN = [];
RMDEV = [];
for i=0:(size(POINTS,1)-1)
    RMMEAN = [RMMEAN;
cellfun(@str2num,result(i*size(APs)+1:(i+1)*size(APs),3)')];
    RMDEV = [RMDEV; cellfun(@str2num,result(i*size(APs)+1:(i+1)*size(APs),4)')];
end

POINTSriferimento=POINTS;
IDPOINTSriferimento=IDPOINTS;
APsriferimento=APs;
RMMEANriferimento=RMMEAN;
RMDEVriferimento=RMDEV;

conn = database(datasourcename,user,password,driver,databaseurl);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Estrazione Dataset di Test
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

vistaPuntiTest = 'punti_test_polo_24ap';
sqlStr = sprintf('SELECT MIN(idrilevazione) OVER (PARTITION BY C ."X", C ."Y")
:: VARCHAR (24) idrilevazione, rssi_media :: VARCHAR (24), rssi_dev :: VARCHAR
(24), COALESCE ("RSSI", "RSSI", %d) :: VARCHAR (24) "RSSI", C ."X" ::
VARCHAR (24), C ."Y" :: VARCHAR (24), C ."MACAddress", samples::VARCHAR (24)
n_samples FROM (SELECT "MACAddress", "X", "Y", timestamp, samples FROM (select
distinct "MACAddress" from %s order by "MACAddress") AS A,(select count(*) over
(partition by "X","Y") samples, timestamp, "X","Y" from %s group by
"X","Y",timestamp order by "X","Y",timestamp) AS B) AS C LEFT JOIN ( SELECT
idrilevazione,stddev_pop("RSSI") OVER w rssi_dev,AVG ("RSSI") OVER w rssi_media,
timestamp,"RSSI", "X", "Y", "MACAddress" FROM %s WINDOW w AS (PARTITION BY
"X", "Y","MACAddress" ORDER BY "X", "Y", "MACAddress", timestamp)) AS D ON C
."X" = D."X" AND C ."Y" = D."Y" AND C ."MACAddress" = D."MACAddress" AND C
.timestamp = D.timestamp order by C."X"::REAL,C."Y"::REAL,C."MACAddress",
C.timestamp',
cellfun(@str2num,RSSIMIN),vistaPuntiTest,vistaPuntiTest,vistaPuntiTest);

curs = exec(conn,sqlStr);
curs = fetch(curs);

result_test_apfiltrati = curs.Data;
close(curs);

result = result_test_apfiltrati;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

APs=unique(result(:,7));

```

```

[C, ia,ic] = unique(strcat(result(:,5),'|',result(:,6)));
PUNTI = sortrows(cellfun(@str2num,result(ia,[1 5 6])),[2 3]);
POINTS=PUNTI(:,2:3);
IDPOINTS=PUNTI(:,1);

testpoints=double.empty(0,size(APs,1)+2);
n_sample= cellfun(@str2num,result(1,8));
incremento = 0;
for i=0:(size(POINTS,1)-1)
    for j = 1:n_sample
        testpoints = [testpoints; cellfun(@str2num,result(incremento+1,5:6))
cellfun(@str2num,result(incremento+j:n_sample:incremento+size(APs,1)*n_sample,4)
')]];
    end
    incremento = incremento + size(APs,1)*n_sample;
    if (size(result,1)>=incremento+1)
        n_sample= cellfun(@str2num,result(incremento+1,8));
    end
end

POINTStest=POINTS;
IDPOINTStest=IDPOINTS;
APstest=APs;

close(conn);

```

8.2 Modulo statistiche errore con algoritmo KNN

Di seguito il codice che simula la localizzazione su tutti i punti di test utilizzando l'algoritmo KNN in matlab. Successivamente al calcolo, verifica l'effettivo scostamento della posizione stimata da quella reale e calcola l'errore medio e la deviazione standard dell'errore:

```
ExtractPointsAreaTestPorteChiuse_2014_05_30
```

```

RMMEAN=RMMEANriferimento;
POINTS=POINTSriferimento;
N=10;
meanerrors=[];
stderrors=[];
for K=1:10

% connessione al DB
conn = database(datasourcename,user,password,driver,databaseurl);

POINTSSTIMATIKNN=[];
tablename = 'errori_localizzazione';
colnames={'fk_id_esperimento','id_rilevazione_test','errore','id_rilevazione_rif
1','id_rilevazione_rif2','id_rilevazione_rif3',...

'id_rilevazione_rif4','id_rilevazione_rif5','id_rilevazione_rif6','id_rilevazion
e_rif7','id_rilevazione_rif8','id_rilevazione_rif9','id_rilevazione_rif10'};

tic
fprintf(1,'testpoint:      ');
errore = [];
exdata = [];

```

```

MAX_RSSI_DEV_CUTOFF = max(max(RMDEVriferimento));
MIN_RSSI_DEV_CUTOFF = min(min(RMDEVriferimento));
while (MAX_RSSI_DEV_CUTOFF-MIN_RSSI_DEV_CUTOFF)>0.001

nzeri=sum((RMDEVriferimento<(MAX_RSSI_DEV_CUTOFF+MIN_RSSI_DEV_CUTOFF)/2))==0,2)
;
    if max(nzeri)<=2
        MAX_RSSI_DEV_CUTOFF = (MAX_RSSI_DEV_CUTOFF+MIN_RSSI_DEV_CUTOFF)/2;
    else
        MIN_RSSI_DEV_CUTOFF = (MAX_RSSI_DEV_CUTOFF+MIN_RSSI_DEV_CUTOFF)/2;
    end
end
RSSI_DEV_CUTOFF = (MAX_RSSI_DEV_CUTOFF+MIN_RSSI_DEV_CUTOFF)/2
IDX=zeros(size(testpoints,1),N);
for i=1:size(testpoints,1)
    [coordKpuntivicini,patternKpuntivicini] =
Kpuntiviciniadunpuntoditest3(RMMEANriferimento,RMDEVriferimento,POINTSriferiment
o, IDPOINTSriferimento,testpoints(i,:),N,-100, RSSI_DEV_CUTOFF);
    IDX(i,:)=coordKpuntivicini(:,3)';
    sumX=sum(coordKpuntivicini(1:K,1))/K;
    sumY=sum(coordKpuntivicini(1:K,2))/K;
    POINTSSTIMATIKNN=[POINTSSTIMATIKNN;sumX sumY];
    [Lia,Locb] = ismember(testpoints(i,1:2),POINTStest,'rows');
    exdata = [exdata; 26 IDPOINTStest(Locb)
pdist2(testpoints(i,1:2),POINTSSTIMATIKNN(i,:)) coordKpuntivicini(:,5)'];
    fprintf(1,'\b\b\b\b%4i',i);
end
toc

errorsKNN=diag(pdist2(testpoints(:,1:2),POINTSSTIMATIKNN));
meanerrorKNN=mean(errorsKNN);
stderrorKNN=std(errorsKNN);
meanerrors=[meanerrors meanerrorKNN];
stderrors=[stderrors stderrorKNN];
end

close(conn);
fprintf(1,'\n%s','Elaborazione completata');

function [coordKpuntivicini,patternKpuntivicini] =
Kpuntiviciniadunpuntoditest3(RMMEANriferimento,RMDEVriferimento,POINTSriferiment
o,IDPOINTSriferimento,testpoint,K, RSSI_AVG_CUTOFF, RSSI_DEV_CUTOFF)%con
distanza senza considerare le componenti che nel punto di riferimento sono a -
110
distanzedaipuntidiriferimento=pdist2(testpoint(1:2),POINTSriferimento);

RMMEAN=RMMEANriferimento;
POINTS=POINTSriferimento;
IDPOINTS=IDPOINTSriferimento;
RMMEANPROV=RMMEAN;
POINTS PROV=POINTS;
IDPOINTS PROV=IDPOINTS;
disttestpointrif=[];
for l=1:size(RMMEANPROV,1)

    % algoritmo modificato
    %rifprov=RMMEANriferimento(l,:).*((RMDEVriferimento(l,:)<RSSI_DEV_CUTOFF));

    % algoritmo KNN Standard con media -110
    rifprov=RMMEANriferimento(l,:);

```

```

    testpointprov=testpoint (find(~isnan (rifprov) & (rifprov~=0))+2);
    if length(testpointprov)>=3
        disttestpointrif=[disttestpointrif; ((sum(((testpointprov-
rifprov (~isnan (rifprov) & (rifprov~=0))) .^2)) .^0.5)/length(testpointprov) 1)];
    end
end
ptsortdistsort=sortrows (disttestpointrif,1);
distanzespazioeuclidsort=sortrows ([distanzedaipuntidiriferimento'
(1:size (POINTSriferimento,1))'],1);
coordKpuntivicini=[];
for i=1:K
    [Lia,LoCb] =
ismember (POINTSPROV (ptsortdistsort (i,2),:),POINTSriferimento,'rows');
    coordKpuntivicini=[coordKpuntivicini; POINTSPROV (ptsortdistsort (i,2),:)
find (distanzespazioeuclidsort (:,2)==LoCb)
distanzespazioeuclidsort (find (distanzespazioeuclidsort (:,2)==LoCb),1)
IDPOINTSPROV (ptsortdistsort (i,2),:)]];
end
patternKpuntivicini=[];
for i=1:K
    patternKpuntivicini=[patternKpuntivicini;
RMMEANriferimento (ptsortdistsort (i,2),:)]];
end

```

8.3 Modulo statistiche errore con algoritmo K-Means

Di seguito il codice che simula la localizzazione su tutti i punti di test utilizzando l'algoritmo KMEANS in matlab. Successivamente al calcolo, verifica l'effettivo scostamento della posizione stimata da quella reale e calcola l'errore medio e la deviazione standard dell'errore:

```

K=10
DV=[];
IDXV=[];
for i=1:size(testpoints,1)
distanzedaipuntidiriferimento=pdist2 (testpoints (i,1:2),POINTSriferimento);
distanzespazioeuclidsort=sortrows ([distanzedaipuntidiriferimento'
(1:size (POINTSriferimento,1))'],1);
DV (i,:) =distanzespazioeuclidsort (1:K,1)';
IDXV (i,:) =distanzespazioeuclidsort (1:K,2)';
end

TESTPOINTS=[];
POINTSPREDICTED=[];

K=10;
for i=1:size (testpoints,1)
    i
    IDXvicinitrovati=[];
    IDX = kmeans (RMMEANriferimento,2);
    [IDXvicinoiniz,D] = knnsearch (RMMEANriferimento,testpoints (:,3:end),'K',1);
    RMMEANriferimento2=RMMEANriferimento;
    POINTSriferimento2=POINTSriferimento;
    vicini=find (IDX==IDX (IDXvicinoiniz (i)));
    while size (vicini,1)>=K
        RMMEANriferimento2=RMMEANriferimento2 (vicini,:);
        POINTSriferimento2=POINTSriferimento2 (vicini,:);
    end
end

```

```

        IDX = kmeans(RMMEANriferimento2,2);
        [IDXvicino,D] = knnsearch(RMMEANriferimento2,testpoints(i,3:end),'K',1);
        vicini=find(IDX==IDX(IDXvicino));
    end
    [Lia,LoCb] =ismember(POINTSriferimento,POINTSriferimento2(vicini,:), 'rows');
    IDXvicinitrovati=[IDXvicinitrovati;(find(Lia==1))'];
    TESTPOINTS=[TESTPOINTS;testpoints(i,1:2)];
    if size(vicini,1)==1
        POINTSPREDICTED=[POINTSPREDICTED;
POINTSriferimento(IDXvicinitrovati,1:2)];
    else
        POINTSPREDICTED=[POINTSPREDICTED;
mean(POINTSriferimento(IDXvicinitrovati,1:2))];
    end
end
errorsKNN=diag(pdistsq(TESTPOINTS,POINTSPREDICTED));
mean(errorsKNN)
std(errorsKNN)

```

16 giugno 2014.m

ExtractPointsAreaTestPorteChiuse_2014_05_30

```

RMMEAN=RMMEANriferimento;
POINTS=POINTSriferimento;
N=10;
K=10;

```

```

conn = database(datasourcename,user,password,driver,databaseurl);

```

```

POINTSSTIMATIKNN=[];
tablename = 'errori_localizzazione';
colnames={'fk_id_esperimento','id_rilevazione_test','errore','id_rilevazione_rif
1','id_rilevazione_rif2','id_rilevazione_rif3',...

```

```

'id_rilevazione_rif4','id_rilevazione_rif5','id_rilevazione_rif6','id_rilevazion
e_rif7','id_rilevazione_rif8','id_rilevazione_rif9','id_rilevazione_rif10'};

```

```

tic
fprintf(1,'testpoint:      ');
errore = [];
MAX_RSSI_DEV_CUTOFF = max(max(RMDEVriferimento));
MIN_RSSI_DEV_CUTOFF = min(min(RMDEVriferimento));
while (MAX_RSSI_DEV_CUTOFF-MIN_RSSI_DEV_CUTOFF)>0.001

```

```

nzeri=sum((RMDEVriferimento<(MAX_RSSI_DEV_CUTOFF+MIN_RSSI_DEV_CUTOFF)/2))==0,2)
;

```

```

    if max(nzeri)<=2
        MAX_RSSI_DEV_CUTOFF = (MAX_RSSI_DEV_CUTOFF+MIN_RSSI_DEV_CUTOFF)/2;
    else
        MIN_RSSI_DEV_CUTOFF = (MAX_RSSI_DEV_CUTOFF+MIN_RSSI_DEV_CUTOFF)/2;
    end
end

```

```

RSSI_DEV_CUTOFF = (MAX_RSSI_DEV_CUTOFF+MIN_RSSI_DEV_CUTOFF)/2

```

```

for i=1:size(testpoints,1)
    IDXvicinitrovati=[];
    IDX = kmeans(RMMEANriferimento,2);
    [coordKpuntivicini,patternKpuntivicini] =
Kpuntiviciniadunpuntoditest4(RMMEANriferimento,RMDEVriferimento,POINTSriferiment
o,testpoints(i,:),N,-100, RSSI_DEV_CUTOFF);
    [Lia,LoCb] =ismember(POINTSriferimento,coordKpuntivicini(1,1:2), 'rows');
    vicini=find(IDX==IDX((find(Lia==1))));

```

```

RMMEANriferimento2=RMMEANriferimento;
RMDEVriferimento2= RMDEVriferimento;
POINTSriferimento2=POINTSriferimento;
while size(vicini,1)>=K
    RMMEANriferimento2=RMMEANriferimento2(vicini,:);
    RMDEVriferimento2=RMDEVriferimento2(vicini,:);
    POINTSriferimento2=POINTSriferimento2(vicini,:);
    IDX=kmeans(RMMEANriferimento2,2
[coordKpuntivicini,patternKpuntivicini] =
Kpuntiviciniadunpuntoditest4(RMMEANriferimento2,RMDEVriferimento2,POINTSriferime
nto2,testpoints(i,:),N,-100, RSSI_DEV_CUTOFF);
    [Lia,LoCb]
=ismember(POINTSriferimento2,coordKpuntivicini(1,1:2),'rows');
    vicini=find(IDX==IDX((find(Lia==1))));
end
    [Lia,LoCb] =ismember(POINTSriferimento,POINTSriferimento2(vicini,:),'rows');
    IDXvicinitrovati=[IDXvicinitrovati;(find(Lia==1))'];
    sumX=mean(POINTSriferimento(IDXvicinitrovati,1));
    sumY=mean(POINTSriferimento(IDXvicinitrovati,2));
    POINTSSTIMATIKNN=[POINTSSTIMATIKNN;sumX sumY];
    fprintf(1,'\b\b\b\b%4i',i);
end
toc
errorsKNN=diag(pdist2(testpoints(:,1:2),POINTSSTIMATIKNN));
meanerrorKNN=mean(errorsKNN);
stderrorKNN=std(errorsKNN);

close(conn);
fprintf(1,'\n%s','Elaborazione completata');

```

8.4 Modulo statistiche errore con algoritmo di apprendimento del rank

Di seguito il codice che simula la localizzazione su tutti i punti di test utilizzando l'algoritmo di apprendimento in matlab. Successivamente al calcolo, verifica l'effettivo scostamento della posizione stimata da quella reale e calcola l'errore medio e la deviazione standard dell'errore:

```

i=1;
indicitestpoints=[];
while i<=size(testpoints,1)
    indicitestpoints=[indicitestpoints;i];
    ifin=find(ismember(testpoints(:,1:2),testpoints(i,1:2),'rows')==1,1,'last');
    if (i==ifin)&(ifin~=size(testpoints,1))

i=find(ismember(testpoints(:,1:2),testpoints(i+1,1:2),'rows')==1,1,'last')+1;
    else
        i=i+1;
    end
end
end

training=zeros(size(indicitestpoints,1)*size(RMMEANriferimento,1),(size(testpoint
ts,2)-2)*2);
group=zeros(size(indicitestpoints,1)*size(RMMEANriferimento,1),1);
a=1;

```

```

for i=1:size(indicitestpoints,1)

distanzedaipuntidiriferimento=pdist2(testpoints(indicitestpoints(i),1:2),POINTSr
iferimento);
    distanzespazioeuclidsort=sortrows([distanzedaipuntidiriferimento'
(1:size(POINTSriferimento,1))'],1);
    for l=1:size(RMMEANriferimento,1)
        training(a,:)=[testpoints(indicitestpoints(i),3:end)
RMMEANriferimento(l,:)];
        group(a)=find(distanzespazioeuclidsort(:,2)==1);
        a=a+1;
    end
end
mdl = ClassificationKNN.fit( training,group);

indicitestpoints2=setdiff([1:size(testpoints,1)],indicitestpoints')';
K=10
DV=[];
IDXV=[];
for i=1:size(indicitestpoints2,1)
distanzedaipuntidiriferimento=pdist2(testpoints(indicitestpoints2(i),1:2),POINTS
riferimento);
distanzespazioeuclidsort=sortrows([distanzedaipuntidiriferimento'
(1:size(POINTSriferimento,1))'],1);
DV(i,:)=distanzespazioeuclidsort(1:K,1)';
IDXV(i,:)=distanzespazioeuclidsort(1:K,2)';
end

tic
indicitestpoints2=setdiff([1:size(testpoints,1)],indicitestpoints')';
test=zeros(size(indicitestpoints2,1)*size(RMMEANriferimento,1),(size(testpoints,
2)-2)*2);
grouptest=zeros(size(indicitestpoints2,1)*size(RMMEANriferimento,1),1);
TESTPOINTS=[];
POINTSPREDICTED=[];
a=1;
IDX=zeros(size(indicitestpoints2,1),K);
for i=1:size(indicitestpoints2,1)
    testrif=zeros(size(RMMEANriferimento,1),(size(testpoints,2)-2)*2);
    grouptestrif=zeros(size(RMMEANriferimento,1),1);

distanzedaipuntidiriferimento=pdist2(testpoints(indicitestpoints2(i),1:2),POINTS
riferimento);
    distanzespazioeuclidsort=sortrows([distanzedaipuntidiriferimento'
(1:size(POINTSriferimento,1))'],1);
    for l=1:size(RMMEANriferimento,1)
        test(a,:)=[testpoints(indicitestpoints2(i),3:end)
RMMEANriferimento(l,:)];
        testrif(l,:)=[testpoints(indicitestpoints2(i),3:end)
RMMEANriferimento(l,:)];
        grouptest(a)=distanzespazioeuclidsort(l,2);
        grouptestrif(l)=find(distanzespazioeuclidsort(:,2)==1);
        a=a+1;
    end
    grouppredicted=predict(mdl,testrif);
    for l=1:K
        IDX(i,l)=find(grouppredicted==l);
    end
    %grouppredicted=predict(ens,testrif);
    TESTPOINTS=[TESTPOINTS;testpoints(indicitestpoints2(i),1:2)];
    POINTSPREDICTED=[POINTSPREDICTED;
POINTSriferimento(find(grouppredicted==1),:)];

```

```

end
toc
errorsKNN=diag(pdist2(TESTPOINTS, POINTSPREDICTED));
mean(errorsKNN)
std(errorsKNN)

meanTPS=[];
for K=1:10
TPs=[];
for i=1:size(indicitestpoints2,1)
TP=size(intersect(IDX(i,1:K),IDXV(i,1:K)),2);
TPs=[TPs TP/K];
end
meanTPS=[meanTPS mean(TPs)];
end
meanTPS
figure;plot(meanTPS, '*-')

meanER=[];
stdER=[];
for K=1:10
TESTPOINTS=[];
POINTSPREDICTED=[];
for i=1:size(indicitestpoints2,1)
TESTPOINTS=[TESTPOINTS;testpoints(indicitestpoints2(i),1:2)];
POINTSPREDICTED=[POINTSPREDICTED; mean(POINTSRiferimento(IDX(i,1:K),1))
mean(POINTSRiferimento(IDX(i,1:K),2))];
end
errorsKNN=diag(pdist2(TESTPOINTS, POINTSPREDICTED));
meanER=[meanER mean(errorsKNN)];
stdER=[stdER std(errorsKNN)];
end

```


9 Appendice E: sorgente documentato del codice dei moduli della applicazione di monitoraggio network-side

Il codice sviluppato per l'applicazione di monitoraggio network side è stato scritto in java per poter essere eseguito nelle alix su cui è installato un sistema operativo linux ubuntu 11.10 e una versione di jdk v1.7.

I moduli utilizzati nell'applicazione riguardano

- avvio di un thread che richiama la funzione iwlist sull'interfaccia desiderata ed effettua il parsing dei dati di interesse
- aggiornamento della statistica
- invio dei dati al database in real time

Per brevità si riporta la struttura del file di configurazione che è necessario passare al jar eseguibile e uno dei metodi dell'applicazione che esegue la scansione ed effettua il parsing dei dati

File: config.txt

```
# #####
#
#   ###                               Network Side (client alix)           ###
#
#   ###                               ###
#
#   ###                               GEOPOS PROJECT                       ###
#
#   #####
#
#
# configuration file for alix device
#
#alix id
alix-name = alix-geopos-1
wlanCard= BCM4318 N°1
mac= 00:03:c9:df:94:6a
#
#interface parameters
chHopping= [1,6,11]
interfaceMonitor=mon0
# tempo tra due cambiamenti di canale in millisecondi
timeHopping=2000
#
#tcpdump configuration
tcpCommand= "tcpdump -i mon0 -e link[0]==0x80"
```

```

#mac filter
filterMac= [00:0f:b5:99:28:47,00:03:c9:d3:4a:c1,00:90:96:4c:f5:a0]
FlagfilterMac=false

# network configuration
ipServer=127.0.0.1
port=1234

# communication type
socket=false
trasmission=UDP

# communication database
#urlHttp=http://localhost/postdb.php
urlHttp=http://194.119.214.82/alix.php
queryDb="timestamp=&sourceMACAddress=&detectedMACAddress=&dev_std=&RSSI=&nsampl"
database=true

channel=1

#nel caso di scansione semplice
simpleScan=true
iwListCommand= iwlist wlan0 scan
#delay tra una scansioni successive in secondi
delayScansion=10

# finestra per la media mobile
winSimpleScan=5

#logga i dati relativi alle scansioni
logFile=false

```

metodo: executeCommand (Package: com.NetworkSide.alixClient Classe: ExecuteShellCmd);

```

private String executeCommand(String command) {

    UDPClient clientSendUdp=null;
    TCPClient clientSendTcp=null;
    StringBuffer output = new StringBuffer();

    if(socket){
        clientSendUdp=new UDPClient(ip,port);
        clientSendTcp=new TCPClient(ip,port);
    }

    Process p;
    try {
        p = Runtime.getRuntime().exec(command);
        BufferedReader reader =
            new BufferedReader(new
InputStreamReader(p.getInputStream()));

        String line = "";
        String totalScansion="";

```

```

while ((line = reader.readLine()) != null) {

//          Cell 01 - Address: 5C:96:9D:6A:56:E7
//          Channel:1
//          Frequency:2.412 GHz (Channel 1)
//          Quality=54/70 Signal level=-56 dBm
//          Encryption key:on
//          ESSID:"Network Wi-Fi di machi"

//intercetta la cell e cicla per recuperare
i dati di interesse
line.contains("Address:")){

    String quality="";
    String rssi="";
    String channel="";
    String essid="";
    String mac="";

    String temp=line.trim();
    String[] parts = temp.split(" ");
    mac=parts[4];

    //nell ipotesi di filtraggio dei mac address
    verifica se incluso o meno nella lista
    //del file di config

    boolean macMatch=false;

    //verifica se è attivo il filtro sul mac address
    e se c'è il match
    if (filterMacFlag &&
filterMacData.contains(mac.toLowerCase())){

        macMatch=true;

    }

    //viene eseguita se c'è il filtro sui mac address e
    contestualmente c'è stato un matching
    if ((!filterMacFlag) || macMatch){

        for (int i = 0; i < 5; i++) {

            line = reader.readLine();

            if (line.contains("Frequency") &&

```

```

line.contains("Channel")){
                                String temp1=line.trim();
                                String[] parts1 = temp1.split(" ");
                                String channel1=parts1[0];

                                channel =channel1.substring(10,
channel1.length());
                                }

                                if (line.contains("Quality=") &&
line.contains("Signal level")){
                                String temp1=line.trim();
                                String[] parts1 = temp1.split(" ");
                                String quality1=parts1[0];
                                String rssil=parts1[3];

                                quality =quality1.substring(8,
quality1.length());
                                rssi =rssil.substring(6, rssil.length());

                                }

                                if (line.contains("ESSID:")){
                                String temp1=line.trim();

                                essid =temp1.substring(6, temp1.length());

                                }

                                } //fine for

                                //dati grezzi relativi alle singole scansioni
                                String out =getTime() + " " +getAlixName()+"
"+getMacWlan()+ " "+essid+ " "+channel+ " "+mac+ " "+rssi+ " "+quality +"\n";
                                //dati aggregati relativi alle singole scansioni
                                totalScansion+=out;

                                String
dataTime=logData.getData()+"_"+logData.getTime();

                                //invio dei dati grezzi relativi alle singole
scansioni al database
                                if (database){
                                String
queryDb=partsQueryDb[0]+"="+dataTime+partsQueryDb[1]+"="+getMacWlan()+partsQuery
Db[2]+"="+mac+partsQueryDb[3]+"="+null+partsQueryDb[4]+"="+rssi+partsQueryDb[5
1]+"="+null";

                                System.out.println(queryDb);
                                dbClient.executeQueryDb(urlHttp,
queryDb);

                                }

                                if (statisticalAnalysis){

```

```

        //inizializza l'hashmap mettendo come key il mac
address e come dati i rispettivi valori di rssi posti in un array list
        String keyMac=mac;
        int RSS_Value=Integer.valueOf(rssi);
        Object key=
HashMapValueApPassiveScanning.get(keyMac);

        if (key==null) {
i
        //inizializza l'array list che conterrà i valori
di rssi
        rssiValuesPassiveScanning = new
ArrayList<Integer>();
        //aggiunge all'arraylist il valore di rssi
rssiValuesPassiveScanning.add(RSS_Value);
        //inserisce valore di mac address e rssi
HashMapValueApPassiveScanning.put(keyMac,
rssiValuesPassiveScanning);

        continue;
}

        //prende l'array list di rssi associato a quel
specifico mac address
        rssiValuesPassiveScanning =
HashMapValueApPassiveScanning.get(keyMac);

        int sizeDataRSSI=
rssiValuesPassiveScanning.size();
        if (sizeDataRSSI<=winSimpleScan-1) {
rssiValuesPassiveScanning.add(RSS_Value);
        }
        else{
rssiValuesPassiveScanning.remove(0);
rssiValuesPassiveScanning.add(RSS_Value);
        }

        } //fine if (statisticalAnalysis){

        }

        } //fine if cell

    } //fine while

System.out.println(totalScansion);

if (logFile){
    //salvataggio dei dati in un file
    logData.saveDataLog(logFileSample,totalScansion);
}

//INVIO DI DATI AGGREGATI
if(socket){

```

```

        //manda i dati al server mediante tcp o udp
        if (UdpTrasmission){
            clientSendUdp.SendDataUdp(totalScansion);}
        if (TcpTrasmission){

clientSendTcp.SendDataTcp(totalScansion);}

    }

    if (statisticalAnalysis){

        //parte relativa al calcolo delle medie e delle varianze
        degli rssi
        Set set=HashMapValueApPassiveScanning.keySet();
        Iterator iter = set.iterator ();

            while ( iter.hasNext () ) {

                //crea un arraylist dove mette i
                valori di rssi associati ad uno specifico mac
                ArrayList<Integer> rssiValues=new
                ArrayList<Integer>();

                //la chiave selezionata, nonchè
                il mac address
                Object key = iter.next ();

                rssiValues=HashMapValueApPassiveScanning.get(key);
                //

                String mac=key.toString();
                int sample=rssiValues.size();

                //

                double mean=
                Statistical.getMean(rssiValues);
                double
                deviation=Statistical.getDeviation(rssiValues);
                //
                String
                devStd=String.format("%.2f", deviation);
                String
                meanRssi=String.format("%.2f", mean);

                String
                dataTime=logData.getData()+"_"+logData.getTime();

                String
                logPassiveScanning= "TIME: "+dataTime +" , MAC: "+key.toString()+ " , RSSI:
                "+String.valueOf(rssiValues)+" , "+ meanRssi+ " , "+devStd+ " , "+ sample+ "\n" ;

                if (logFile){

//salvataggio dei dati in un file

                logData.saveDataLog(logFileMean,logPassiveScanning);

```

```
}
```

```
System.out.println (logPassiveScanning);
```

```
query al database //prepara la  
interazione con un database mediante comunicazione http //prevede un
```

```
if
```

```
(database){
```

```
String  
queryDb=partsQueryDb[0]+"="+dateTime+partsQueryDb[1]+"="+getMacWlan()+partsQuery  
Db[2]+"="+key.toString()+partsQueryDb[3]+"="+devStd+partsQueryDb[4]+"="+meanRssi  
+partsQueryDb[5]+"="+sample;
```

```
System.out.println(queryDb);
```

```
dbClient.executeQueryDb(urlHttp, queryDb);
```

```
}
```

```
} //while (
```

```
iter.hasNext () ) {
```

```
    } //fine if statisticalAnalysis
```

```
    } catch (Exception e) {  
        e.printStackTrace();  
    }
```

```
    return output.toString();
```

```
}
```

10 Bibliografia

- [1] Rapporto Tecnico ICAR 14/01 Deliverable Geopos 1
- [2] An explanation for unexpected 802.11 Outdoor Link-level Measurement Results
Domenico Giustiniano, Giuseppe Bianchi, Luca Scalia and Ilenia Tinnirello Università degli Studi di Roma - Tor Vergata, Dipartimento di Ingegneria Elettronica, Italy Università di Palermo, Dipartimento di Ingegneria Elettrica, Elettronica e delle Telecomunicazioni, Italy
- [3] Are 802.11 link quality broadcast measurements always reliable?
Domenico Giustiniano - Università di Roma "Tor Vergata", Italy
Giuseppe Bianchi - Università di Roma "Tor Vergata", Italy
- [4] A Theoretical Analysis of Wireless Localization Using RF-based Fingerprint Matching
Jie Yang, Yingying Chen - Department of Electrical and Computer Engineering Stevens Institute of Technology
Hoboken, NJ 07030
- [5] P. Bahl and V. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in Proc. of Infocom'2000, Tel Aviv, Israel, Mar. 2000, vol. 2, pp. 775–584
- [6] CASTRO, P., CHIU, P., KREMENEK, T., AND MUNTZ, R. A, "Probabilistic Location Service for Wireless Network Environments". Ubiquitous Computing 2001 (September 2001)
- [7] Roos, T., Myllymaki, P., Tirri, H., Misikangas, P. and Sievanen, J., 2002. "A probabilistic approach to WLAN user location estimation". International Journal of Wireless Information Networks, 7 (3), 155-163
- [8] Ramadurai, V. & Sichitiu, M. L. (2003), Localization in Wireless Sensor Networks: A Probabilistic Approach., in Weihua Zhuang; Chi-Hsiang Yeh; Olaf Droegehorn; C.-T. Toh & Hamid R. Arabnia, ed., 'International Conference on Wireless Networks', CSREA Press, pp. 275-281
- [9] Youssef, M.A., Agrawala, A. and Shankar, U., 2003. WLAN location determination via clustering and probability distributions. First International Conference on Pervasive Computing and Communications, 143-150
- [10] Porifido, Maria Bianca. Localizzazione device-free in ambiente indoor con reti di sensori ZigBee, tesi di laurea magistrale, Università di Pisa, Facoltà di Ingegneria, Corso di Laurea in Ingegneria delle Telecomunicazioni
- [11] Ekahau, 2010. <http://www.ekahau.com>, April 2010.
- [12] Mirowski, P., Milioris, D., Whiting, P. and Kam Ho, T. (2014), Probabilistic Radio-Frequency Fingerprinting and Localization on the Run. Bell Labs Tech. J., 18: 111–133
- [13] Charalampos Papamanthou, Franco P. Preparata, and Roberto Tamassia. 2008. Algorithms for Location Estimation Based on RSSI Sampling. In Algorithmic Aspects of Wireless Sensor

Networks, Indor P. Fekete (Ed.). Lecture Notes In Computer Science, Vol. 5389. Springer-Verlag, Berlin, Heidelberg 72-86

[14] Youssef, M., 2004. HORUS: A WLAN-Based indoor location determination system, Ph.D. Dissertation, University of Maryland, 2004

[15] King, T., Kopf, S., Haenselmann, T., Lubberger, C., Effelsberg, W., 2006. COMPASS: A Probabilistic Indoor Positioning System Based on 802.11 and Digital Compasses, 1st WiNTECH, Sept 2006, 34-40

[16] A Practical Path Loss Model For Indoor WiFi Positioning Enhancement
Atreyi Bose and Chuan Heng Foh School of Computer Engineering Nanyang Technological University Singapore

[17] DorFin: WiFi Fingerprint-based Localization Revisited
Chenshu Wu , Zheng Yang, Zimu Zhou, Yunhao Liu and Mingyan Liu School of Software and TNList, Tsinghua University CSE, Hong Kong University of Science & Technology Department of EECS, University of Michigan

[18] The Airplace Indoor Positioning Platform for Android Smartphones: C. Laoudias, G. Constantinou, M. Constantinides, S. Nicolaou, D. Zeinalipour-Yazti and C. G. Panayiotou
KIOS Research Center for Intelligent Systems and Networks, University of Cyprus
Department of Computer Science, University of Cyprus

[19] Cisco Wireless Control System Configuration Guide
Cisco Systems, Inc.

[20] Wi-Fi Based Positioning Algorithm for Indoor Environment using 802.11 Standards
International Journal of Scientific & Engineering Research, Volume 5, Issue 3, March-2014
Sikandar Khan, Adeel Akram, Saeeda Usman