

Localizzazione tramite una rete di sensori *wireless* utilizzando tecniche non-parametriche.

Damiano Picco, Marco Prendin, Damiano Vanti

Abstract — Questo documento è la relazione finale del progetto svolto per il corso di Progettazione di Sistemi di Controllo tenuto dal prof. L. Schenato dell'Università di Padova.

In questo progetto viene sviluppato il problema della localizzazione di un nodo mobile tramite una rete di sensori wireless in modo distribuito e con tecniche non parametriche. Un approccio Bayesiano e in particolare l'utilizzo della teoria sui processi Gaussiani ha portato allo sviluppo di un algoritmo in grado di stimare la posizione di un nodo mobile grazie alla conoscenza della potenza del segnale radio ricevuto dallo stesso da parte di nodi fissi preposti nei due ambienti che sono stati analizzati: i laboratori dell'Istituto Nazionale di Fisica Nucleare (INFN) di Legnaro e i laboratori della sede di Ingegneria dell'Informazione (DEI) di Padova.

L'algoritmo è stato poi riformulato sia in forma centralizzata che in forma decentralizzata, andando anche a sviluppare tecniche multi-hop.

Index Terms — Localizzazione, Bayesiano, WSN, Processo Gaussiano, inferenza, covarianza, multi-hop, connettività, centralizzato, decentralizzato.

I. INTRODUZIONE

L'obiettivo del progetto è la determinazione della posizione, all'interno di una rete di sensori wireless (Wireless Sensor Network, WSN), di un nodo mobile tramite la conoscenza della misura della potenza (Received Signal Strength, RSS) del segnale emesso da sensori fissi preposti nell'ambiente (nodi ancora) e grazie anche alla conoscenza di un training set di misure di potenza-posizioni effettuate in precedenza. Dal punto di vista teorico l'obiettivo è quello di estrapolare un modello continuo da un set di dati che sia un sottoinsieme dei dati sul quale vogliamo

applicare il modello; si tratta perciò di un procedimento di tipo induttivo, in quanto si cerca di ricostruire un modello generale basandosi su un insieme limitato di dati sperimentali. Come sarà approfondito in dettaglio nel seguito, in questo progetto si farà uso della teoria sui processi Gaussiani, in quanto i calcoli richiesti per l'inferenza statistica diventano particolarmente semplici.

Questo progetto è nato dall'idea di migliorare i risultati ottenuti in precedenti lavori, in particolare con riferimento al progetto sviluppato nei laboratori dell'Istituto Nazionale di Fisica Nucleare (INFN) di Legnaro [3] nato dalla necessità di instaurare all'interno dei laboratori una rete che permettesse, in caso di emergenza, la localizzazione ed il tracking in tempo reale degli operatori di soccorso presenti nei locali colpiti, in modo da poterne guidare le mosse dall'esterno e facilitarli nell'intervento.

Il progetto è stato in seguito esteso ai laboratori al piano terra della sede di Ingegneria dell'informazione dell'Università di Padova (DEI), in particolare si è cercato di sfruttare tutta la rete già cablata presente nei laboratori nonostante fossero stati già intrapresi studi sulla connettività e sul numero minimo di sensori necessari per una buona prestazione (in termini di connettività) di tale rete [4]; studi che potranno essere ripresi in uno sviluppo futuro per la realizzazione di una rete efficiente e a minor costo.

Lo stato dell'arte nella localizzazione e tracking riporta, come sarà descritto in maggior dettaglio nella prossima sezione, svariati metodi e strategie per la soluzione di tale problema. I metodi utilizzati lungo tutto il progetto fanno uso, come già introdotto, di una mappa di potenza RSSI del segnale emesso e ricevuto rispettivamente dai nodi ancora e dal nodo mobile, diversificandosi a seconda della quantità di informazione utilizzata nella stima della posizione del nodo mobile da parte dei vari nodi ancora (saranno sviluppati algoritmi sia centralizzati che decentralizzati e in quest'ultimi si illustreranno anche tecniche multi-hop) e di come il nodo mobile "richiede" la propria posizione.

Questo documento è organizzato nel modo seguente. Nella Sezione II è contenuta una rassegna di alcuni metodi di localizzazione e *tracking* basati su reti wireless. La Sezione III è lasciata ai posteri.

II .STATO DELL'ARTE

Qualsiasi sistema di navigazione e di tracking ha come presupposto principale la localizzazione dell'oggetto di interesse; è bene ricordare che l'ambito al quale ci riferiremo d'ora in poi sono le WSN (*Wireless Sensor Network*) cioè reti senza fili formate da un insieme molto numeroso di dispositivi autonomi, detti nodi o mote, in grado di comunicare via radio tra loro e spesso dotati di sensori (come ad esempio per la temperatura e umidità) e in grado di instaurare delle connessioni ben distribuite e capillari con l'obiettivo di monitorare attentamente un'area territoriale.

Le WSN sono caratterizzate da un imponente numero di sensori, con una direzione del traffico dati non precisa; normalmente ogni nodo spedisce pacchetti a tutti i dispositivi in ascolto, e non ad alcuni soltanto.

Notevole importanza riveste l'aspetto energetico: utilizzandoli come sensori e posizionandoli anche in luoghi non facilmente o frequentemente raggiungibili, vengono spesso alimentati a batterie e diventa quindi necessario assicurare loro un'autonomia discreta in base al servizio richiesto, per evitare onerosi tempi di sostituzione (nel caso il numero di sensori sia elevato) o lunghi spostamenti per raggiungerli tutti.

In sintesi possiamo elencare le caratteristiche fondamentali che distinguono una rete WSN:

- nodi di dimensioni estremamente contenute e dalle risorse di memoria e calcolo limitate;
- interazione con l'ambiente circostante tramite sensori;
- scambio di dati mediante rete wireless;
- lunga autonomia energetica (dell'ordine di mesi o anni);
- contesti operativi eterogenei ed ostili, anche inaccessibili all'uomo;
- topologia dinamica di rete;
- funzionamento autonomo, non costantemente monitorato, e conseguente politica di gestione dei guasti.

Ritornando al nostro problema, in letteratura le tecniche di localizzazione sono classificate secondo svariati criteri, come ad esempio in base alla grandezza fisica utilizzata: onde radio, impulsi sonori, infrarossi; o in base al parametro che fornisce le informazioni tra i nodi, in particolare si distinguono:

- *Angle of Arrival AoA*: basato sulla stima dell'angolo relativo tra due nodi;
- *Time of Arrival ToA*: basato sul tempo impiegato dal segnale radio per propagarsi da un nodo ad un altro;
- *Time Difference of Arrival TDoA*: basato sull'intervallo di tempo tra l'arrivo di un segnale radio e un segnale ad ultrasuoni emesso da un nodo;
- *Received Signal Strength Indicator RSSI*: basato sulla potenza ricevuta-inviata dai diversi nodi.

Le prime tre tecniche presentate richiedono dispositivi specifici per la loro implementazione: ad esempio dispositivi ad ultrasuoni, dispositivi di sincronizzazione, telecamere ecc.

La loro realizzazione ha portato interessanti risultati, si ricordi come esempio il progetto Cricket [], dove per localizzare un nodo mobile il sistema misura la differenza fra il tempo di volo del segnale elettromagnetico ed il segnale ultrasonico prodotto dal nodo mobile.

La precisione ottenuta con questa tecnica di localizzazione è ragguardevole, in quanto compresa nell'ordine del centimetro, ma si è verificato che i componenti preposti alla generazione degli impulsi ultrasonici sono molto sensibili alle vibrazioni meccaniche e che quindi, l'inserimento di tali attuatori nei nodi mobili soggetti a spostamenti e a continue vibrazioni meccaniche, causa la rottura degli stessi con la conseguente inutilizzabilità del nodo mobile.

Questo e altri motivi legati al costo hanno provveduto ad assicurare ai metodi basati sulla RSSI una maggiore popolarità, malgrado i risultati ottenuti con tali dispositivi non siano altrettanto allettanti di quelli ottenuti con le già citate tecniche.

Le tecniche di localizzazione basate sull'RSSI si distinguono in due classi principali:

- modello di Canale: l'algoritmo stima la distanza relativa del nodo mobile dai nodi ancora e con tecniche di triangolazione individua la posizione dello stesso.
- mappa di Potenza RSSI: ogni nodo ancora crea una mappa di potenza che può essere ottenuta attraverso un modello analitico sulla propagazione del segnale radio oppure da una serie di misure di potenza raccolte in diverse posizioni (learning measure) così da permettere un confronto con la misura test. I principali sistemi sviluppati in questa categoria sono RADAR [] e MoteTrack [] che andiamo a descriveremo brevemente in seguito.

Il progetto RADAR [] opera in due fasi:

- 1) *fase off-line*, in cui si effettua una raccolta dati che descriva la distribuzione nell'ambiente della potenza del segnale emesso dai nodi `ancora (*Received Signal Strength, RSS*);
- 2) *fase on-line*, in cui le misure rilevate vengono confrontate con i dati a priori ed elaborate per di determinare la posizione del nodo mobile.

La prima fase viene attuata in due modi: effettuando una serie di misurazioni sul campo oppure basandosi su un modello di propagazione del segnale radio. Nel secondo caso tipicamente il modello è il seguente:

$$P(d) = P(d_0) - 10n_p \log\left(\frac{d}{d_0}\right) - C WAF$$

dove nella formula $P(d_0)$ rappresenta il valore della potenza del segnale elettromagnetico che il nodo rileva ad una distanza d_0 dal nodo ancora che genera il segnale elettromagnetico.

Tale modello prevede inoltre la stima sul campo di due parametri: n_p e WAF.

Il primo rappresenta il fattore secondo cui decresce la potenza del segnale elettromagnetico all'aumentare della distanza, mentre il secondo tiene conto dell'attenuazione di potenza del campo elettromagnetico causata dal passaggio delle onde attraverso i muri (*Wall Attenuation Factor*).

Anche in presenza di ambienti complessi in cui vi siano molte interferenze agenti sulla propagazione del segnale radio, il sistema RADAR permette di

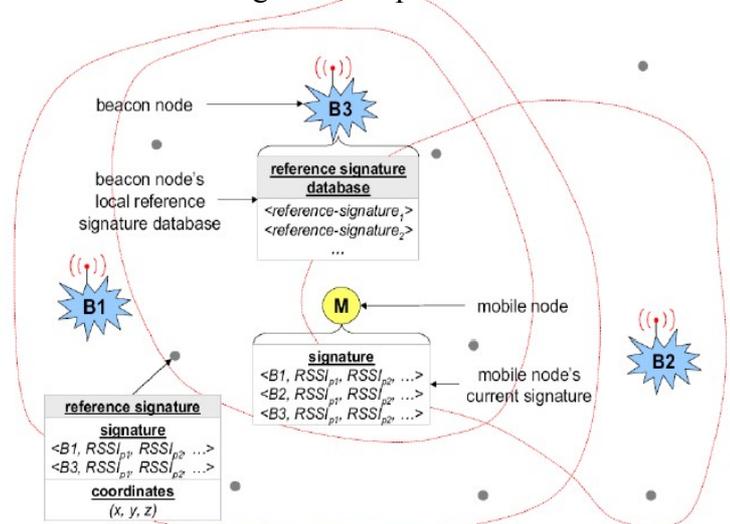
avere un notevole grado di accuratezza (lo scarto medio si aggira sui 2/3 metri).

Il progetto MoteTrack è un sistema di localizzazione decentralizzato, basato sulla misura della potenza del segnale elettromagnetico prodotto da alcuni nodi ancora. Per determinare la posizione di un nodo mobile il sistema utilizza una mappa, creata attraverso la raccolta di misure svolte a priori sul campo, che associa ogni punto del piano al valore di potenza del campo elettromagnetico che un nodo mobile rileverebbe in quella coordinata. MoteTrack non utilizza quindi alcun tipo di modello che legghi la potenza del segnale ricevuto (RSS, Received Signal Strength) ad un qualche valore di distanza.

L'insieme dei valori di RSS rilevati da un nodo mobile, unitamente all'identificativo (ID) del nodo ancora che ha generato il campo elettromagnetico a cui si riferiscono i dati, è detta signature. L'insieme di una signature e delle coordinate spaziali in cui essa è stata raccolta è detto invece reference signature. L'insieme delle reference signature raccolte sono memorizzate fra i nodi della rete in modo da minimizzare l'occupazione di memoria per il singolo nodo e contemporaneamente massimizzare la ridondanza dei dati.

Durante la fase di localizzazione, un nodo mobile M, dopo aver raccolto una propria signature, utilizza le reference signature memorizzate nei nodi ancora da esso rilevabili per stimare la propria posizione. La posizione del nodo è dunque determinata attraverso una media pesata delle distanze dalle coordinate contenute nelle reference signature utilizzate e, i pesi utilizzati nella media, sono proporzionali alle distanze fra la signature rilevata dal nodo M e quelle contenute nelle reference signature.

In figura è riportato un esempio di raccolta di una signature da parte di un nodo mobile M e l'invio di alcune reference signature da parte dei nodi ancora.



Sia RADAR che MoteTrack sono basati su un modello deterministico della mappa di potenza; nel progetto che andremo a descrivere in dettaglio nella prossima sezione includeremo anche dell'informazione relativa alla varianza delle misure rilevate, in questo modo la posizione del nodo mobile è ottenuta calcolando la maggiore probabilità di una posizione data la potenza ricevuta, quindi sfruttando un approccio di tipo Bayesiano.

III. REGRESSIONE E PROCESSI GAUSSIANI

– INTRODUZIONE

Dal punto di vista teorico, il problema affrontato in questo progetto è l'apprendimento di una mappa di ingresso-uscita da un set di dati empirici (training dataset). In letteratura il problema è stato largamente studiato, prendendo il nome di *regressione* nel caso in cui l'uscita sia continua, e *classificazione* in caso contrario.

Trattando con la regressione, il nostro obiettivo sarà quindi quello di estrapolare un modello continuo che legghi la potenza rilevata dal nodo mobile dai vari sensori ancora (input, x) alla posizione corrente dello stesso (uscita, y) in un sistema di riferimento noto. Chiamando N il numero di nodi ancora preposti nell'ambiente, l'obiettivo sarà quindi quello di ricavare una funzione $g(y)$ tale per cui

$$g(y) = f^{-1}(x) : \mathbb{R}^N \rightarrow \mathbb{R}^2$$

Per fare questo, ricavato sperimentalmente un set di misure ingresso-uscita, divideremo tale set in due insiemi: *training set* grazie al quale estrapoleremo la funzione $g(y)$ e *test set* grazie al quale testeremo l'efficacia della funzione trovata.

L'idea fondamentale, che svilupperemo inizialmente grazie all'inferenza bayesiana e che poi estenderemo grazie ai processi gaussiani, sta nell'assegnare ad ogni possibile funzione, che dato come ingresso la potenza rilevata dal nodo mobile dai vari sensori ancora dia in uscita la posizione corrente dello stesso, una probabilità a priori sulla base di assunzioni che verranno fatte sulle funzioni in gioco (ad esempio la regolarità) andando poi a modificare tale probabilità grazie al training set acquisito sperimentalmente. Appare evidente il problema della presenza di un insieme infinito di funzioni plausibili, problema che osserveremo in un primo momento parlando di inferenza bayesiana e che risolveremo in

un secondo momento estendendo il nostro problema ai processi gaussiani.

Ricordando che un processo gaussiano non è altro che la generalizzazione della densità di probabilità Gaussiana, dove si trattano in questo caso non più scalari o vettori bensì funzioni, non è difficile a questo punto dare un'idea grafica della situazione che andremo a trattare; si pensi per semplicità ad un problema di regressione scalare, dove l'ingresso x produce l'uscita $f(x)$. In figura 1.1(a) sono rappresentate un numero finito di funzioni estratte casualmente dalla distribuzione a priori specificata da un particolare processo Gaussiano (si osservi come tale processo, attraverso la sua *covarianza*, favorisca funzioni regolari). La probabilità a priori rappresenta la nostra idea a priori sul tipo di funzioni che ci aspettiamo di osservare, assumendo che la varianza di tale probabilità non dipenda da x è evidente che la regione di incertezza (rappresentata dalla parte ombreggiata) è quella riportata in figura 1.1(a). Le cose cambiano se sono disponibili osservazioni di coppie di ingresso-uscita, ad esempio siano note $(x_1, y_1), (x_2, y_2)$. Ora vogliamo considerare solo funzioni che passano esattamente per i due punti specificati; tali punti insieme alla probabilità a priori danno luogo a quella che viene chiamata probabilità a posteriori, che porta i risultati visibili in figura 1.1(b).

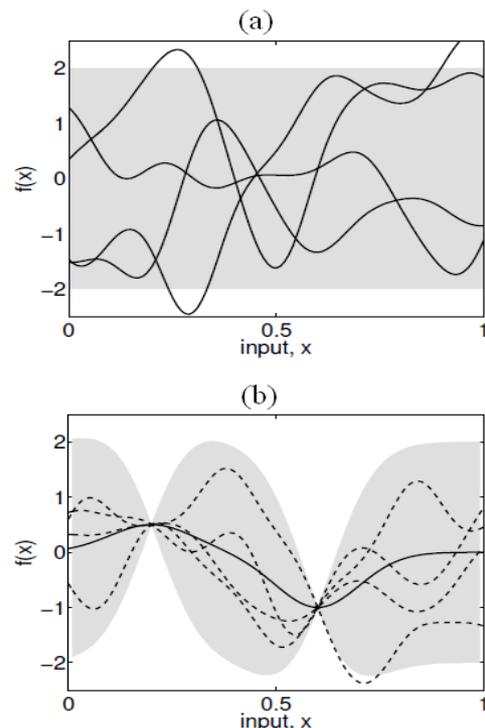


Figura 1.1: Il riquadro (a) mostra quattro funzioni campione estratte dalla probabilità a priori, il riquadro (b) mostra la situazione dopo l'osservazione di due datapoint.

Come studieremo approfonditamente nella parte dedicata ai processi Gaussiani, la specificazione della probabilità a priori è importante perché fissa le proprietà delle funzioni che considereremo per l'inferenza. Sopra abbiamo sfiorato l'argomento parlando della varianza, che comunque non è l'unica caratteristica che può essere specificata e manipolata. Molto importante è osservare che la regolarità e la stazionarietà delle funzioni in figura 1.1(a) è indotta dalla funzione *covarianza* del processo gaussiano generato; come si vedrà in seguito il problema di *learning* nei processi gaussiani si traduce esattamente nel problema di trovare proprietà adatte per le funzioni covarianza.

similmente sui nodi vicini e non modifica dunque la similarità delle misure rilevate dai nodi.

TECNICHE DI LOCALIZZAZIONE BAYESIANE

Come si `e detto in precedenza, MoteTrack basa il proprio funzionamento sull'ipotesi fondamentale che due nodi vicini debbano rilevare delle signature simili fra loro, rilevare cioè, gli stessi valori di potenza di campo elettromagnetico. Tale ipotesi, contrariamente a quanto avviene col modello di campo utilizzato nel progetto RADAR, non impone alcun legame fra la potenza del campo elettromagnetico e la distanza dal trasmettitore a cui esso viene rilevato e nemmeno, richiede che tale valore di campo sia costante per identici valori di distanza dal trasmettitore.

E' chiaro dunque che una tecnica di localizzazione che si basi sulla similarità delle misure di campo da parte dei nodi, sia più robusta ai disturbi dell'ambiente rispetto ad una che consideri ogni misura come un evento indipendente dalle altre misure, dato che, generalmente, il disturbo influisce