



UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI MATEMATICA

Corso di Laurea Magistrale in Informatica

Indirizzo Sistemi

**PATHS: UN SERVIZIO LOCATION-BASED PER
PERCORSI ALTERNATIVI**

**(PATHS: A LOCATION-BASED SERVICE FOR ALTERNATIVE
ROUTES)**

Laureando

Tobia Zorzan

Relatore

Prof. Claudio Palazzi

Co-relatore

Prof.ssa Ombretta Gaggi

ANNO ACCADEMICO 2014/2015

Indice

1	Introduzione	1
2	Contesto e lavori correlati	3
2.1	Web ² e Opportunistic Sensing	3
2.2	Gestione dei dati di movimento	8
2.2.1	Ricostruzione delle traiettorie	10
2.3	Versione precedente e lavori correlati	12
3	PathS Client	15
3.1	Requisiti	15
3.2	Tecnologie	15
3.3	Stato dell'arte	15
4	PathS Server	17
4.1	Requisiti	17
4.2	Componenti	17
4.2.1	Ricezione dei campioni	17
4.2.2	Elaborazione dei campioni	17
4.2.3	Interrogazione	18
4.3	Tecnologie e linee guida	18
5	Ricezione dei campioni	19
5.1	Formato	19
5.2	Implementazione	19
5.3	Esempi	19
6	Elaborazione dei campioni	21
6.1	Servizi di Cartografia	21
6.2	Persistenza ed elaborazione dei dati GIS	21
6.3	Algoritmo di Map Matching	21
6.3.1	ST-MapMatching	22

6.3.2	Pre-elaborazione del percorso - step 1	22
6.3.3	Identificazione dei candidati - step 2	22
6.3.4	Valutazione e selezione dei candidati - step 3	22
7	Routing	23
7.1	Implementazione	23
7.2	Percorsi calcolati	23
7.2.1	Servizio Map Quest	23
7.2.2	Shortest Path	23
7.2.3	Percorsi con <i>label</i>	23
8	Conclusione	25
8.1	Risultati	25
8.2	Miglioramenti ed Evoluzioni	25
	Appendice	27
A	Manuale installazione	29
A.1	Ambiente	29
A.2	Dipendenze	29
A.2.1	Database	29
A.2.2	Java	30
A.2.3	Repository	30
A.3	Configurazione	30
A.4	Avvio	31
A.5	Note	31
	Bibliografia	33

Sommario

I recenti sviluppi dei dispositivi *smartphone*, in particolare l'inclusione di sensori sempre più evoluti, ha portato alla definizione di un nuovo paradigma denominato *Web²* (*Web Squared*). La diffusione capillare di questi strumenti consente di creare una rete di sensori distribuiti in grado di raccogliere una ingente mole di dati. I dati stessi possono essere quindi elaborati e fornire il supporto a nuovi servizi che non si basano sulle sole informazioni *statiche* ma su un contesto *dinamico* modificato dall'interazione degli utenti. Il progetto *PathS* vuole essere un prototipo che segue questo paradigma. La sua componente *client* raccoglie informazioni aggiuntive sui percorsi pedonali percorsi dagli utenti. La parte *server* elabora le informazioni raccolte e mantiene una base di dati in grado di supportare servizi di interrogazione che adottano criteri aggiuntivi a quelli tradizionali. Si espone quindi quali sono stati i principi e le valutazioni eseguite nella realizzazione del componente *PathS Server*, i risultati raggiunti e le possibili aree di miglioramento individuate.

Capitolo 1

Introduzione

W.I.P. Introduzione al progetto. Presentazione dello scopo principale e degli attori coinvolti. Rassegna del contenuto dei capitoli.

Capitolo 2

Contesto e lavori correlati

In questo capitolo si presenta il contesto in cui si è sviluppato il progetto *PathS* ed alcuni concetti fondamentali collegati all'ambito di applicazione. Si citano inoltre progetti con finalità simili e approcci ritenuti interessanti punti di riferimento.

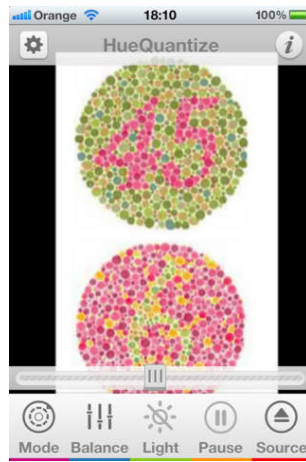
2.1 Web² e Opportunistic Sensing

Una rete di sensori o *Sensor Network* è composta di elementi autonomi e distribuiti atti al monitoraggio di determinati parametri ambientali. Ciascun nodo ha capacità autonome di calcolo, percezione e misurazione dell'ambiente circostante, comunicazione ed eventualmente di mobilità. Molte ricerche si sono focalizzate in questo ambito, in particolare con lo scopo di definire un sistema affidabile e decentralizzato per la comunicazione e il supporto a servizi in questo scenario. Raramente questo tipo di soluzioni sono andate oltre l'ambito della sperimentazione fino all'applicazione su larga scala. Il recente sviluppo tecnologico e la diffusione capillare dei dispositivi *smartphone* ha radicalmente modificato la situazione, aprendo la possibilità ad importanti evoluzioni. Di fatto anche un dispositivo base ha tutte le caratteristiche necessarie, essendo dotato di un sistema GPS, moduli di connessione alla rete dati e numerosi sensori tra cui quello di prossimità, luminosità, microfono, fotocamera e accelerometri. In questo modo si possono coprire ampie zone geografiche con un buon numero di dispositivi, offrendo sia l'occasione di applicare le ricerche sviluppate nell'ambito delle *sensor networks*, sia di esplorare scenari del tutto nuovi. Ciò che rende ancora più interessante l'applicazione di una rete pervasiva di dispositivi, non è solo la possibilità di migliorare i servizi per gli utilizzatori diretti, ma la possibilità di studiare scenari del tutto nuovi a giovamento dell'intera società.

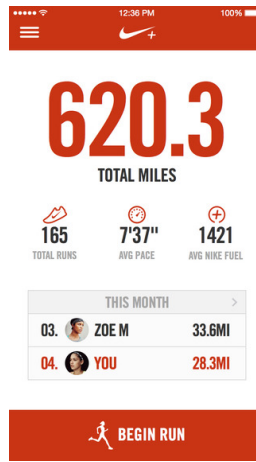
Pensiamo a quella che è già stata una grande evoluzione del mondo di internet, ovvero il *Web 2.0*. Il suo avvento è fatto coincidere con la diffusione dei social networks, alla cui base c'è l'elevato livello di interazione tra il sito Web e gli utenti stessi. In ogni caso è richiesta una interazione diretta dell'utilizzatore, il quale consapevolmente contribuisce ad aumentare la base delle informazioni. Lo sviluppo di alcune tecnologie specifiche (tra cui *AJAX*) ha permesso lo scambio di dati in background fra Web browser e server consentendo l'aggiornamento dinamico della pagina senza esplicito ricaricamento da parte dell'utente. Fondamentalmente si è passati da un Web statico a uno di tipo dinamico e sociale in cui l'utente non è più solo un lettore e un consumatore passivo di contenuti, ma è il principale creatore di essi. La diffusione capillare di questi dispositivi con sensori avanzati, apre la strada al paradigma *Web²* in cui la quantità di dati, raccolta anche senza l'esperienza utente diretta, non aumenta in modo lineare ma letteralmente esplode con andamento esponenziale. Diventa quindi di fondamentale importanza le modalità con cui si raccoglie questa mole di dati e i principi di design con cui si imposta lo sviluppo di applicazioni in questo contesto. Una buona analisi tecnica e un'architettura di riferimento per lo sviluppo di servizi *Web²* basati su reti di sensori *smartphone* è fornita da Calma, Palazzi e Bujari in [1]. Nell'articolo è introdotto e sviluppato il concetto di *opportunistic sensing* il quale è ritornato molto utile anche nel progetto *PathS*. Gli autori espongono due criteri fondamentali secondo cui categorizzare le applicazioni in ambito *Web Squared*: il primo riguarda la scala di rilevamento, il secondo è il grado di coinvolgimento degli utenti. Considerando la scala di rilevamento possiamo distinguere una applicazione in:

- **personale**: progettata per utenti singoli e i risultati vengono visualizzati solo all'utente stesso, senza la loro condivisione. Un esempio di queste può essere *Dankam*, applicazione che aiuta nell'identificazione dei colori le persone affette da daltonismo;
- **di gruppo**: ad esempio *Nike+* o *Runtastic*, applicazioni in cui un gruppo di persone con interessi affini condivide informazioni spesso sfruttando la connessione ai popolari social network;
- **di comunità**: i dati sono raccolti da una vasta gamma di persone e i risultati condivisi pubblicamente per renderli disponibili a tutti gli utenti. Un'esempio di applicazione è Google Maps e la relativa segnalazione dei tratti trafficati.

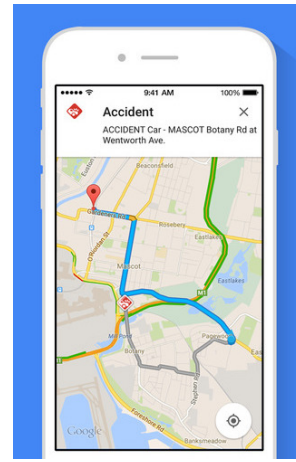
Valutando invece il coinvolgimento con l'utente, le applicazioni di questa categoria si possono suddividere in:



(a) Dankam



(b) Nike+



(c) Google Maps

- **partecipative:** ovvero la raccolta dei dati avviene tramite il coinvolgimento esplicito e diretto dell'utente. Ci si affida quindi all'entusiasmo e alla relazione diretta con gli utilizzatori i quali volontariamente eseguono l'applicazione e raccolgono i dati necessari. Solitamente le informazioni che derivano da questa tipologia sono molto accurate e ben distribuite anche per la possibilità di una opportuna pianificazione. La controparte è che spesso non è facile ottenere una sufficiente base di copertura se non si dispone del bacino di utenza e delle risorse necessarie.
- **opportunistiche:** in cui l'acquisizione delle informazioni avviene tramite i sensori del dispositivo in determinate situazioni di utilizzo, anche se in quel momento non è l'attività principale. La raccolta avviene quindi in *background* informando, ma non richiedendo necessariamente l'attenzione, dell'utente. In qualche modo si "distrae" l'utilizzatore con una attività primaria dall'alto livello di gradimento e nel frattempo si approfitta della situazione per catturare le informazioni necessarie. L'aspetto negativo di questo approccio è principalmente tecnico, dato che la sovrapposizione di più attività comporta un maggiore dispendio di risorse del dispositivo (*cpu*, batteria, rete ...) e al tempo stesso gli algoritmi che devono interpretare i dati provenienti dai sensori e identificare le situazioni idonee sono particolarmente complessi. Questo tipo di approccio risulta particolarmente utile per le applicazioni *di comunità* che, pur raccogliendo pochi campioni, possono contare su un'ampio bacino di utenza.

Nello sviluppo del progetto *PathS* ed in particolare per la progettazione della

componente *client* si è pensato di sfruttare entrambi questi aspetti. Tramite il coinvolgimento diretto di alcuni studenti e il coordinamento del Prof. Palazzi e dei suoi collaboratori, è stata organizzata una campagna di campionamento massivo della zona universitaria e dei dintorni. Questo ha consentito sia di raccogliere numerosi *feedback* riguardo il funzionamento del sistema che di raccogliere una consistente base di informazioni per l'area di riferimento. Per integrare e ampliare questa base di partenza, si è pensato ad un approccio di tipo opportunistico presentato in dettaglio nel capitolo successivo.

Un altro aspetto fondamentale proposto nello stesso articolo è il principio di architettura software secondo cui impostare un servizio basato sul paradigma *Web*². Concepire una applicazione mobile di questo tipo non è banale, in quanto affiancare le due attività di raccolta intensiva di dati dai sensori e la condivisione di essi può portare a diversi problemi. Uno degli aspetti fondamentali riguarda la gestione delle risorse (spesso limitate) del dispositivo. Un utilizzo troppo intenso delle componenti relative ai sensori così come una trasmissione dati frequente può portare ad esaurire in breve tempo la batteria, o in alcuni casi a sovraccaricare l'unità di calcolo del dispositivo, causando un disservizio nelle funzionalità principali (come ricevere le chiamate). Una delle soluzioni proposte è prima di tutto individuare e separare le componenti sulla base del ruolo che devono svolgere per raggiungere il risultato atteso. Nel caso delle applicazioni *Web*² è individuato un sistema architetturale basato su tre *layer* di competenza che sono:

- **sensing layer**: che necessariamente risiede nel dispositivo e si occupa della raccolta dei dati grezzi dai sensori;
- **learning layer**: che svolge il ruolo di processare i dati raccolti e derivarne i risultati;
- **release layer**: che si occupa di fornire i risultati ottenuti all'utente.

Le problematiche che riguardano il primo *layer* riguardano principalmente l'integrazione con gli ambienti di sviluppo (*SDK*) e le modalità di accesso alle componenti del dispositivo. Spesso gli strumenti messi a disposizione degli sviluppatori non sono così a basso livello da consentire un'utilizzo ottimale. E' necessario spesso trovare dei *workaround* specifici, piattaforma per piattaforma, che portino al giusto compromesso tra la frequenza di campionamento, la precisione dei dati raccolti e l'utilizzo delle risorse. Per quanto riguarda il *learning layer* si possono invece trovare le soluzioni più diverse. Non è indispensabile che questa logica applicativa risieda necessariamente nel dispositivo ma è possibile delegare l'attività ad un sistema esterno (es. server cloud) in cui la disponibilità di risorse è maggiore ed è possibile adottare algoritmi più complessi. In alcuni casi si suggerisce anche un approccio

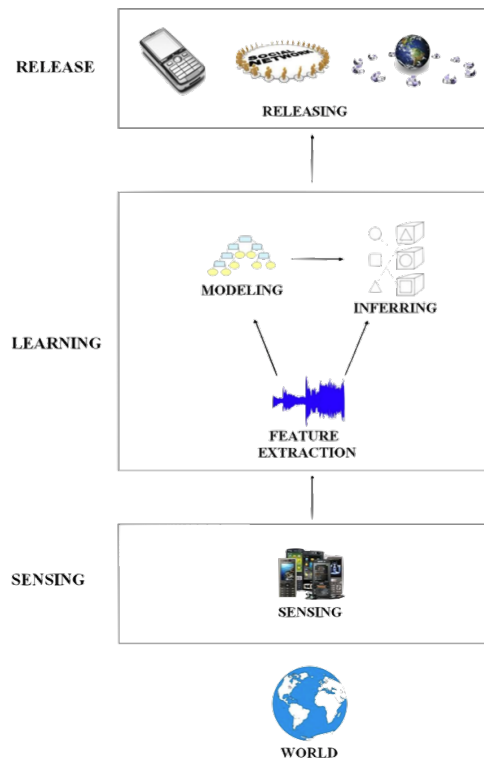


Figura 2.2: Architettura a *layer* di una applicazione Web²

“misto” in cui i dati vengono pre-processati dal dispositivo cercando di estrarre delle *feature* che vengono quindi comunicate all’esterno ed elaborate. Il *release layer* si occupa di mostrare il risultato all’utente finale e, a seconda del tipo di applicazione, potrebbe avere come destinazione sia un singolo cliente così come una vasta comunità. Per questo motivo anche in questo caso le soluzioni proposte possono riguardare sistemi esterni o ad esempio il collegamento a vaste reti di distribuzione tipo *social-network*.

Nel progetto *PathS* ritroviamo tutte queste problematiche e le soluzioni adottate seguono in molti casi i principi generali qui presentati. Ad esempio le modalità di campionamento GPS eseguite dal *client* sono state tarate al fine di ottenere risultati accettabili minimizzando l’uso delle risorse. Dal punto di vista dell’architettura del sistema si è deciso di elaborare tutti i dati nel sistema *server* esterno, dove i dati raccolti vengono processati e utilizzati per derivare le informazioni di sintesi. Infine agli utenti finali si comunicano i risultati sia in forma diretta (con opportuno protocollo di comunicazione con il client) che in forma più estesa tramite l’accesso a dei servizi web.

2.2 Gestione dei dati di movimento

Un'altra tematica affrontata in quanto affine all'ambito del progetto è quella dei *Mobility Data*. Alcuni esempi di *mobility data* sono forniti in [4] da Pelekis e Theodoridis e possono essere:

- i dati provenienti da una conversazione telefonica cellulare (la compagnia telefonica ha informazioni aggiuntive in merito al posizionamento del dispositivo);
- i dati provenienti da un dispositivo GPS durante una determinata attività (la combinazione tra la posizione e il momento in cui è rilevata);
- i dati scambiati tra i veicoli di una *vehicular ad hoc network* (*VANET*);
- i dati raccolti da un sistema di *radio-frequency identification* (*RFID*).

Come vediamo il concetto di *mobility data* è molto vario e può essere applicato a numerosi contesti. Più in generale si considera riguardante questa tematica tutti i casi in cui si combinano assieme i dati dell'asse temporale con i dati spaziali. L'evoluzione del movimento di un oggetto nel tempo è diventato recentemente una importante tematica di ricerca. La gestione di questi dati in domini separati è ben consolidata: da una parte gli *Spatial Databases* sono in grado di gestire in modo efficiente i dati posizionali e le operazioni di interrogazione su di essi, così come dall'altra parte i *Temporal Databases*. Tuttavia la combinazione simultanea di questi due insiemi può aprire ad importanti prospettive di applicazione. Prendiamo ad esempio l'analisi del traffico in una rete di trasporto cittadina. E' possibile (ed è già stato realizzato in diversi studi) dotare un numero consistente di veicoli circolanti di un dispositivo GPS. Il dispositivo registra le informazioni in merito alla posizione con una frequenza sufficientemente dettagliata (ad esempio 0.2 Hz, una campionamento ogni 5 secondi). La raccolta di queste informazioni può portare ad un *set* di dati che se interpretato correttamente può fornire un supporto a richieste del tipo:

- **analisi del traffico:** quanti veicoli affollano un determinato tratto in un momento specifico? Qual'è il tempo di attesa medio ad un semaforo oppure funziona correttamente l'effetto "onda verde"?
- **servizi *location-aware*:** qual è l'attività commerciale più vicina alla mia posizione attuale o al percorso che ho intenzione di seguire? Quali dei miei amici *facebook* è in prossimità della città che sto visitando?

Quelli riportati sono solo alcuni esempi, ma più in generale lo studio di *mobility data* e di *database* di traiettorie può supportare lo sviluppo di *Location Based Services* ed applicazioni *Location-* e *Mobility-Aware*. Per ***Location Based Service*** si intendono tutti quei servizi che forniscono informazioni ai propri utenti sulla base della posizione geografica corrente. Se il servizio è caratterizzato da un altro grado di interazione tra gli utenti stessi e la formazione di una sorta di rete fra gli stessi, allora si può parlare di ***Location-based Service Networking***. Una tassonomia di questi servizi è proposta in tabella 2.1 e deriva dal grado di mobilità (stazionario o mobile) degli attori coinvolti ovvero l'utente che interroga il servizio e gli oggetti interrogati a database.

Database Objects \ Reference Object	<i>Stationary</i>	<i>Mobile</i>
	<i>Stationary</i>	<i>Mobile</i>
<i>Stationary</i>	routing	guide-me
<i>Mobile</i>	find-me	get-together

Tabella 2.1: Tassonomia di applicazioni *location-aware*

Come vediamo l'ambito del progetto *PathS* ha molti aspetti in comune con quelli trattati nel *mobility data managment*. Tuttavia le problematiche che si intendere risolvere (almeno nelle prime fasi del progetto) riguardano la prima classe di problemi ($\langle stationary, stationary \rangle$) relativamente più conosciuta e trattata in letteratura.

Modellando la rete di trasporto (*transportation network*) come un grafo $G = (N, E)$ composto da nodi N e archi E , l'operazione di *routing* consiste tipicamente nel cercare un percorso ottimale che porti dalla sorgente S alla destinazione T dove $S, T \in N$. Tecnicamente per queste tematiche sono disponibili diversi algoritmi, principalmente basati sulla soluzione del problema di flusso *Shortest Path*. Un adeguamento del calcolo dei pesi utilizzati negli archi della rete, consente di ottenere diverse modalità di navigazione ad esempio la più veloce, la più breve o, come è di recente interesse, la più ecologica.

Per questi motivi nel corso dello sviluppo del progetto *PathS* ci si è allontanati da quelle che sono le problematiche e gli strumenti specifici nell'ambito del *mobility data managment*, seguendo piuttosto altri spunti da cui derivare l'implementazione. Ciò non toglie che l'ambito di applicazione è molto coerente con questo tema e l'adozione di queste tecniche può tornare molto utile nelle possibili evoluzioni del sistema, in particolare se si intende utilizzare i

dati per successive analisi o interrogazioni che vadano oltre la tematica del *routing*.

2.2.1 Ricostruzione delle traiettorie

L'ampia diffusione di dispositivi dotati di sistema GPS combinata con lo sviluppo di adeguate tecniche di memorizzazione, processamento ed interrogazione dei dati ha portato alla produzioni di immense quantità di dati *location-aware*. Tuttavia la derivazione di informazioni significative da questi dati grezzi non è per nulla banale. Nella trasformazione di questi dati si incontrano diverse problematiche, tra cui la gestione dei segnali di rumore o non accurati, la semplificazione dei *set* di dati ed in particolare l'identificazione delle traiettorie come sequenza ordinata di posizioni campionate. La ricostruzione di una traiettoria a partire da un insiemi di dati grezzi è una operazione fondamentale che si rende necessaria prima di qualsiasi altra elaborazione o analisi dei dati.

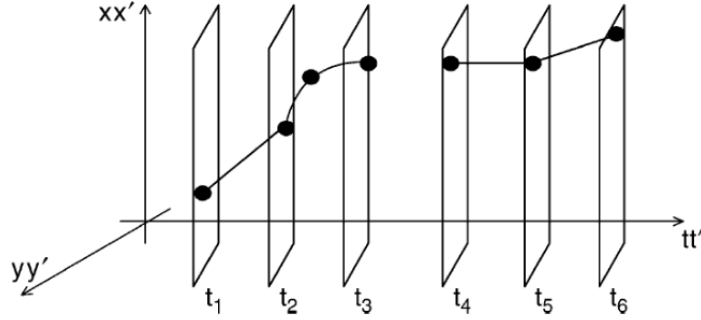


Figura 2.3: *Sliced representation* di un oggetto in movimento.

Come presentato in [4, capitolo 3.1], una definizione teorica del movimento di un oggetto può essere quella di una funzione dallo spazio temporale $I \subseteq \mathbb{R}$ allo spazio geografico $S \subseteq \mathbb{R}^2$:

$$I \subseteq \mathbb{R} \rightarrow S \subseteq \mathbb{R}^2 : t \rightarrow l(t)$$

dove con $l(t)$ si intende la posizione in cui si trova l'oggetto all'istante t . in questo modo la traiettoria effettiva seguita da un oggetto può essere definita come:

$$T_{act} = \{t, l(t) | t \in \mathbb{I}\} \subset \mathbb{R}^2 \times \mathbb{R}$$

Questo rappresenta una curva continua nella realtà, tuttavia la forma che ci si trova a rappresentare nei calcolatori è la sua versione finita, definita come

sequenza di coppie di valori spazio-tempo:

$$T = \{ \langle p_1, t_1 \rangle, \langle p_2, t_2 \rangle, \dots, \langle p_n, t_n \rangle \}$$

dove $p_i \in \mathbb{R}^2, t_i \in \mathbb{R}, 1 \leq i \leq n$ e $t_1 < t_2 < \dots < t_n$. Il risultato è che una traiettoria può essere rappresentata con un modello che si definisce *sliced representation* come quello in figura 2.3. Questo modello decompone lo sviluppo temporale in frammenti definiti *slice*, tali per cui questa evoluzione può essere descritta da una qualche funzione semplice, ad esempio l'interpolazione lineare.

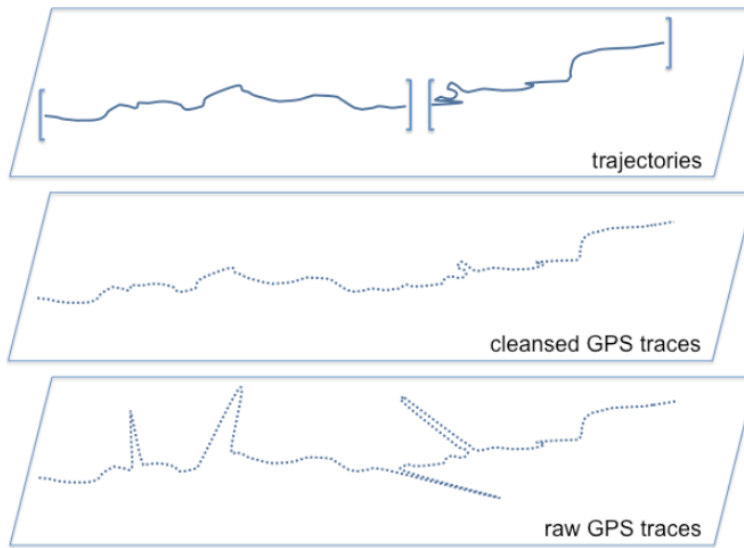


Figura 2.4: Il processo di *trajectory reconstruction* da dati grezzi GPS.

In generale il problema di *trajectory reconstruction* si decompone di due operazioni principali:

- *data cleansing*: è la fase in cui si ripuliscono i dati grezzi dai segnali di rumore o dagli eventuali *outlier*, è possibile eseguire operazioni di arrotondamento o altre elaborazioni;
- *trajectory identification*: ha l'obiettivo di interpolare la sequenza di posizioni raccolte al fine di approssimare il movimento continuo dell'oggetto.

In particolare per la seconda operazione le modalità sono tutt'altro che semplici e delineate. Sono disponibili diversi approcci, alcuni che si basano sulle caratteristiche geometriche della rappresentazione dei dati, altri si basano su sistemi probabilistici, altri ancora sono sistemi ibridi.

Anche in *PathS* si incontra la problematica fondamentale di ricostruire le traiettorie individuate dai campionamenti. L'approccio che si è valutato più opportuno (considerate le caratteristiche dei dati sorgente) è stato quello di utilizzare un algoritmo di *map matching*. Il caso di *trajectory reconstruction* che si affrontato è stato quindi quello di ricondurre ciascun campionamento ad una posizione corrispondente nella rete di trasporto. Utilizzando la rappresentazione di grafo $G = (V, E)$ composto di archi e vertici, per ciascuna posizione $\langle p_i, t_i \rangle$ della traiettoria che non appartiene alla rete, si è individuata la corrispondente coppia $\langle p_i^i, t_i^i \rangle$ in cui $p_i^i \in e_j$.

2.3 Versione precedente e lavori correlati

L'idea di sviluppo del progetto non è del tutto nuova, ma prende spunto da un lavoro precedente svolto sempre all'interno dell'Università di Padova dallo studente Lorenzo Teodori seguito dal Prof. Palazzi. Il progetto chiamato *Path2.0* voleva essere un sistema partecipativo per fornire indicazioni di navigazione accessibile per utenti con disabilità. L'idea di base è quella sviluppata in *PathS*, ovvero raccogliere informazioni da un insieme di utenti per poi riutilizzarle nel suggerire dei percorsi pedonali. Gli obiettivi in cui questo progetto estende e sviluppa l'idea iniziale sono i seguenti:

- estendere il concetto di informazione aggiuntiva sul percorso, non trattare solo l'accessibilità (valori 1, 0) ma delle *label* più generiche;
- implementare soluzioni che rispondessero con percorsi diversi dal semplice percorso accessibile e percorso breve;
- fornire una base di dati elaborati che consenta non solo il *routing*, ma anche l'analisi o l'applicazione ad altri scenari in chiave sociale.

Dal punto di vista tecnologico, si è preferito fare tesoro dell'esperienza precedente e strutturare un sistema completamente nuovo per le seguenti motivazioni:

- il sistema SDK utilizzato nello sviluppo dell'applicazione *mobile* era molto datato, richiedeva un profondo aggiornamento e la revisione di alcune modalità di sviluppo;
- tutta la soluzione era molto legata e del tutto dipendente dai servizi API di Google Maps (anch'essi da aggiornare);
- il progetto di partenza non presentava dei caratteri architettonici sufficientemente delineati da consentirne l'utilizzo come base, in particolare per le esigenze di espansione ed evoluzione ad altri contesti;

Un progetto analogo a *PathS* nel quale si utilizza però il microfono del dispositivo è Crowd++ [5]. Lo scopo dell'applicazione sviluppata è di individuare dati aggiuntive riguardo l'attività sociale dell'utente e l'interazione tra *speaker* e pubblico in un evento. I suoni registrati dal microfono sono utilizzati per individuare il numero di parole pronunciate o il numero di persone che stanno parlando, cercando quindi di derivare le informazioni necessarie. Altro esempio da cui si è tratto spunto è il progetto PartecipAct [2]. Sviluppato recentemente presso l'Università di Bologna, anch'esso ha come obiettivo quello di studiare il potenziale inesplorato della collaborazione tra utenti, sfruttando gli smartphone come strumento di interazione e interconnessione. Coinvolgendo studenti e volontari in un esperimento di raccolta dati, si sono raccolte informazioni sugli utenti e sul contesto in cui si trovano. Tutti i dati vengono memorizzati in un database online cercando rispondere a domande del tipo: "dove passano la maggior parte del tempo i nostri utenti?". Come vediamo l'approccio è molto simile a quello di *PathS*, nonostante le modalità di svolgimento e il fine ultimo siano diversi.

Capitolo 3

PathS Client

Tra le due componenti in cui si articola il progetto, la prima sviluppata è stata la componente *client*. Questa parte consiste sostanzialmente di una applicazione per dispositivi *smartphone* Android, completamente riscritta a partire da alcuni sviluppi precedenti per poter assolvere alle nuove funzioni specifiche per il progetto *PathS*. Si riassumono quali sono i requisiti che sono stati identificati e quali sono stati i criteri principali considerati nello sviluppo del software.

3.1 Requisiti

Presentazione degli obiettivi riguardanti la parte client:

- raccogliere campioni tramite sensore di luminosità e microfono;
- offrire indicazioni di percorso;
- coinvolgere utente (tramite esperienza Augmented Reality).

3.2 Tecnologie

Riassunto delle tecnologie adottate per la realizzazione della componente *client*: AndroidSDK, Wikitude.

3.3 Stato dell'arte

Risultati ottenuti con la parte client. Funzionamento ed eventuali limiti riscontrati.

Capitolo 4

PathS Server

Introduzione alla componente *server* di *PathS*.

4.1 Requisiti

Dettaglio dei requisiti per la parte server: raccolta campioni, persistenza, interfacciamento con il client, associazione dei campioni ai percorsi, servizio di routing.

4.2 Componenti

Identificazione delle componenti logiche del server sulla base delle funzioni che deve svolgere.

4.2.1 Ricezione dei campioni

La componente di ricezione campioni deve colloquiare con il client. Necessità di definire un formato chiaro e *standard*, facilmente implementabile con qualsiasi tecnologia *client* ed eventualmente da altri sorgenti dati.

4.2.2 Elaborazione dei campioni

I campioni così come sono ricevuti dai client sono in formato grezzo e non consentono di sfruttare le informazioni in esse contenute. Il passo principale per l'utilizzo dei dati è quello di associare ciascun dato ad un segmento della rete di trasporto.

4.2.3 Interrogazione

Lo scopo pensato per l'utilizzo delle informazioni raccolte è quello di fornire servizi *routing* alternativi, che tengano conto dei dati raccolti per suggerire percorsi che vanno oltre la semplice regola del percorso più breve. Il componente deve quindi:

- definire un formato di colloquio con il client per i percorsi;
- implementare un algoritmo di routing;
- modificare l'algoritmo di routing affinché consideri le informazioni aggiuntive.

4.3 Tecnologie e linee guida

Tecnologie e librerie selezionate per l'applicazione globale al sistema server (Play! Framework, PostgreSQL, LeafletJS, Bootstrap). Presentazione dei criteri non funzionali con i quali è stato sviluppato il server (suddivisione delle responsabilità, astrazione, estensibilità).

Capitolo 5

Ricezione dei campioni

5.1 Formato

Definizione e modifiche al formato di comunicazione.

5.2 Implementazione

Modalità di implementazione e deserializzazione dei dati.

5.3 Esempi

Esempio di invio dal client e risultati ottenuti.

Capitolo 6

Elaborazione dei campioni

6.1 Servizi di Cartografia

Requisiti per la selezione di un servizio di mappe su cui basarsi. Analisi comparativa e aspetti principali di:

- Google Maps
- Openstreet Map

Il servizio di interrogazione *Overpass API*, funzionalità e tipo di *query* implementata.

6.2 Persistenza ed elaborazione dei dati GIS

Libreria scelta per la persistenza dei dati PostGIS, modello dati del DB e tipi utilizzati. Operazioni di interrogazione supportate dall'estensione del database. Libreria scelta per la manipolazione dei dati in ambiente *JAVA:JTS*. Operazioni necessarie ed esempi di utilizzo.

6.3 Algoritmo di Map Matching

Perchè è necessario eseguire una operazione di *map matching*. A cosa servono questo tipo di algoritmi, panoramica sullo stato dell'arte e le soluzioni possibili.

6.3.1 ST-MapMatching

Esposizione delle caratteristiche principali dell'algoritmo come presentato in [3].

6.3.2 Pre-elaborazione del percorso - step 1

Calcolo della bounding box, recupero della rete di trasporto coinvolta e inserimento a sistema.

6.3.3 Identificazione dei candidati - step 2

Definizione e modalità di calcolo dei candidati. Operazioni implementate ed esempi di esecuzione.

6.3.4 Valutazione e selezione dei candidati - step 3

Applicazione della valutazione spazio-temporale dei campioni. Algoritmo di selezione dei candidati da assegnare ai campioni.

Capitolo 7

Routing

7.1 Implementazione

Modalità generali di implementazione e interfaccia *Router*. Pre-elaborazione e gestione dei segmenti a sistema tramite operazione di *noding* e creazione della topologia (nodi e archi).

7.2 Percorsi calcolati

7.2.1 Servizio Map Quest

Interrogazione del servizio Map Quest come riscontro e *fallback*.

7.2.2 Shortest Path

Calcolo del percorso minimo con la libreria PGRouting.

7.2.3 Percorsi con *label*

Calcolo dei percorsi valutando le label di luminosità e rumorosità. Viste e funzioni di supporto sul database. Modalità di calcolo dei costi, spiegazione delle formule. Requisiti del servizio di routing. Definizione del formato di colloquio (GeoJSON) e interpretazione lato client.

Capitolo 8

Conclusione

8.1 Risultati

Risultato complessivo del progetto. Funzionamento end-to-end con screenshot ed esempi.

8.2 Miglioramenti ed Evoluzioni

Limiti e compromessi dell'attuale implementazione. Aree di miglioramento ed evoluzione, tra cui:

- suddivisione ulteriore dei tratti di percorso;
- alternative algoritmi di map-matching;
- altri algoritmi di routing;
- servizi aggiuntivi al routing (analisi dei posti più frequentati, funzioni “get-together” e *social oriented*).

Appendice

Appendice A

Manuale installazione

Si riassumono i requisiti necessari all'installazione del sistema e le modalità di configurazione.

A.1 Ambiente

Il sistema è stato sviluppato in ambiente *Linux* e per la sua installazione ed esecuzione si consiglia il sistema operativo *Ubuntu Linux ver. 14.04 LTS*. Il requisito non è vincolante, ma semplifica l'installazione e la gestione dei pacchetti necessari per le dipendenze. In termini di risorse il sistema richiede almeno 2 unità di calcolo a 64 bit e 2GB di memoria RAM disponibile.

A.2 Dipendenze

L'ambiente di esecuzione si compone di due elementi principali, il **database** contenente tutte le informazioni persistenti del sistema e la **logica applicativa** sviluppata con tecnologia *Java*.

A.2.1 Database

E' necessario utilizzare un database *PostgreSQL ver 9.3* o superiore. L'installazione può avvenire sia tramite il gestore dei pacchetti del sistema operativo che in modalità manuale. Oltre al pacchetto standard, sarà necessario installare due estensioni del database, in particolare:

- **PostGIS**: una estensione che consente la memorizzazione di dati spaziali e fornisce altre funzioni di manipolazione di dati geografici (<http://postgis.net>);

- **PGRouting**: una libreria che estende ulteriormente le funzioni del database geospaziale introducendo algoritmi di *routing* e procedure di supporto (<http://pgrouting.org>).

Nei sistemi *Ubuntu Linux* l'installazione del servizio di database e di tutte le dipendenze può avvenire tramite l'esecuzione dei comandi:

```
sudo add-apt-repository ppa:georepublic/pgrouting-unstable
sudo apt-get update
sudo apt-get install postgresql-9.3-pgrouting
```

Seguire le indicazioni presentate a riga di comando, in particolare per le prime operazioni in cui si aggiunge un *repository* non ufficiale al gestore dei pacchetti.

A.2.2 Java

Per eseguire il codice con cui è stata sviluppata la logica applicativa del server, è necessario disporre di un ambiente *Java Virtual Machine* compatibile. Si consiglia di installare il pacchetto *Oracle Sun JDK ver. 1.7.0* secondo le modalità previste dal sistema operativo usato. Non si garantisce l'esecuzione del software in ambiente *Java 8* nè con altre *JVM* diverse da quella Oracle SUN.

Nello sviluppo è stato utilizzato il framework di supporto *Play! Framework ver. 1.2.7*. Installare il pacchetto seguendo le indicazioni sul sito ufficiale (<https://www.playframework.com/documentation/1.2.7/install>). La modalità standard consiste nello scaricare l'archivio e scompattarlo in una posizione nota. Le librerie necessarie all'esecuzione del codice sono incluse nel *repository* del codice e saranno recuperate tramite i successivi passi di installazione.

A.2.3 Repository

Il codice sviluppato è disponibile tramite *repository* pubblico *GIT* all'indirizzo <https://github.com/tzorzan/path-server.git>. Clonare il *branch master* in locale per ottenere una copia aggiornata del codice da eseguire.

A.3 Configurazione

Eseguire il *restore* del database iniziale. Il file da cui eseguire il ripristino è presente nella posizione *install/paths.backup* ed è in formato archivio. Eseguendo questa operazione si crea un database con la struttura corretta ed

un set di dati di partenza. Il ripristino può essere eseguito tramite tool grafici o da linea di comando utilizzando le modalità previste da *PostgreSQL* (<http://www.postgresql.org/docs/9.3/static/backup.html>).

A questo punto è necessario istruire l'applicazione fornendo il puntamento al database da utilizzare. Questo può essere fatto impostando la variabile d'ambiente *DATABASE_URL* prima di avviare l'applicazione. Impostare il valore sostituendo i parametri necessari:

```
export DATABASE_URL=postgres://<utente>:<password>@<db_address>:<db_port>/<db_name>
```

Impostare le credenziali di un utente che possiede diritti completi sullo schema, dato che alcune procedure richiedono la cancellazione e la creazione di alcune tabelle.

Per installare le dipendenze (librerie *Java*) necessarie all'esecuzione è necessario posizionarsi nella directory base del progetto e lanciare il comando:

```
play deps
```

Saranno scaricate tutte le dipendenze necessarie e salvate nella cartella */lib* usata durante l'esecuzione del software.

A.4 Avvio

Avviare l'applicazione tramite il comando:

```
play run
```

In questo modo si avvia il server e resta collegato al terminale in uso.

Per avviare in *background* il servizio e quindi lasciare il server attivo usare i comandi di avvio e spegnimento:

```
play start
```

```
play stop
```

Il software utilizzerà come *log* di sistema il file: */logs/system.out*.

A.5 Note

E' necessario riconfigurare i client mobile affinché puntino al nuovo servizio server avviato. Questa operazione allo stato attuale può essere effettuata modificando il file sorgente */com/path3ar/model/Utilities.java* e rieseguendo packaging ed installazione dell'applicazione.

Bibliografia

- [1] Giampiero Calma, Claudio E. Palazzi e Armir Bujari. “Web Squared: Paradigms and Opportunities”. In: *Proc. of the International Workshop on DIstributed SIMulation and Online gaming (DISIO 2012) - ICST SIMUTools 2012*. Desenzano, Italy, 2012.
- [2] Giuseppe Cardone et al. “The ParticipAct Mobile Crowd Sensing Living Lab: The Testbed for Smart Cities”. In: *IEEE Communications Magazine* (ott. 2014).
- [3] Yin Lou et al. “Map-Matching for Low-Sampling-Rate GPS Trajectories”. In: *ACM SIGSPATIAL GIS 2009*. Association for Computing Machinery, Inc., nov. 2009. URL: <http://research.microsoft.com/apps/pubs/default.aspx?id=105051>.
- [4] Nikos Pelekis e Yannis Theodoridis. *Mobility Data Management and Exploration*. Springer-Verlag New York, 2014.
- [5] Chenren Xu et al. “Crowdsensing the Speaker Count in the Wild: Implications and Applications”. In: *IEEE Communications Magazine* (ott. 2014).