

UNIVERSITÀ DEGLI STUDI  
DI MODENA E REGGIO EMILIA

Facoltà di Scienze Matematiche, Fisiche e Naturali  
Corso di Laurea Magistrale in Informatica

Tecniche Communication-saving  
per l'Acquisizione Dati  
nel Sistema di Trasporto Intelligente Pegaus

Marcello Pietri  
Sintesi della Tesi di Laurea

discussa nell'Anno Accademico 2009/2010  
il giorno 22 Settembre 2010

## - Motivazioni della tesi -

I problemi relativi al traffico delle nostre città, caratterizzati dall'aumento del consumo di carburante, dall'inquinamento acustico e ambientale, dall'elevato tasso d'incidenti e dalla conseguente generazione di congestioni, sono problemi ben noti che comunemente deteriorano la qualità di vita delle persone. Lo sviluppo di nuovi concetti del trasporto e della mobilità sono stati promossi dal 7th UE Framework Programme, con l'obiettivo di sviluppare innovative ed efficaci iniziative, riunendo tutti gli elementi di un trasporto pulito, a basso consumo energetico, sicuro ed intelligente.

L'Intelligent Transportation System (ITS) definisce le tecnologie tese ad integrare l'Information and Communications Technology (ICT) alle infrastrutture dei trasporti e dei veicoli, al fine di aumentare la sicurezza, e di ridurre l'inquinamento, i tempi di trasporto, il numero di incidenti e i consumi di carburante. L'ITS nasce quindi dalla necessità di gestire i problemi causati dalla congestione del traffico attraverso una sinergia delle nuove tecniche informatiche per la simulazione, il controllo in tempo reale e le reti di comunicazione.

Nell'ambito ITS è inserito il progetto PEGASUS, il cui obiettivo è sviluppare un sistema di trasporto intelligente realizzando una piattaforma infotelematica per fornire soluzioni di mobilità per una gestione del traffico efficiente ed efficace. Questo progetto nasce dalla collaborazione di diverse aziende, tra cui Octotelematics e Meta System, e dalla collaborazione di diversi gruppi di ricerca universitari tra cui ISGroup, presso il quale è stato svolto il seguente lavoro di ricerca.

Nel contesto così descritto, ogni veicolo appartenente al sistema è equipaggiato con un *On-Board Unit* (OBU), ovvero un'unità di bordo che comunica tramite GPRS con un centro di controllo, e via rete WiFi con le altre unità di bordo nelle vicinanze. La comunicazione mediante GPRS tra OBU e centro di controllo è denominata V2I, ovvero *Vehicle to Infrastructure*, mentre la comunicazione via rete WiFi tra le varie unità di bordo è denominata V2V, ovvero *Vehicle to Vehicle*.

Considerando che nella realtà sono circa settecentomila i veicoli circolanti dotati di OBU, e se ne prevedono almeno un milione in tempi brevi, il numero di comunicazioni e la quantità di dati scambiati risultano ostativi. L'obiettivo delle tecniche *communication-saving* è appunto quello di ridurre il traffico di dati sia tra le varie unità di bordo che tra unità di bordo e centrale operativa.

Come nel caso della compressione dati, vi possono essere tecniche *lossy*, ovvero con perdita d'informazione, che sfruttano le ridondanze nell'utilizzo dei dati, oppure *lossless*, ovvero senza perdita d'informazione, che sfruttano le ridondanze nella codifica del dato. Nel caso delle tecniche *communication-saving* saranno utilizzate principalmente metodologie *lossy*, mentre procedimenti *lossless* potranno essere adottati nella fase immediatamente successiva per comprimere i dati generati dalle prime.

Nello scenario così descritto è inserito questo progetto di ricerca, i cui obiettivi principali sono la riduzione delle comunicazioni V2I mediante l'utilizzo di tecniche *communication-saving* e mediante strategie di aggregazione V2V. L'implementazione è stata sviluppata in modo da essere portabile sull'architettura reale di ogni OBU e la fase di analisi ha permesso di scegliere le migliori tecnologie in base al relativo contesto.

Scendendo nei particolari, per analizzare e valutare i processi di invio e ricezione di dati, e per studiare la formazione di *cluster* tra OBU, è stato necessario progettare e sviluppare un simulatore, poiché non sarebbe stato possibile utilizzare i mezzi fisici necessari per questo scopo, sia per questioni pratiche che per questioni economiche.

La validazione dei risultati ottenuti è quindi sostenuta dalla validità dei dati in input, i quali sono stati generati da un simulatore di traffico professionale della PTV Vision chiamato Vissim, operante su scenari pressoché identici alla viabilità reale.

Lo sviluppo e l'implementazione delle tecniche *communication-saving* effettuato ha consentito la portabilità del codice sull'architettura reale di ogni OBU in modo agevole e veloce, mentre la dettagliata analisi sperimentale dei dati ottenuti dalla simulazione ha permesso di decretare quali di queste tecniche sono effettivamente utilizzabili, e anche di decidere quali sono le migliori per i servizi richiesti nei differenti scenari urbani del nostro paese.

## - Metodi di indagine -

Lo sviluppo del progetto riflette fedelmente la struttura presentata nel documento di tesi, a sua volta suddiviso in due parti: la prima parte riporta il caso di studio, mentre la seconda riporta il progetto l'implementazione e l'analisi sperimentale seguita dalle conclusioni e dagli sviluppi futuri.

La struttura dettagliata riflette quindi la seguente indicizzazione:

- studio del problema e definizione degli obiettivi, con la presentazione delle tematiche e delle problematiche che verranno affrontate nel corso della tesi;
- ricerca in letteratura, con la descrizione delle tecnologie utilizzate per l'acquisizione dei dati sul traffico e per lo sviluppo del simulatore;
- analisi e progetto del software, con la progettazione del simulatore secondo le specifiche SRS e il linguaggio di modellazione UML;
- implementazione, con lo sviluppo pratico di quanto delineato con il progetto del software;
- analisi sperimentale, con l'analisi dei dati ottenuti dalle simulazioni e la relativa deduzione degli aspetti chiave di questo progetto.

### Studio del problema:

I problemi affrontati nello sviluppo di questo progetto riguardano principalmente la mancanza di dati reali sul traffico e un ambiente di simulazione dove poter testare le tecniche *communication-saving* e le strategie V2V proposte. Partendo dalle specifiche del *paper Data Management Issues for Intelligent Transportation Systems* di F. Mandreoli, R. Martoglia, W. Penzo, S. Sassatelli presentato nel 18th Italian Symposium on Advanced Database Systems (SEBD 2010), pagine 198–209, 2010, si è proseguito estrapolando e definendo tutti gli obiettivi necessari alla risoluzione del problema.

### Definizione degli obiettivi:

Gli obiettivi principali del progetto, acquisiti dal documento di cui sopra, sono risultati essere i seguenti:

- ridurre il carico sulle comunicazioni V2I utilizzando tecniche per il risparmio di dati;
- indicare dopo un'attenta analisi le migliori strategie *communication-saving* utilizzabili in pratica;
- ridurre il carico sulle comunicazioni V2I utilizzando comunicazioni V2V;
- stimare la necessaria quantità percentuale di veicoli necessari per rendere possibile un efficiente utilizzo delle comunicazioni V2V;
- utilizzare comunicazioni e strategie V2V per ridurre il carico sulle comunicazioni V2I.

Dagli obiettivi principali sono emersi gli ulteriori seguenti obiettivi necessari per la realizzazione di un corretto lavoro:

- riuscire a implementare le tecniche descritte in modo da rendere il codice portabile sull'architettura reale senza troppi disagi;
- cercare e/o creare nuove tecniche *communication-saving* eventualmente utilizzabili;
- confrontare tra loro pregi e difetti delle diverse tecniche;
- ottenere stime il più verosimili possibile alla realtà;
- definire delle strategie e delle politiche di *clustering* tra veicoli (ed eventualmente tra *cluster*).

## **Ricerca in letteratura:**

Per aumentare la produttività e migliorare la qualità del lavoro svolto è stata effettuata una ricerca esaustiva di quanto già presente in letteratura. Da sottolineare che una ricerca in letteratura è un passo dettagliato e organizzato di ricerca di tutto il materiale disponibile su un determinato argomento. Le fasi principali della ricerca bibliografica svolta sono state la selezione delle fonti di informazione, quali biblioteche, riviste, libri, atti dei convegni, dissertazioni e tesi, e l'identificazione del materiale pertinente. Dai quarantasette documenti selezionati, presentati in bibliografia, si sono delineate le tecnologie da utilizzare e i principi guida per lo sviluppo del software e per l'analisi sperimentale.

## **Analisi e progetto del software:**

Avendo così appreso il problema nella sua interezza ed avendo acquisito le necessarie conoscenze per risolverlo, in questa fase si sono utilizzate le specifiche SRS per l'analisi e il linguaggio di modellazione UML per il progetto. Sulla base della specifica dei requisiti prodotta dall'analisi, il progetto definisce come tali requisiti saranno soddisfatti, entrando nel merito della struttura che dovrà essere data al sistema software che deve essere realizzato.

## **L'implementazione**

La fase di implementazione deriva dalla fase di progettazione del software, infatti, per comprendere la struttura del progetto sono stati presentati i *package* ed i *class diagram* definitivi dello standard UML, che rispecchiano fedelmente la struttura del codice. Il codice sviluppato segue le specifiche definite in *The Java Language Specification* di Guy Steele, James Gosling, Bill Joy and Gilad Bracha, Sun Microsystems, third edition, 1996-2005, ISBN 0-321-24678-0, un modello di programmazione in stile GNU. La politica di utilizzo degli oggetti e di entità ad alto livello è stata condizionata dall'esigenza futura di portare parte del codice sviluppato sull'architettura reale delle più di settecentomila OBU in circolazione.

Il software, oltre ad essere stato testato in tutte le sue parti mediante apposite classi ed essere risultato corretto nella totalità dei casi, è stato progettato in ottica di configurabilità, espandibilità e usabilità futura.

## Analisi sperimentale

L'analisi sperimentale con l'ausilio del simulatore sviluppato, intesa anche come scelta degli scenari, dei parametri e delle tecniche su cui effettuare prove sperimentali, è il passo conclusivo e di maggior rilevanza di questo progetto.

In merito agli scenari usati, il primo passo per intraprendere una corretta analisi sperimentale è stato quello di scegliere gli scenari in modo da rispecchiare la realtà nel modo più fedele possibile. Non sempre questo è possibile, soprattutto nel nostro caso, dato che per scenari si ci riferisce ad esempio a tutti i veicoli che circolano nelle strade della città di Bologna. La soluzione adottata è stato l'utilizzo di un simulatore di traffico professionale della PTV Vision denominato Vissim, che però ha richiesto a sua volta la costruzione e la modifica di mappe, anche a livello di incroci, impianti semaforici e veicoli. I dati ottenuti da questo simulatore di traffico sono stati processati dal simulatore da me interamente progettato e realizzato.

In merito ai parametri usati, questi sono stati definiti inizialmente sulla base dei documenti presenti in letteratura, in modo da poter confrontare parte dei risultati ottenuti. Successivamente i parametri utilizzati sono derivati dall'esperienza acquisita dall'ingente numero di simulazioni precedenti. Le scelte effettuate in merito sono dunque giustificate sulla base dei risultati di test preliminari e dello stato dell'arte attuale. Oltre a tutto ciò per fissare i parametri con cui fare analisi serve anche definire una linea guida, ovvero un parametro base a cui fare riferimento per i confronti, in modo da sapere sempre se il lavoro svolto ottiene risultati positivi, neutri o negativi. L'altro aspetto da definire riguarda l'utilizzo dei dati trasmessi, infatti, se non si definiscono dei servizi che li utilizzano non si potrà mai sapere cosa serve e cosa non serve. Il criterio stabilito per valutare la qualità dei dati trasmessi riguarda quindi la percentuale di errore relativa alle distanze percorse ed alle velocità, considerando, ad esempio, servizi di tipo assicurativo, che basano il premio sulle distanze percorse, servizi di tipo informativo, quali la stima dei tempi di percorrenza di un tratto stradale e servizi di tipo preventivo, che verificano le velocità in determinati tratti stradali. In tutti i casi comunque, ciò che si valuta è la precisione sia delle distanze che delle velocità trasmesse, rispetto quelle reali del veicolo. Per calcolare le distanze reali si è fatto riferimento ai dati riportati dal simulatore di traffico, ovvero la distanza percorsa dal veicolo durante la simulazione stessa ed il suo tempo di vita. Dividendo la distanza percorsa per il tempo di vita si è ottenuta la velocità media del veicolo.

Le statistiche dettagliate derivate dall'analisi sperimentale sono state in primo luogo ottenute tramite fogli di calcolo, poi fatte calcolare dal programma stesso, in modo da ottenere una piena validazione dei risultati. Al termine di ogni analisi e di ogni gruppo di analisi sono state inoltre riportate delle note e delle deduzioni conclusive in merito.

Infine, dall'analisi acquisita dai test effettuati, si ci è accorti che i risultati ottenuti, oltre alla loro correttezza nella quasi totalità dei casi, salvo piccole incongruenze dovute a fattori pressoché irrilevanti al nostro scopo, erano in molti casi migliorati, sia per merito delle tecniche *communication-saving* sia per merito delle altre migliorie apportate.

## - Risultati ottenuti -

Durante lo svolgimento del progetto sono stati portati a termine tutti gli obiettivi prefissati e sono state sviluppate delle funzionalità aggiuntive che hanno permesso l'analisi di tecniche *communication-saving* maggiormente avanzate, al fine di ricercare le migliori da applicare nel contesto reale, in particolare:

- il numero delle comunicazioni V2I è stato ridotto utilizzando tecniche *communication-saving* senza compromettere il *trade-off* dati inviati - errore commesso;
- sono stati codificati, compressi e studiati i flussi di dati provenienti dalle OBU verso il centro di controllo, al fine di conoscere le potenziali quantità di dati con cui si avrà a che fare nella realtà;
- dopo un'attenta analisi sono state individuate le migliori strategie *communication-saving* utilizzabili nei contesti reali richiesti;
- sono state studiate ed implementate nuove tecniche *communication-saving* al fine di apportare miglioramenti a quelle inizialmente proposte;
- l'implementazione delle tecniche descritte è stata svolta in modo da rendere agevole la portabilità del codice sull'architettura reale;
- sono state confrontate le diverse tecniche al fine di individuare pregi e difetti di ognuna;
- sono stati analizzati alcuni scenari reali per mostrare le quantità necessarie di veicoli che rendono possibile un efficiente utilizzo delle comunicazioni V2V;
- sono state definite strategie e politiche di *clustering* tra veicoli;
- è stata effettuata la stima della riduzione del carico sulle comunicazioni V2I, derivata dall'utilizzo di comunicazioni V2V.

Considerando che tutto il progetto è stato svolto in un'ottica di estensione e usabilità orientate agli sviluppi futuri del progetto Pegasus, sono di seguito presentate alcune proposte.

In particolare, per quanto riguarda le tecniche *communication-saving*:

- queste si dovrebbero testare su scenari molto più estesi, con dati sul traffico aggiornati e forniti da enti certificati, al fine di validare definitivamente le analisi svolte;
- la codifica sviluppata avvantaggia un'ulteriore codifica in stile Huffmann, la quale potrebbe essere implementata insieme ad altre tecniche di compressione, al fine di ridurre i dati inviati alla centrale;
- sullo stile della tecnica *Linear Regression*, la quale ha ottenuto ottimi risultati, potrebbero essere sviluppate tecniche più complesse quali *Multivariate adaptive regression splines* o *Segmented regression*.

Per quanto riguarda invece le possibili estensioni delle funzionalità dell'OBU:

- si potrebbero sviluppare ulteriormente gli algoritmi per il *clustering* V2V al fine di ottimizzare le comunicazioni ed i gruppi formati;
- si potrebbe studiare una tecnica di auto-selezione adattativa, in modo da ottimizzare l'utilizzo delle tecniche *communication-saving*; la validità di questo algoritmo sarebbe verificabile direttamente dall'OBU stessa, essendo questa a conoscenza dei dati inviati alla centrale.

La conclusione ultima di questa ricerca vedrebbe l'implementazione di quanto sviluppato in questo progetto sull'architettura reale di ogni OBU, a loro volta impiantate su veicoli realmente circolanti negli scenari urbani delle nostre città.