

Piattaforma per Location Based Services di logistica e mobilità interrogabile in linguaggio naturale

Guido GENTILE (*), Lorenzo MESCHINI (*), Giuseppe MUSSUMECI (**)
Maurizio BOMBARA (***), Davide CALÍ(***), Ivana CALÍ(***), Giuseppe TROPEA (***),
Luca Zagarella (***) , Antonio Pantò(****)

(*) Dipartimento di Idraulica Trasporti e Strade, Università “La Sapienza” di Roma, Via Eudossiana n.18 - 00184 Roma, tel. 0644585737, e-mail lorenzo.meschini@uniroma1.it

(**) Dipartimento di Ingegneria Civile e Ambientale, Università di Catania, V.le A.Doria n.6 - 95125 Catania, tel. 0957382212, e-mail gmussume@dica.unict.it

(***) BC S.r.l. Software Company, Via Caronda n.136 – 95125 Catania tel. 0957286481, e-mail info@bcsoftware.it

(****) NetSense S.r.l., Via Enrico Pantano n.93 – 95129 Catania, tel. 3474705746, e-mail info@netsenseweb.com

Sommario

Le basi di dati a supporto della logistica e della mobilità individuale sono spesso poco fruibili o poco correlate fra loro. In questo lavoro si presenta la realizzazione di un sistema integrato per l'accesso in linguaggio naturale a tali dati, capace di lavorare insieme a modelli di calcolo per la previsione dei flussi di traffico o di supporto per la gestione di alcune emergenze di protezione civile, e fruibile anche tramite dispositivi mobili agganciati ad una piattaforma di localizzazione.

Il prototipo realizzato è quindi in grado di rilevare in tempo reale tramite GPS o triangolazione WiFi la posizione utente, inoltrare al sistema richieste del tipo: *qual è il percorso più veloce per via Giusti*, ed ottenere la mappa e i dati in risposta.

Abstract

Databases which support Location Based Services and personal mobility are often difficult to query and hold loosely uncorrelated data. With this work we have achieved an integrated SW/HW prototype enabled with Natural Language query capabilities upon these data. We have integrated the NL engine together with two different Decision Support Systems, which are based on powerful analytical models for vehicular traffic flow analysis and for the efficient management of certain civil emergencies. The prototype is queried through portable devices, like PDAs or cellphones, connected to a custom localization engine able to transmit real-time GPS or WiFi location information of the user. Queries such as: *which is the fastest path to reach via Giusti?* can be forwarded to the server, and graphical results map and data table are received and displayed on the smartphone.

1- Introduzione

Le informazioni disponibili in forma strutturata nelle basi di dati a supporto della logistica e della mobilità individuale sono spesso, oltre che parziali nello spazio e nel tempo, non accessibili con interfacce semplici, tali da consentire agli utenti di esplorare i legami fra i vari strati informativi, né integrate con sistemi che tengano conto della posizione degli utenti sul territorio. Inoltre i dati sono spesso relativi a determinate sezioni stradali e tipologie di veicoli non coprendo dunque l'intera rete, oppure riguardano stati della stessa presenti e passati, ma non futuri.

Obiettivo del presente progetto è quello di rendere tali informazioni maggiormente fruibili, con particolare riferimento al supporto alle decisioni ed alle strategie di mobilità, alla possibilità di suggerire agli utenti (enti addetti alla supervisione del sistema, operatori logistici, singoli utenti) percorsi ottimali con riferimento alle distanze o alle condizioni di traffico effettivamente presenti e previste sulla rete di trasporto, corredati da una stima del tempo di percorrenza atteso. L'esplorazione delle informazioni avviene tramite un'intuitiva interfaccia in linguaggio naturale,

che consente di porre richieste del tipo: *dove sono i magazzini meno forniti?, dov'è il distributore GPL con self-service più vicino a piazza Garibaldi?, quali sono le strade meno trafficate intorno alla stazione?*

Il sistema è costruito a partire dal software d'analisi del linguaggio naturale *FuLL (Fuzzy Logic and Language)* di BC S.r.l., integrato ai modelli previsionali e decisionali sviluppati dal Dipartimento DITS dell'Università degli Studi di Roma "La Sapienza" e dal Dipartimento DICA dell'Università di Catania, ed inserito in una infrastruttura di rete e rilevamento della posizione tramite tecnologie GPS e WiFi sviluppate da NetSense S.r.l., allo scopo di ottenere una piattaforma flessibile e potente.

1- Architettura della piattaforma

Nello schema seguente è rappresentata una architettura di massima dei moduli funzionali e degli scambi di dati che avvengono.

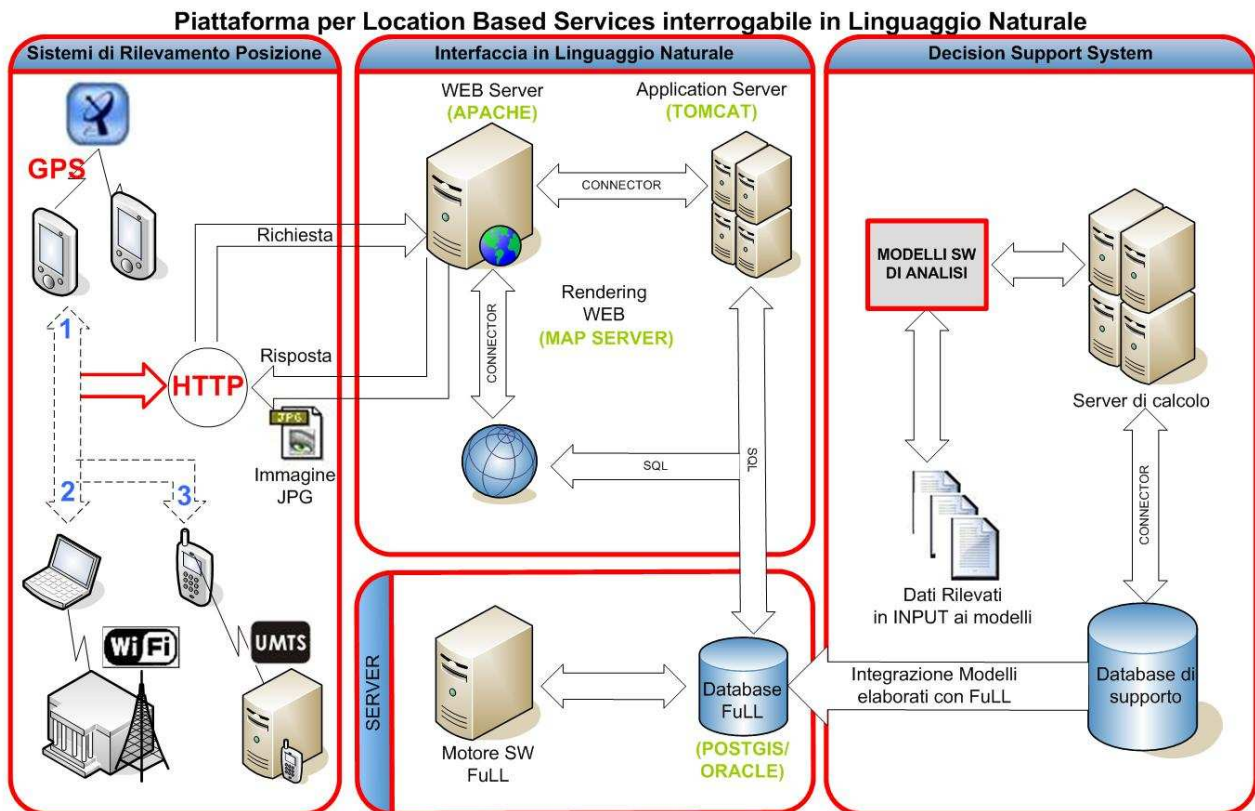


Figura 1 - Schema dell'architettura della piattaforma

1.2.- Potenziamento dell'interfaccia in Linguaggio Naturale

Dal punto di vista dell'interfaccia in linguaggio naturale all'intero sistema, le difficoltà maggiori incontrate per il prototipo hanno riguardato l'introduzione dei concetti di posizione spaziale dell'utente e di orario della richiesta, da un lato, e la gestione semantica del concetto di *percorso* da un punto origine arbitrario ad un arbitrario punto destinazione, dall'altro.

Il problema del percorso è legato alla natura ambivalente di tale concetto, rispetto alla usuale classificazione ontologica che abbiamo solitamente seguito nel mappare una base di dati alla conoscenza sintattica e semantica che di essa possiede il motore in linguaggio naturale di FuLL. Tale classificazione serve ad includere tutti gli aspetti e gli oggetti registrati nel DB, all'interno di categorie quali *funzione, entità, attributo, relazione, valore fuzzy* (Bartolini et al., 2006; Giovannetti et al., 2006), che costituiscono l'ossatura della Ontologia Semantica di FuLL. Nel caso del concetto *percorso*, questa classificazione risulta difficoltosa poiché esso deve essere mappato sia su una funzione di calcolo dinamica sia su di una entità che possieda una moltitudine di suoi attributi propri (come ad esempio l'orario di validità del percorso, il tempo di percorrenza, la lunghezza, o

altro) spesso solo implicitamente specificati nelle richieste. In sostanza, dal punto di vista semantico il concetto di percorso è sia una funzione (che prende tre parametri in ingresso, ovvero origine e destinazione e orario di partenza) e che deve essere valutata secondo la logica fuzzy insita in FuLL (percorso veloce, lento, brevissimo, etc...), sia una entità geometrica da materializzare nel contesto delle relazioni con le altre entità e dotata di attributi propri.

Per gestire questa ambivalenza è stato necessario introdurre nella catena dei moduli che costituiscono il motore di interfaccia in LN di FuLL, un nuovo passo intermedio di trattamento post-semantico della Forma Logica, la quale rappresenta il primo momento di traduzione dell'output del parser sintattico in concetti ontologici legati fra loro, successivamente traducibile in uno specifico dialetto SQL spaziale e temporale. In questo nuovo modulo, la *funzione percorso* viene re-interpretata come *entità percorso* tutte le volte che ciò è necessario.

Lo stesso approccio è stato impiegato per la gestione del concetto di utente. In tutte quelle richieste dell'utente dove risulti implicita la volontà di conoscere un percorso per il quale l'origine non è stata specificata integralmente, il modulo post-semantico provvede a manipolare la Forma Logica introducendo i necessari legami col concetto di utente e della sua posizione.

Si noti che l'Ontologia Semantica in questo caso è stata opportunamente modificata per includere la nuova entità *Utente*, con tutti i suoi attributi (la geometria per rappresentarlo sulla mappa, le coordinate spaziali rilevate, etc...). Il modulo post-semantico ha lo scopo di affinare la comprensione che il sistema ha della domanda immessa dall'utente tramite il dispositivo portatile, e perciò dotato di capacità di localizzazione.

1.3 - Modelli decisionali per le reti di trasporto

Il sistema di modelli proposto in questo ambito consente la fornitura in tempo reale di informazioni in merito allo stato del traffico su una rete di trasporto, con particolare riferimento all'individuazione di percorsi ottimali ed alla stima dei tempi di percorrenza previsti fra una qualsiasi coppia di punti (origine-destinazione) sulla rete stessa.

Specificatamente, tale sistema è in grado di:

- a) simulare il funzionamento del sistema di trasporto ottenendo una stima dello stato completo della rete di trasporto e dell'evoluzione di tale stato nel futuro;
- b) determinare, sulla base di tale stato variabile nel tempo, una stima dei tempi di percorrenza sulla rete fra una qualsiasi coppia di punti in funzione dell'istante di partenza;
- c) individuare, in base ai tempi e costi di percorrenza sulla rete, un percorso ottimale sulla rete fra una qualsiasi coppia di punti.

La metodologia adottata per la stima dei tempi di percorrenza e per il calcolo dei percorsi ottimali rientra nell'ambito dei *modelli dinamici di analisi e simulazione dei sistemi di trasporto*. I *modelli di assegnazione alle reti di trasporto con dinamica intraperiodale*, (*within-day Dynamic Traffic Assignment (DTA) models*) sono stati oggetto in questi ultimi dieci anni di un'intensa ricerca scientifica, e costituiscono una recente e fondamentale innovazione rispetto ai tradizionali modelli statici. In particolare, il modello di assegnazione dinamica adottato in questa sede è stato oggetto di diverse pubblicazioni internazionali (Bellei et al., 2005; Gentile et al. 2006) e presenta caratteristiche innovative ed uniche, tali da fornire un metodo di assegnazione dinamica molto efficiente in termini di risorse utilizzate e tempi di calcolo. Questo modello risolve il problema dell'assegnazione alle reti di trasporto con dinamica intraperiodale (DTA, *Dynamic Traffic Assignment*) su reti di grandi dimensioni, affrontando esplicitamente la simulazione del fenomeno dello spillback. A valle di tale modello, uno specifico algoritmo per la ricerca dinamica del percorso ottimo (in grado cioè di individuare il miglior percorso in presenza di tempi e costi di percorrenza variabili nel tempo) consente di fornire risposte a quesiti inerenti la mobilità.

1.4 - Modelli di supporto alla protezione civile

Le infrastrutture di trasporto, e in particolare le reti stradali, hanno un ruolo essenziale per la gestione ottimale delle emergenze di protezione civile, poiché è solo grazie ad esse che è possibile organizzare eventuali evacuazioni delle popolazioni interessate dall'evento calamitoso e far

confluire rapidamente i soccorsi necessari.

Numerosi studi sono stati condotti nell'ultimo decennio presso il DICA dell'Università di Catania sulla modellazione in ambiente GIS delle condizioni di rischio delle reti stradali e la simulazione dello stato di funzionalità post evento (D'Andrea et al., 2005, Cafiso et al. 2005). Sono stati studiati e definiti modelli di analisi del rischio in relazione a diversi possibili scenari di pericolosità (sismica, idrogeologica, da incendio, vulcanica, da trasporto di sostanze pericolose) con applicazioni concrete alla rete stradale della Sicilia orientale e, con maggiore livello di dettaglio, della provincia di Catania. Da questi studi emerge un concreto contributo alla modellistica del rischio e delle sue componenti (pericolosità, esposizione, vulnerabilità), ma, oggettivamente, i risultati a cui è possibile pervenire rimangono fruibili solo da parte di utenti esperti nella consultazione di SIT realizzati con le diverse piattaforme proprietarie impiegate in campo tecnico-scientifico. Succede così che, malgrado alcune realizzazioni concrete di SIT perfettamente funzionali alle esigenze della protezione civile, questi strumenti di supporto alle decisioni non vengano realmente utilizzati, perché richiedono competenze specialistiche troppo spinte e generalmente non disponibili all'interno delle diverse Organizzazioni che compongono la "protezione civile" (forze di polizia, esercito, vigili del fuoco, organizzazioni sanitarie e di volontariato, ...). L'applicazione che si presenta nell'ambito del presente lavoro segue una prima esperienza di collaborazione tra il DICA e BC Software sul tema dell'interrogazione in linguaggio naturale di SIT finalizzati alle esigenze della protezione civile (Condorelli et al., 2006). Esempi di quesiti in lingua italiana: *quali ponti potrebbero essere crollati?*, ovvero *quali sono i tratti di strada dove è probabile che non riesca a passare un mezzo pesante?*, *quale è il percorso più breve per raggiungere il fronte di fuoco?*, *in quanto tempo raggiungo la destinazione assegnata?*, *come raggiungo l'area di ammassamento più vicina alla stazione ferroviaria?*, *quali tratti di strada potrebbero risultare dissestati*". In questo lavoro, in particolare, sono stati utilizzati i modelli messi a punto da diversi ricercatori del DICA e ampiamente diffusi nella letteratura specialistica sulla valutazione delle condizioni di funzionalità post evento della rete stradale, per consentire, ad utenti non esperti di interrogare il SIT anche con dispositivi mobili agganciati ad una piattaforma di localizzazione.

1.5 - Tecnologie di localizzazione GPS e WiFi

I servizi basati sull'informazione di localizzazione dell'utente e sul tracciamento del suo movimento si basano su tecnologie fortemente dipendenti dal contesto applicativo e dalle caratteristiche specifiche dell'area fisica di erogazione del servizio. Le tecnologie maggiormente affermate negli ambiti outdoor e indoor sono rispettivamente il GPS e il WiFi. Il primo garantisce una copertura globale per applicazioni outdoor mentre il secondo ha bisogno di una infrastruttura di rete che, pur potendo essere predisposta sia per applicazioni indoor che outdoor, è naturalmente più diffusa per servizi di localizzazione indoor. Un sistema completo e flessibile deve integrare o prevedere l'integrazione di differenti sistemi di localizzazione basati su diverse tecnologie. La soluzione sviluppata integra l'accesso alle diverse tecnologie di localizzazione, che il singolo dispositivo portatile può avere a disposizione, con il sistema di generazione della query in linguaggio naturale e le API esposte dal server FuLL. I sistemi di localizzazione previsti possono essere di diverso tipo dipendentemente dall'ambito fisico in cui si opera e dalle tecnologie realmente supportate dal dispositivo specifico; in ambienti outdoor la scelta più flessibile è quella basata su GPS perché cominciano ad essere sempre più numerosi i dispositivi portatili (PDA e smartphone) che lo integrano o che si interfacciano facilmente con un'unità GPS esterna tramite Bluetooth. L'eventuale disponibilità di altri sistemi (localizzazione tramite Wi-Fi o per mezzo di coordinante fornite dalla rete cellulare stessa, così come è previsto dalla tecnologia UMTS) è comunque resa trasparente all'utente, così come il sistema di trasmissione per interfacciarsi al server FuLL (connessione con SMS/MMS, GPRS, bluetooth, WiFi). Il dispositivo è programmato per rilevare le tecnologie utilizzabili e seleziona la migliore e più opportuna in maniera automatica. La portabilità della soluzione proposta è stata massimizzata basandosi sul linguaggio Java, versione Micro Edition.

6- Applicazioni sperimentali

E' stato realizzato un prototipo nel quale i dati sulla rete stradale vengono integrati con un sistema di rilevamento posizione a bordo dei veicoli o in dotazione agli utenti, e con un sistema WiFi/GPRS di trasmissione delle informazioni al server di elaborazione. Per la sperimentazione, nel presente lavoro è stato utilizzato un cellulare smartphone Nokia N73. La localizzazione utente avviene tramite latitudine e longitudine attuali dello smartphone Java. Sul server funzionano i Decision Support System basati sui modelli di simulazione dinamica, integrati con i moduli di interfaccia in linguaggio naturale con capacità di analisi sintattica e semantica. I dati di test sono reali e riguardano le reti stradali delle città di Bologna e Catania; sono stati forniti dalle Amministrazioni Provinciali, che si ringraziano per la collaborazione.

6.1 – Città di Bologna

Per Bologna il DB di riferimento è costituito da una porzione dei dati attualmente gestiti dalla piattaforma FuLL per la Provincia di Bologna (strade, stazioni ferroviarie, parcheggi, distributori di carburante, musei), esteso con le tabelle dei risultati dei calcoli dei modelli di traffico, necessarie a comprendere lo stato del sistema di trasporto e per prevedere la sua evoluzione nel tempo. Il prototipo è in grado di intercettare il concetto di percorso ottimo fra la posizione attuale dell'utente ed una destinazione specificata esplicitamente; ad esempio: *qual è il distributore più vicino? Qual è il percorso più veloce/migliore/breve per la stazione?*. Viene selezionato il percorso ottimo in base all'orario in cui l'utente formula la richiesta, ovvero la fascia oraria di riferimento è sottintesa essere quella più vicina al tempo attuale di ricevimento della richiesta da parte del server. Il concetto di vicino viene inteso prendendo in considerazione il tempo di percorrenza ottimo per un utente a bordo di un veicolo stradale, calcolato dal modello dinamico, su una porzione della rete stradale di Bologna.



Figura 2 – Interfaccia utente con esempio di richiesta in linguaggio naturale e risultati

Il prototipo è inoltre in grado di selezionare i migliori percorsi possibili fra una specifica origine e più destinazioni, incrociando altri fattori desiderati dall'utente, come ad esempio il fatto che i percorsi siano nei pressi di un museo, o che le destinazioni siano nei pressi di una specifica strada. Questa possibilità deriva dalla gestione fuzzy-logic delle variabili in gioco, e dalla loro combinazione tramite operatori logici di tipo fuzzy-AND. Tutte le variabili di interesse vengono "fuzzificate" scegliendo 5 intervalli utili; ad esempio il concetto di "grado di congestione" del tratto stradale viene gestito distinguendo fra *liberi*, *molto scorrevoli*, *scorrevoli*, *trafficati* e *congestionati*.

E' quindi possibile chiedere, ad esempio: *quali sono i tratti congestionati nei pressi della Stazione Centrale? Ci sono distributori AGIP vicini ai percorsi per il museo X?*

6.2 – Città di Catania

Il prototipo è stato anche provato sui dati della applicazione sperimentale FuLL per la Provincia di Catania, contenente dati relativi al dominio Protezione Civile ed i modelli relativi a vulnerabilità e rischio sismico. In questo contesto il presente lavoro ha permesso di estendere l'applicazione esistente con funzionalità LBS, consentendo di introdurre nello scenario l'utente e la sua posizione sulla mappa. E' stato quindi possibile inserire domande in linguaggio naturale del tipo *quali sono i ponti a elevato rischio sismico vicini?*, la quale comporta un AND fuzzy fra i concetti fuzzy di rischio sismico (elevato) e distanza (breve) dalla posizione dell'operatore che formula la richiesta.

7- Conclusioni

L'introduzione del nuovo Modulo Post-Semantico di FuLL apre la strada alla gestione algoritmica delle ambivalenze del linguaggio naturale, aumentando la precisione complessiva. In prospettiva sarà importante riuscire ad affinare le capacità semantiche in modo da gestire domande in cui l'utente possa specificare esplicitamente degli orari o degli intervalli temporali.

Sarà inoltre importante la realizzazione di un prototipo sperimentale più robusto ed evoluto in collaborazione con soggetti interessati all'utilizzo della piattaforma, anche estesa e specializzata in altri ambiti e domini applicativi.

In particolare, per applicazioni critiche riguardanti la gestione delle emergenze, sono state poste le basi per un'integrazione con la piattaforma SEISMIC, sviluppata congiuntamente dal CNIT e dal Dipartimento della Protezione Civile della Regione Sicilia (Morabito et al., 2007).

Bibliografia

- Bartolini R., Caracciolo C., Giovanetti E., Marchi S., Pirrelli V. (ILC CNR) & Renso C., Spinanti L. (ISTI CNR) & Lenci A. (Dip. Linguistica Univ. Pisa) (2006), "Creation and Use of Lexicons and Ontologies for NL Interface to databases", *Atti "Fifth international conference on Language Resources and Evaluation"*, Genova.
- Bellei G., Gentile G., Papola N. (2005), "A within-day dynamic traffic assignment model for urban road networks", *Transportation Research B* 39, 1-29.
- Cafiso S., Condorelli A., Mussumeci G. (2005), "Functional analysis of the urban road network in seismic emergency. GIS application on Catania city.", volume "Seismic prevention of damage. A case study in a Mediterranean city" (editor M.Maugeri) series: *Advances in earthquake engineering*, vol.14 (ISBN:1-84564-004-7).
- Condorelli A., Marletta O., Mussumeci G. (DICA Univ. Catania) & Bombara M., Calì D., Calì I., Cartelli V., Tropea G. (BC S.r.l.) (2006), "Interrogazioni in linguaggio naturale di banche dati GIS della Protezione Civile", *Atti della 10a Conferenza Nazionale ASITA 14-17 Novembre 2006*, Bolzano.
- D'Andrea A., Cafiso S., Condorelli A. (2005), "Methodological considerations for the evaluation of seismic risk on road network", *Pageoph (Pure and Applied Geophysics)*, volume 162 No.4, pp. 767-782, Pageoph Topical Volumes "Messina Seismological Observatory Memoriale Volume" editing di A.Bottari, D.Mayer-Rosa, J.Ibanez, M.Maugeri; edizioni Birkhauser, Berlino.
- Gentile, G., Meschini, L., Papola, N. (2006) Spillback congestion in dynamic traffic assignment: a macroscopic flow model with time-varying bottlenecks. *Transportation Research B* 39, accettato per la pubblicazione, in stampa.
- Giovanetti E., Masserotti M.V., Renso C., Spinanti L. (ISTI CNR) & Bombara M., Calì D., Calì I., Tropea G. (BC S.r.l.) (2006), "Interrogazioni in linguaggio naturale a basi dati eterogenee: l'ontologia del sistema FuLL nei GIS", *Atti della Conferenza Tematica AMFM 2006 - 21-22 Settembre 2006*, Roma.
- Morabito G., Narcisi R., Palazzo S., Scoto F., Cocina S., Formosa F. (2007), "SEISMIC: a System for Environmental Integrated Sensing, Monitoring and Interactive Control", *Proc. of WRECOM 2007*, Rome, October 1-2, 2007.