

Interoperabilità dei dati geografici in ambiente WebGIS

Francesco Pirotti (*), Dario Dominico (**), Alberto Guarnieri (*), Antonio Vettore (*)

(*) Centro Interdipartimentale di Cartografia, Fotogrammetria, Telerilevamento e GIS
Università degli Studi di Padova - via dell'Università 16 - 35020 Legnaro (PD)
email: (francesco.pirotti antonio.vettore alberto.guarnieri)@unipd.it

(**) Servizio Cartografia e Gis, Provincia di Brescia, via Milano 13 25126 Brescia
email: ddominico@provincia.brescia.it

Sommario

Per ottimizzare la risorsa dei dati geografici è ormai fondamentale la condivisione ed armonizzazione degli stessi. La rappresentazione del territorio sta diventando sempre più ricca grazie a tecnologie che forniscono grandi quantità di informazioni rappresentate sia mediante il modello raster che mediante il modello vettoriale. Questo ha fortemente accelerato il settore ricerca e sviluppo nell'ambito dei sistemi *backend* webgis per poter condividere dati geografici in tempo reale avvantaggiandosi della velocità della rete e dei sistemi informativi utilizzati per la visualizzazione. Allo stato dell'arte ci sono diverse soluzioni che consentono di applicare strategie di accesso sia per la consultazione che per l'interrogazione e la modifica del dato da remoto. In questa nota si illustrano le novità riguardanti le metodologie di interoperabilità dei dati geografici, in particolar modo per mezzo di sistemi webgis, illustrando i diversi approcci sviluppati ed enumerandone i corrispondenti vantaggi. Vengono anche descritti gli sforzi fatti a livello nazionale ed internazionale per definire standard di rappresentazione e trasmissione dei dati condivisi dai vari *stakeholders* (enti pubblici, utenti privati, professionisti).

Verrà riportato un esempio operativo nell'ambito di una collaborazione tra il portale della Provincia di Brescia ed il portale multifunzione di COGEME Informatica che integra due diversi portali condividendo dati mediante una soluzione che utilizza protocolli standard di servizi per mantenere uno sviluppo separato dei due sistemi conservando livelli informativi in comune.

Abstract

To optimize geographic data distribution collaborative strategies are now quite common. Webgis portals and similar internet-based solutions, give users the possibility to access a wide range of such data. Both vector and raster type models are shared using such portals. The growing availability of this information, calls for the development of shared protocols to manage, store and access the geographic data. To this day there are several approaches for sharing data over different web servers and over end users. This paper will give a short presentation of the state of the art of these approaches and the illustrate a study case from a project between two web portals which provide cadastre and town-scale registry information.

Introduzione

La quantità di dati geografici disponibile ad oggi è in continuo aumento e questo ha portato alla necessità di trovare valide strategie di condivisione e trasferimento dei dati, come anche di metodi di modifica ed elaborazione che seguano protocolli standard.

I protocolli di servizio per l'accesso ai dati geografici tramite la rete sono stati implementati dal Open Geospatial Consortium (OGC) a partire dal 1997 con l'obiettivo di fornire degli standard di riferimento per la richiesta di dati geografici attraverso la rete (Doyle, 1997). Il lavoro del OGS è risultato in una serie di comandi standard per accedere a servizi di tipo diverso (Cuthbert, 1998) di cui si illustrano in seguito i più comunemente utilizzati (Open Geospatial Consortium 2010a,2010b):

1. WMS - *Web Map Service*: fornisce immagini (raster) create da dati geografici. Un profilo di WMS proposto da OSGeo – Open Geospatial Foundation, (2010), denominato WMS-C, indica ritagli standard memorizzati in spazio *cache*. L'OSGeo (2010) ha anche compilato le specifiche TMS – *Tile Map Service* per la standardizzazione dell'organizzazione di questo approccio.
2. WFS - *Web Feature Service*: fornisce e permette di modificare (WFS-T) dati vettoriali tramite GML/XML.
3. WCS - *Web Coverage Service*: fornisce oggetti informative su un'area.
4. WPS - *Web Processing Service*: servizio di accesso a funzioni di elaborazione su una macchina remota.
5. CSW - *Web Catalog Service*: fornisce accesso ad informazioni di catalogo.
6. SFS - *Simple Features – SQL*.
7. GML - *Geography Markup Language*: formato XML per informazioni di tipo geografico.
8. KML - *Keyhole Markup Language*: format basato sullo schema XML per etichettature e visualizzazioni di dati geografici su piattaforme gis in 2D o 3D.

Per quanto riguarda i dati geografici che devono essere condivisi e caricati tramite internet, i fattori limitanti i tempi di caricamento possono essere presenti in uno dei seguenti passaggi:

1. l'elaborazione dei dati per produrre il formato da servire alla rete (i.e. da file in formato shape in XML per i vettoriali, oppure da un'immagine GeoTIFF ad una certa risoluzione ad immagini in formato jpeg/png/ecc a risoluzione diversa per i dati raster);
2. la messa in rete dei dati tramite protocollo http o simili grazie a applicativo *web server* dedicato;
3. la trasmissione dati attraverso la rete, quindi la velocità del *provider* e della rete dell'utente oltre che la velocità dell'*hardware* di rete dell'utente;
4. l'elaborazione del dato ottenuto dalla rete e disegno su schermo (*rendering*) dell'utente che visualizza il dato.

La seguente nota riporta uno studio sui diversi software che possono leggere i dati forniti da questi servizi, oltre che una discussione critica delle varie strategie ad oggi utilizzate per fornire degli strumenti che consentono di operare, tramite rete, su dati distribuiti da servizi dedicati.

Materiali e metodi

Per questa sperimentazione sono stati utilizzati dati provenienti da due sistemi definibili come server di mappe, ovvero due server che forniscono informazioni geografiche agli utenti nella rete internet. Il primo server, presso COGEME Informatica, ed il secondo, gestito dalla Provincia di Brescia. Essi forniscono dati a livello comunale e sovra-comunale per l'accesso a informazioni di interesse gestionale, come i piani urbanistici e la cartografia nel formato database topografico. I protocolli di servizio per dati geografici utilizzati sono il WMS ed il WFS rispettivamente nella loro versione 1.1.1 e 1.1.0.

Vengono utilizzati due *dataset* di riferimento uno riguardante dati vettoriali, gestiti sia tramite WMS che tramite WFS, ed uno riguardante dati raster gestiti tramite WMS e WMSC. L'area

considerata è un sottoinsieme di comuni della regione Lombardia, per lo più situati nella provincia di Brescia, Pavia e Crema. In [tabella 1](#) è presente una descrizione delle caratteristiche dei dati. Il dato vettoriale viene testato utilizzando diverse tipologie di dati (punto, linea e poligono) aventi dimensioni diverse per valutare quali possono essere i limiti nel caricamento e nella gestione. I dati vettoriali da parte del server COGEME Informatica sono in formato Postgis in tabelle adeguatamente indicizzate in un database Postgresql versione 8.4, mentre da parte del server della Provincia di Brescia sono nel geodatabase gestito tramite ArcSde ed Oracle. La pubblicazione dei servizi WMS nello standard Ogc versione 1.3.0 avviene tramite ArcGis Server versione 9.3.1. Per quanto riguarda il modello raster sono state utilizzate ortofoto della zona a risoluzione 0.25 m ottenute durante una ripresa aerea nel 2009, rielaborate per ottimizzarne la lettura e l'elaborazione. La rielaborazione ha riguardato la creazione di GeoTIFF con ricampionamenti a varie scale per creare piramidi di *overviews* di veloce caricamento. In [figura 1](#) sono visibili le ortofoto con sovrapposti i confini amministrativi caricati da servizio WFS tramite portale webgis.

Tabella 1 – Dati utilizzati per la fase di valutazione dei metodi di condivisione (DBT=Database Topografico).

Nominativo	Tipo	n° vertici/elementi	Descrizione
CA	Poligono	1595923/1567	Confini amministrativi da Regione Lombardia
CAS	Poligono	214390/1567	Confini amministrativi semplificati*
Acc. Est.	Punto	27307	Accessi esterni da classe DBT
Incidenti	Punto	1613	Posizione di incidenti di veicoli**
Assi stradali	Polilinea	6376/36572	Elemento stradale da classe DBT
Ortofoto	Raster	NA	491 files, 10000x16000 pixel/file, 0.25 m risoluzione, formato GeoTIF c/ <i>overviews</i> .

* Semplificati mediante funzione “simplify” nel gruppo di funzioni Postgis.

**Progetto in collaborazione con la Provincia di Brescia per la valutazione del rischio incidenti.

I diversi applicativi messi a confronto permettono di accedere ai servizi e si dividono essenzialmente in due gruppi; i primi tre sono applicativi desktop che permettono di caricare i dati dalla rete utilizzando i servizi testati, mentre gli altri vengono lanciati da un *browser* (e.g. Firefox, IExplorer o Safari) tramite portali webgis adeguatamente implementati. Gli applicativi testati, tutti *open source* tranne il secondo, sono riportati in [tabella 2](#), con le loro caratteristiche. Nei risultati si riporterà un *benchmarking* delle velocità di caricamento dei diversi dati condivisi con i servizi in questione.

Tabella 2 – Gli applicativi utilizzati per i test comparativi.

Applicativo	Caratteristiche
¹ GvSIG	Sviluppato in JAVA – desktop – largo spettro di funzionalità
² Gaia	Sviluppato in C++ - desktop – specifico per servizi geografici web
³ UDig	Sviluppato in JAVA – desktop – largo spettro di funzionalità
⁴ Open Layers	Sviluppato in Javascript – web-based – specifico per servizi geografici <i>thick-client</i>
⁵ Open Scales	Sviluppato in ActionScript – web-based – specifico per servizi geografici <i>thick-client</i>
⁶ Mapserver	Sviluppato in C – web-based – specifico per servizi geografici <i>thin-client</i>

Thick-client intende che il PC dell'utente si prende carico di buona parte dell'elaborazione (e.g. rendering, interrogazioni, tematizzazione del livello); *thin-client* al contrario intende che questi vengono eseguiti dal server e solo il risultato viene fornito all'utente.

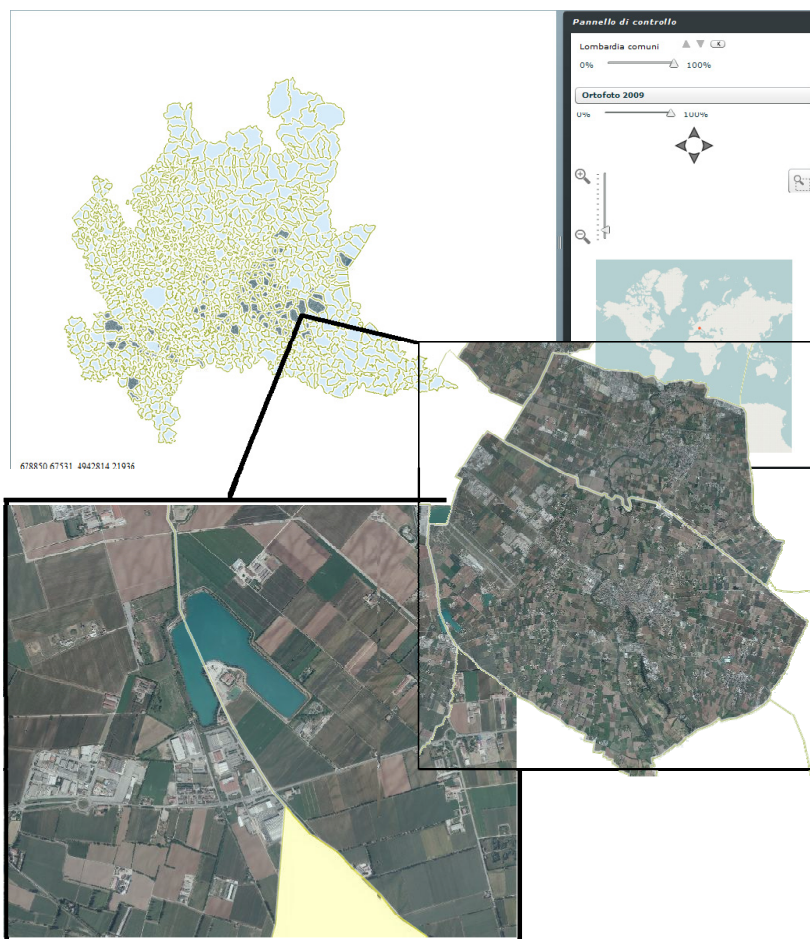


Figura 1 – Confini Amministrativi della Regione Lombardia con tutte le ortofoto da servizio WMS – applicativo OpenScales.

Tabella 3 – Tempi di caricamento (secondi).

Applicativo Elemento	1 GvSIG	2 Gaia	3 UDig	4 OpenLayers	5 OpenScales	6 Mapserver
CA	?	?	?*	396	302	9
CAS	21	16	12	63	34	4
Acc. Est.	15	15	16	52	58	3
Incidenti	4	5	4	4	3	2
Assi stradali	31	30	24	95	62	4
Ortofoto ¹	9	8	2 ¹	2 ¹	2 ¹	3

¹ Viene utilizzato il meccanismo Tile Mapping Service (TMS) con cache su disco (Tilecache, 2010) per tutti tranne che per GvSIG e Gaia che non lo supportano.

* Non riesce caricare tutti i comuni, ma a scale più grandi (x3) visualizza i comuni compresi nella vista.

? Caricamento non riuscito, applicazione chiusa dopo 5 minuti.

Risultati

In [tabella 3](#) vengono riportati i tempi di caricamento nelle diverse combinazioni tra dati e applicativi. Questo schema permette di avere un'idea di quali siano i limiti che sono posti ad oggi all'interoperabilità nel riguardo di questi dati. Per quanto riguarda i dati di tipo raster, le velocità di caricamento sono molto elevate grazie ad accorgimenti e strumenti adeguati come l'organizzazione dei files di immagine in squadri regolari e l'ulteriore ricampionamento in squadri a diversa risoluzione grazie a servizi di tipo WMS-C/TMS; nella sezione successiva viene ulteriormente discusso questo aspetto. Dalla tabella si evince invece come la difficoltà sta nella rappresentazione dei dati vettoriali. Questi, infatti, vengono disegnati su schermo – passaggio fondamentale per eventuali successive analisi/operazioni – in base al numero di vertici che compongono ogni oggetto, a prescindere dalla scala effettiva di visualizzazione. Questo porta ad un vantaggio solo nel caso il numero totale di oggetti/vertici sia limitato. Circa 10^3 vertici in un PC standard ad oggi (i.e. 4Gb RAM e processore 2.53 GHz e velocità di caricamento dati di una rete ADSL) hanno tempi di caricamento accettabili. Fino a 10^4 vertici i tempi sono sostenibili ma non confortevoli, mentre oltre si hanno serie difficoltà.

Discussione e conclusioni

Questo lavoro non vuole essere una digressione esaustiva sullo stato dell'arte di questi servizi e della loro performance, ma una valutazione sugli approcci comunemente utilizzati per operare su dati geografici. Dai risultati ottenuti si può vedere come la gestione di dati raster mediante WMS supportato dal profilo WMS-C può usufruire di molti strumenti che rendono la velocità di caricamento elevata. Come è già noto i modelli raster possono essere messi a disposizione e caricati in modo relativamente veloce, ma non hanno la capacità di fornire i dati associati che invece possono essere forniti con i modelli vettoriali. L'approccio di fornire dati di tipo vettoriale mediante WMS può essere vantaggioso dal punto di vista della velocità, ma porta come svantaggi a dover ricorrere a metodi più complessi per l'interrogazione, a non poter essere modificata direttamente ed inoltre ad avere una qualità di rappresentazione più bassa.

Servizi thin-clint basati su Javascript come OpenLayers hanno il vantaggio di essere relativamente semplici da implementare, ma hanno come limite il *rendering* di un numero elevato di oggetti vettoriali.

Openscales è basata su flash, quindi un plugin aggiuntivo proprietario, con le conseguenze che questo comporta. Nonostante tutto permette tempi di caricamento leggermente più accessibili per quanto riguarda dimensioni significative di dati vettoriali. Anche la velocità di navigazione/interazione con un numero elevato di elementi caricati è stato osservato essere più confortevole.

Per tutti e due gli applicativi *web-based* comunque, quando si devono disegnare più di qualche migliaio di vertici per la rappresentazione vettoriale, i tempi di caricamento diventano insostenibili, ed inoltre la memoria necessaria rallenta le altre interazioni. E' stato inoltre osservato che i tempi di *rendering* degli elementi dipendono da come questi vengono rappresentati. Nel caso di elementi di tipo punto, nel caso di OpenLayers ed OpenScales, ad esempio, i tempi sono significativamente lunghi in quanto la rappresentazione standard viene fatta con un cerchio con bordo pieno e riempimento trasparente. La creazione di questo oggetto moltiplicata per il numero di punti porta a tempi lunghi di caricamento, oltre che un rallentamento della navigazione in sé. Un'attenta valutazione dei modelli di rappresentazione (i.e. modelli più elementari) potrebbe portare ad un significativo vantaggio.

I tempi di creazione di viste più veloci sono stati rilevati su UMN Mapserver. La sua velocità consiste nel riportare tutti i livelli in un raster finale che viene fornito all'utente. Questo elimina il problema di dover caricare un eventuale alto numero di oggetti (i.e. vertici di elementi vettoriali) tutti sulla vista dell'utente, come invece viene fatto nei due applicativi sopra menzionati. UMN Mapserver fornisce un robusto sistema di creazione di mappe e viste in modo estremamente veloce,

ma lascia al programmatore la parte di implementazione degli strumenti di interfacciamento utente/sistema.

In futuro è probabile che la gestione di questi dati segua una strategia “*peer-to-peer*” o simili ovvero che i dati potranno essere distribuiti, anche in copie ridondanti, su più server e condivisi a blocchi a seconda della disponibilità di banda della rete e carico del server. Inoltre sarà sempre più importante un approccio di segmentazione geo-spaziale dei dati vettoriali, ovvero un sistema che decida quali dati fornire a seconda dell’area visualizzata e della dimensione dei dati da caricare. Ad esempio, in uno scenario dove un intero comune ha potenzialmente migliaia di curve di livello, non si può pensare di farle caricare tutte quando l’utente si collega con una vista che comprenda tutto il comune. Un applicativo dovrà decidere di visualizzare una frazione del totale, ma questa frazione dovrà essere campionata in modo tale da dare una corretta visione d’insieme (e.g. una curva ogni dieci). Ad oggi i servizi per i dati vettoriali permettono una strategia basata sul *bounding box* (bbox) che caricano solo gli elementi all’interno della vista caricata dall’utente, oppure permettono di limitare arbitrariamente il numero di elementi da caricare, senza però eseguire alcun ricampionamento “intelligente” (i.e. caricano i primi n elementi trovati, anche se questi sono distribuiti tutti in un unico punto del territorio, mancando quindi di rappresentare uniformemente la zona).

Riferimenti bibliografici

Cuthbert, A. (1998), User Interaction with Geospatial Data. Open GIS Consortium

Doyle, Allan (1997), WWW Mapping Framework. Open GIS Consortium

Open Geospatial Consortium (2010a), <http://www.opengeospatial.org/standards/wfs>. Accesso: 2010-07-26.

Open Geospatial Consortium (2010b), <http://www.opengeospatial.org/standards/wms>. Accesso: 2010-07-26.

Tilecache (2010), <http://tilecache.org/> Accesso: 2010-07-22.

OSGeo – Open Source Geospatial Foundation (2010), <http://www.osgeo.org/> Accesso: 2010-07-22.