

GEO MEDIA

www.rivistageomedia.it

Rivista bimestrale - anno 14 - Numero 4/2010
Sped. in abb. postale 70% - Filiale di Roma

La prima rivista italiana di
geomatich e geografia intelligente

N°4
2010



GENERAZIONE PDA USABILITÀ E TECNOLOGIE

- La gestione del Rischio nei Trasporti
Scenari e Strumenti per la Mitigazione
del Rischio
- iPhone e Applicazioni Geomatiche
Una Guida alle Applicazioni più
Interessanti
- L'eccellenza Cartografica è Marchigiana
Il Premio Geoportali 2009 va
al Comune di Jesi
- MDVLab e INGV: in volo sull'Antartide
Droni e Telerilevamento al Servizio
dell'ambiente

I Cluster Geografici

di Maurizio Rosina

Un interessante esempio di uso di gruppi di informazioni (cluster) geografici che fa riferimento ai database alfanumerici georiferibili della pubblica amministrazione che sono cresciuti ad un ritmo impressionante negli ultimi anni ci porta nel mondo del progetto GEOPOI nato dal laboratorio di ricerca della SOGEI mostrandoci gli ultimi risultati di una affascinante ricerca.

«In linguistica si chiama effetto 'rumore': succede quando il troppo stroppia, quando sei sommerso da una valanga di informazioni e il messaggio fatica a farsi strada, come fosse bloccato da un gigantesco ingorgo. I massmediologi parlano di 'data smog' [...]» (Grasso 2000).

Il contesto in cui la frase precedente è stata scritta è quello dei media radio-televisivi ma la loro valenza è molto più generale, come più generale deve essere considerato il concetto di *informazione* presente in tali affermazioni. In qualsiasi campo e disciplina l'eccesso di informazioni genera 'rumore' e quando ciò accade è spesso conveniente suddividere le tante, troppe informazioni in gruppi; questi gruppi vengono generalmente chiamati, soprattutto nella comunità tecnico-scientifica anglofona, *cluster*. Ciascun cluster contiene e raggruppa solo quelle entità/informazioni che soddisfano degli specifici criteri. L'esistenza stessa di un cluster quindi, oltre alla funzione di raggruppamento, 'segnala' che vi sono entità informative di base che soddisfano gli specifici criteri definiti per appartenere a quel cluster.

I Geographic Cluster

L'articolo tratterà esclusivamente di entità (rappresentative di oggetti, eventi o valori) singolarmente georiferibili su mappa tramite una propria (unica) coppia di coordinate, ed i criteri di classificazione di queste entità saranno quelli di vicinanza geografica e di correlazione/interconnessione. Ogni entità georiferita sarà quindi rappresentata da un POI (*Point Of Interest*). Tali entità, assoggettate ai criteri sopra citati, verranno raggruppate in gruppi (i cluster), e data la caratteristica di georiferibilità delle entità componenti, chiameremo i nostri gruppi *Geographic Cluster*. In conseguenza ai criteri di classificazione attivati sulle entità, vengono individuati e si farà riferimento a due macro-categorie di *Geographic Cluster*:

- *Geographic Cluster Territoriale*: la condizione primaria e sempre necessaria delle entità costituenti sarà il loro appartenere tutte ad uno stesso specifico ambito territoriale (ad esempio essere tutte contenute entro uno specifico confine amministrativo o entro una porzione di territorio contenuta entro tale confine). Sulla base della tipologia dell'ambito territoriale distingueremo, in questa prima macro-categoria, tra:
 - *Geographic Cluster Territoriale Amministrativo* (cluster di 'tipo A'): gruppo di entità contenute entro uno specifico confine/ambito amministrativo;
 - *Geographic Cluster Territoriale Sub-Amministrativo*

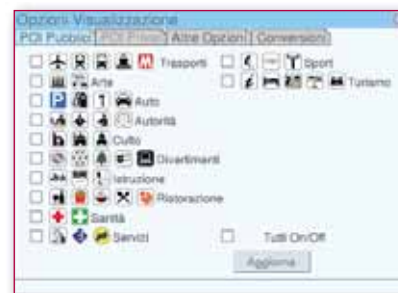
(cluster di 'tipo S'): gruppo di entità contenute entro una porzione di territorio distinguibile come un livello sub-amministrativo, ovvero presenti in una porzione di territorio che non contiene totalmente uno specifico e predefinito confine/ambito amministrativo;

- *Geographic Cluster Funzionale* (cluster di 'tipo F'): macro-categoria nella quale la sola condizione necessaria è l'esistenza di interrelazioni funzionali tra le entità costituenti, potendo ciascuna entità essere ovunque georiferita.

I Geographic Cluster in GEOPOI

GEOPOI (*GEOcoding Points Of Interest*) è un framework – ottenuto dalla ricaduta della ricerca applicata di Sogei e coperto da brevetto – che mette a disposizione di più settori della Pubblica Amministrazione (PA), in modo centralizzato e condiviso, una WUI (*Web User Interface*) univoca realizzata secondo lo standard SVG (*Scalable Vector Graphics*) emanato dal W3C (*World Wide Web Consortium*); questa interfaccia propone una visione del territorio italiano come continuum territoriale rappresentativo della viabilità. La WUI di GEOPOI è corredata da funzionalità generali di base: dinamica sui dati cartografici (*zooming, panning, ecc.*), ricerca di città, strade e relativi civici, gestione di percorsi, ecc., e ne può essere richiesta la personalizzazione, con loghi e funzionalità sviluppate ad hoc, per ciascun settore della PA. Su tale strato cartografico/viario di base sono localizzati oltre 200.000 punti di interesse *pubblici* (identificativi di hotel, cinema, teatri, musei, ecc.), la cui visione/fruizione è generalmente consentita a chiunque e su di esso i vari settori della PA possono georiferire i propri *layer* di *dati privati* (classi di POI e/o di poligoni privati) associando a ciascun punto o poligono privato più hyperlink in grado di attivare applicazioni – generalmente *web-oriented* – preesistenti e/o realizzate specificamente. L'accessibilità ai vari layer e classi di dati privati è ovviamente riservata – tramite meccanismi di accredito e di profilazione sulla base di privilegi – ai soli utenti autorizzati dagli specifici settori titolari delle varie classi di dati privati.

Figura 1 – La gestione in classi e sottoclassi (le icone) dei POI pubblici in GEOPOI.



Ciascuna entità puntuale in GEOPOI è rappresentata da un POI, distinguibile dalla propria specifica icona, appartenente ad una sottoclasse a sua volta contenuta in una classe di POI (Figura 1).

Nel framework GEOPOI il problema dei cluster è affrontato secondo le logiche di classificazione e di definizione delle tipologie dei cluster precedentemente esposte.

I Cluster di tipo territoriale in GEOPOI

Per la costituzione e la gestione dei cluster di tipo A (Cluster Territoriali Amministrativi) è necessario poter far riferimento a limiti amministrativi riconosciuti, quanto più possibile univoci e con l'ulteriore capacità di 'tassellare' l'intero territorio nazionale senza discontinuità e senza *overlap* tra i vari poligoni costituenti i tasselli stessi. Confini amministrativi con tali caratteristiche e con crescente grado di granularità dei poligoni costituenti, sono quelli relativi rispettivamente a Regioni, Province, Comuni ed alle sezioni di censimento ISTAT. In figura 2 viene proposta la piramide delle tassellazioni amministrativo/geografiche italiane, ordinata, dall'alto verso il basso, per livelli di granularità crescente dei poligoni costituenti ciascuna tassellazione. La piramide propone, inoltre, il livello di clusterizzazione primario di ingresso (*entry level*) più idoneo per ciascuna PAC o PAL. Il risultato in GEOPOI di una gestione di cluster di tipo A a livello regionale risulta essere quello proposto in figura 3.

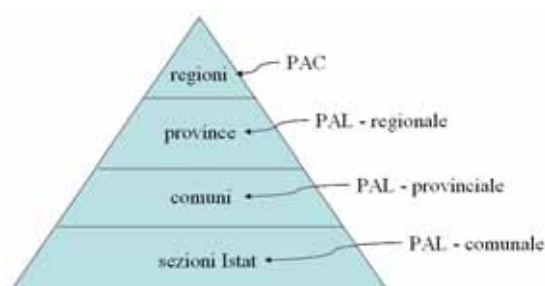


Figura 2 – PAC/PAL: la piramide delle tassellazioni amministrativo/geografiche italiane, ordinata, dall'alto verso il basso, per livelli di granularità crescente dei poligoni costituenti ciascuna tassellazione, ed i confini/ambiti amministrativi primari (*entry level*) di clusterizzazione per le PAC/PAL.

I cluster in GEOPOI presentano due specificità, spesso richieste dagli utenti e sovente presenti nelle clusterizzazioni di natura geografica: l'indicazione del numero delle entità/POI componenti proposta direttamente sull'icona di ciascun simbolo di cluster e la rappresentazione dei cluster con icone di differenti colori. In GEOPOI i tre differenti colori delle icone sono rappresentativi della percentuale di POI che ciascun cluster raggruppa rispetto al cluster con la massima numerosità di POI presente nella porzione di territorio in esame. Ad esempio detto *M* il cluster con la massima numerosità di POI tra quelli presenti nella porzione di territorio in esame, ad esso viene attribuita l'icona rossa; vengono quindi attribuite le ulteriori icone rosse e celeste ai vari altri cluster che raggruppano rispettivamente un numero di entità maggiore del 66% e minore del 33% rispetto a quelle di *M* (Figura 4). Analoga gestione delle icone dei cluster si ha a livello provinciale (Figura 5).

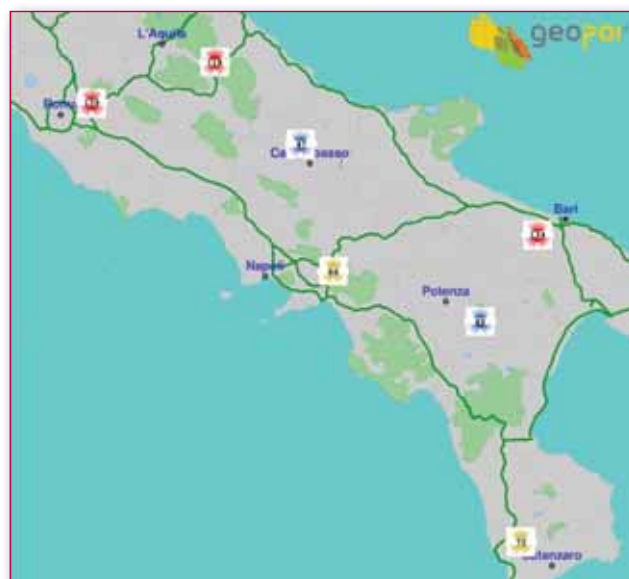


Figura 4 – I cluster con massima, intermedia e minima numerosità dinamicamente automaticamente ridefiniti – sempre a livello regionale – al variare, rispetto alla precedente figura 3, della porzione di territorio in esame.



Figura 3 – Cluster di tipo A a livello regionale in GEOPOI.

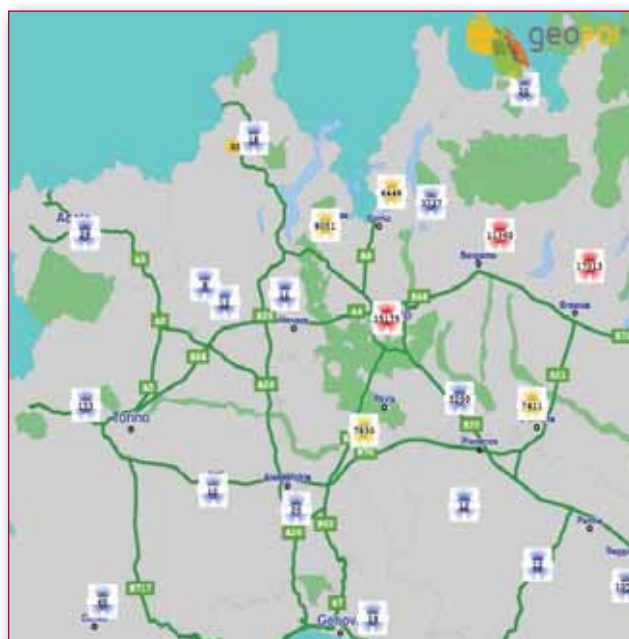


Figura 5 – Cluster di tipo A a livello provinciale in GEOPOI.

Generalmente non è necessario proporre clusterizzazioni per tutti i livelli sottostanti il livello primario di ingresso, ovvero proporre cluster di tipo A riferiti ad ambiti amministrativi sempre più ristretti. Sulla base delle competenze e delle necessità della specifica PA, risulta conveniente che ciascuna Amministrazione individui e selezioni, rispetto alla piramide proposta, il livello amministrativo sottostante il proprio entry level in grado di proporre una *granularità dei poligoni ritenuta adeguata*. Tra questi due livelli si opererà tramite cluster di tipo A – ovvero tramite clusterizzazione su ambiti amministrativi. Al di sotto del livello amministrativo con la massima granularità di tassellazione selezionato si proporranno cluster di tipo S, ovvero cluster territoriali di tipo Sub-Amministrativo (Figura 6).



Figura 6 – La situazione in cui una PAC individua quale livello amministrativo primario (entry level) il regionale e come livello amministrativo (sottostante) di massima granularità delle tassellazioni adeguato alle proprie competenze/necessità il livello provinciale, al di sotto del quale la proposizione dei cluster avverrà per ambiti sub-amministrativi, nel caso in oggetto per ambiti sub-provinciali.

I cluster di tipo Sub-Amministrativo sono realizzati sulla base di una griglia di mattonelle (*tiles*) di eguali dimensioni le quali logicamente tassellano (a *run-time*) la porzione di territorio in esame. Per ciascuna mattonella viene proposto il relativo cluster di tipo S (Figura 7 nella quale, a solo scopo illustrativo, è stato sovrapposta la griglia che determina l'array di mattonelle). La numerosità delle mattonelle, ovvero la numerosità di righe e colonne della griglia, è gestibile in modo parametrico, sulla base delle esigenze della specifica PA.



Figura 7 – Cluster di tipo S, calcolati rispetto ad un array di 9 mattonelle.

In particolare la figura 7 evidenzia come i simboli dei cluster S si presentino variamente posizionati nell'ambito di ciascuna mattonella, in quanto ciascun simbolo di cluster viene georiferito nel baricentro dei POI che esso raggruppa.

I Cluster di tipo Funzionale in GEOPOL

La gestione dei cluster di tipo F in GEOPOL è una *composizione e specializzazione* di funzionalità più generali relative ai calcoli a run-time, sia della stella di distanze lineari tra un POI selezionato e tutti gli altri POI presenti nella porzione di territorio in esame che della minima copertura convessa (*convex hull*) dell'insieme di punti presenti nella porzione di territorio in esame. Lo specializzare la composizione di queste due funzionalità all'insieme di entità costituenti un cluster F conduce alla rappresentazione proposta in figura 8. I cluster di tipo F possono essere composti anche da entità rappresentate da POI appartenenti a sottoclassi diverse, sempre rispettando il vincolo che ciascun cluster F sia composto da entità tra le quali sussistano interrelazioni funzionali.



Figura 8 – Il cluster F costituito da un insieme di ipotetiche antenne costituenti una rete GNSS.

Il criterio di arresto

I cluster sono di estrema utilità quando vi è un eccesso di informazioni, che generano rumore. Al di sotto di una determinata numerosità di POI non si ha rumore ma informazione immediatamente fruibile. E' quindi conveniente che venga definito un criterio di arresto della proposizione di cluster (sia di tipo A che di tipo S) basato sulla numerosità delle entità strettamente contenute nella porzione di territorio in esame. La tecnica più semplice, ma nel contempo assai efficace, è quindi quella che vede la definizione di un valore minimo di numerosità, al di sotto del quale non viene operata/proposta alcuna clusterizzazione e vengono, viceversa, proposte direttamente le singole entità georiferite. Una volta definito tale valore di soglia, una sua *frazione* propria può anche essere utilmente utilizzata come criterio di arresto in ciascuna mattonella in cui è suddiviso il territorio relativamente ai cluster di tipo S; questa frazione sarà ovviamente proporzionale al numero di mattonelle. Il criterio di arresto interessa anche le entità dei cluster F, in quanto in GEOPOL esse godono della proprietà di essere automaticamente gestite anche sotto forma di cluster A ed S, soprattutto a piccola scala. Quando incidono su vaste porzioni di territorio, la proposizione di cluster di tipo F avviene fruendo di tecniche di attivazione a run-time di ciascun cluster sulla base della selezione di una sua qualsiasi entità componente; ciò attiva la visualizzazione del cluster tramite una automatica gestione della scala di presentazione ed il filtraggio dei POI non interessati.

Cluster e grandi numeri

Quando i numeri in gioco non sono nell'ordine delle poche centinaia o migliaia di entità, ma piuttosto delle centinaia di migliaia o dei milioni di entità, allora non è proponibile – se non altro in termini di efficienza elaborativa – che la logica di clusterizzazione possa operare totalmente a run-time. Un modo alternativo, efficiente e di grande flessibilità, è operare una pre-elaborazione delle entità, clusterizzandole rispetto ai tre livelli apicali (regione, provincia e comune) della piramide descrittiva dei confini/ambiti amministrativi (Figura 2). Tale metodica gode del vantaggio che le entità gestite dalle PA derivano, in massima parte, da dati alfanumerici strutturati nei quali è presente una connotazione di attribuzione amministrativo/territoriale e/o di toponomastica. Solo se strettamente necessario l'attività di pre-elaborazione includerà anche il livello base della piramide (i confini



Figura 9 – Dati alfanumerici aggiornati h24 e pre-trattati per una efficiente gestione tramite cluster.

delle sezioni censuarie ISTAT); ciò sia perché tale estrema granularità di clusterizzazione sovente non è necessaria ed anche perché l'informazione relativa al codice della sezione ISTAT non è un dato così diffusamente e direttamente presente quanto, ad esempio, quello relativo alla toponomastica. La tecnica di pre-elaborazione permette la gestione di DB alfanumerici contenenti milioni di dati, oggetto di elevati livelli di movimentazione (inserimento, aggiornamento cancellazione) sia come percentuale *tout court* dei dati movimentati nell'arco dell'anno che come frequenza delle movimentazioni nell'arco delle 24h (Figura 9).

Conclusioni

L'effetto rumore, ovvero l'essere travolti da un eccesso di informazioni, è uno dei fenomeni che sempre più affligge le grandi organizzazioni man mano che con il progredire dell'informatizzazione esse si accorgono della numerosità e valenza del *thesaurus di informazioni e dati* che posseggono e gestiscono. Il modello di cluster presentato contribuisce ad aumentare la valenza, la fruibilità e la flessibilità di uso del framework GEOPOI, in quanto mette a disposizione degli aderenti alla community istituzionale che lo utilizzano, un nuovo potente strumento per dipanare il groviglio formato da vasti insiemi di dati, frequentemente movimentati, che incidono sul territorio. Le tipologie di cluster utilizzate permettono di organizzare in modo efficace ed efficiente le informazioni georiferibili e georiferite, consentendone una comprensione e fruizione che si adegua agli ambiti territoriali, ai compiti ed alle necessità di ciascuna Amministrazione.

Mappare qualsiasi cosa.

Ovunque. Sempre. Fino a 1 cm di precisione.

ashtech

ARVA
tec



Caratteristiche

- Accuratezza sub-metrica, decimetrica o centimetrica
- Estremamente leggero e compatto con GSM/GPRS integrato
- S.O. Windows Mobile 6.5 e fotocamera integrata da 3 Mpixel
- Impermeabile e antiurto (IPX7)
- Comunicazione estesa via Bluetooth o WiFi

MobileMapper™ 100

MobileMapper™ 100 è l'ultimo gioiello di casa Ashtech ed è l'unico GPS presente sul mercato che si distingue per la sua modularità: L1 GPS, L1 o L1+L2 GPS & Glonass con le seguenti opzioni: RTK, Post-Processing e software GIS MobileMapper Field. Tutte le opzioni si possono integrare secondo le proprie necessità, attivandole semplicemente con codici di sblocco. Il tecnico GIS, partendo con lo strumento "base", può avere precisioni submetriche. Il topografo, con le opzioni Glonass ed RTK, può avere precisioni centimetriche in tempo reale. Un'altra caratteristica molto importante è il modem interno GSM/GPRS per ricevere la correzione NTRIP.

MobileMapper™ 100 è l'unico GPS portatile che può operare in RTK L1 con l'antenna integrata! La tecnologia BLADE™, proprietaria di Ashtech, permette al ricevitore di operare in condizioni difficili tipo sottobosco o canyon urbani.

A completare le caratteristiche professionali ci sono il sistema operativo Windows Mobile 6.5, la fotocamera integrata da 3 Mpixel e l'estrema robustezza per operare in ambienti ostili.

MobileMapper™ 6

MobileMapper™ 6 è uno strumento completo e ideale per acquisizione e aggiornamento dati GIS. Soddisfa le esigenze di chiunque ha bisogno di un GPS efficiente, produttivo ed economico per raccogliere dati in campo. MobileMapper™ 6, è molto facile da usare ed è pienamente compatibile con i più comuni software GIS. L'opzione Post-Processing permette di avere precisioni inferiori al metro. La fotocamera integrata da 2 Mpixel ed un microfono con altoparlante permettono di arricchire la raccolta dati con immagini e note vocali.



Caratteristiche

- Adatto ad ambienti ostili (IPX7)
- Windows Mobile 6.2
- Fotocamera digitale da 2 megapixel
- Connettività Bluetooth
- Possibilità Post-Processing

www.arvatec.it

Tel. e Fax 0331 464840 - NetFax 178 2223807

Ringraziamenti

Un doveroso ringraziamento va agli ingg. G. Caputo e F. Minuti, che hanno realizzato le funzionalità descritte, ed all'ing. A. Bottaro, che ha permesso ed incoraggiato la formulazione del modello di cluster e l'implementazione nel framework GEOPOI.

Abstract

Geographic Clusters

Over the past decade, public alphanumeric database have been growing at exceptional rate. Most of data can be georeferenced, so that is possible gaining new knowledge from such databases. The contribution of this paper is two-fold. We first present a model of *geographic clusters*, which uses only geographic and functionally data properties. The model is useful to process huge amount of public/government data, even daily upgrading. After that, we merge the model into the framework GEOPOI (GEOcoding Points Of Interest), and show some graphic map results.

Riferimenti

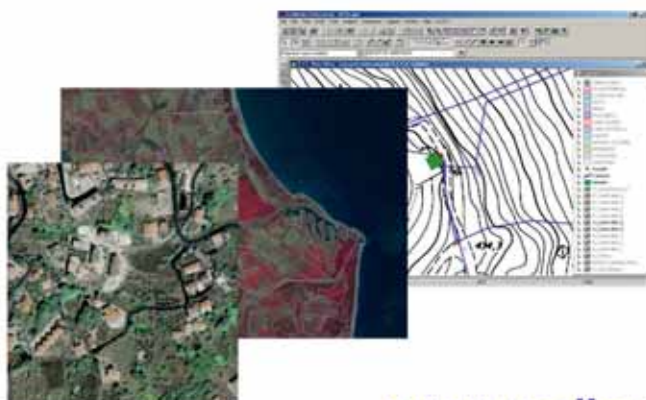
- Grasso A., *Radio e Televisione – Teorie, analisi, storie, esercizi*, pag. 198, ed. Vita e Pensiero, 2000.
- Bottaro A., Rosina M., *GEOPOI per la geocodifica dei punti di interesse*, in GEOmedia 2-2007, ed. Geo4all, Roma.
- Bottaro A., Rosina M., *GEOPOI: la cartografia dell'ipertesto*, GEOmedia 3-2007, ed. Geo4all, Roma.

Autore

MAURIZIO ROSINA

MROSINA@SOGEL.IT

SOGEL – SOCIETÀ GENERALE DI INFORMATICA
INNOVAZIONE, RICERCA E SVILUPPO



www.epsilon-italia.it

LE NOSTRE ECCELLENZE:

- FORMAZIONE SPECIALISTICA SU TECNOLOGIE GIS OPEN SOURCE
- SVILUPPO DI APPLICAZIONI GIS E WEBGIS PERSONALIZZATE
- REALIZZAZIONE DI INFRASTRUTTURE DATI TERRITORIALI (SDI) CONFORMI ALLA DIRETTIVA INSPIRE

EPSILON ITALIA SRL, VIA PASQUALI 79, 87040 MENDICINO (CS)
TEL . 0984.631949 - FAX 0984.631767 - EMAIL: info@epsilon-italia.it