Towards a R-centric architecture for multi-purpose geographical analysis on heterogeneous multi-source data

Arlette ANTONI, Thierry DHORNE and Yann LE GUYADEC

Université de Bretagne Sud
Université Européenne de Bretagne
Lab-STICC / Valoria

UseR! - 8, 9 and 10th July 2009, Rennes - France
Geographical Information Systems (GIS) are used to display, manipulate, and analyse geographical (map) data. Geographical data are data that contain a reference to a place (on the earth). Our original purpose was to (try to) provide a R-centric architecture to manage geographical information, represent geographical information, and execute geographical analysis.
Classical GIS

Collecting ───────> Analysis ───────< GIS ───────> Storage

GIS ───────> Vizualisation
GISes and Analytics

- a very interesting paper

**de Andrade Neto, Ribeiro and Fook** *Integration of Statistics and Geographic Information Systems: the R/TerraLib case*

- GIS have limited capacity to perform advanced analysis
- statistical softwares have limited access to GIS technology

★ hence the need for integration
  - full integration
  - loose coupling
  - close coupling
new evolutions need to take **time** into account

- in various ways
  - varying environments: coverage
  - real-time acquisition
  - spatio-temporal point processes
  - spatio-temporal trajectories (both GPS assisted)
  - time-moving surfaces and volumes
  - ....

- and
  - multi-sensor approaches (visible, IR, radar,....)
New Paradigm - II

- the respective part of analysis with respect to
  - collecting
  - storage
  - visualization
  is growing and growing
  - need to develop more and more tools
  - interest in collaborative design and conception
Analytics-centric GIS

Analysis

 ↔ Collecting

 ↔ Storage

 → Visualization

★ idea to become R-centric
Collecting

- static raster = images (many formats)
- dynamic raster = videos (mpeg and consistent)
- attributes (native)
extracting ⇒ collecting + analytics

- manual coordinate entry
- attributes (extracted)
- error detection
- reference coordinate
Data storage and management

- spatio-temporal databases
  - spatial query, time query, attribute query
  - time access optimization (or not)
- batch or on-line process (stream analysis)
- specific management (copy, subset, merge,...)
- changing projection system
What does R provide?

- for data storage
  - R and relational data bases
  - R and spatial data bases (PostGIS, QGIS)
  - mapproj (maps projection)
Restitution

- map design and layout (logical)
- map printing (physical)
- graphics
- videos
- metadata (generated by analysis)
What does R provide?

- for image analysis
  - adimpro: Adaptive Smoothing of Digital Images
  - biOps(GUI): Image processing and analysis analytics
  - pixmap: Bitmap Images ("Pixel Maps") interactive
  - RImageJ: R bindings for ImageJ external
  - ReadImages: Image Reading Module for R and... medical
  - rimage: Image Processing Module for R
What does R provide?

- for maps management
  - mapdata: Extra Map Databases
  - mapproj: Map Projections
  - maps: Draw Geographical Maps
  - maptools: Tools for reading and handling spatial objects
What does R provide?

- for visualization
  - rgl: 3D visualization device system (OpenGL)
  - shapefiles: Read and Write ESRI Shapefiles
  - Iplots: interactive plots
Analysis

- usual (GIS)
  - interpolation
  - connectivity, proximity and adjacency
  - map design cartography
  - bounding dissolve, spatial data overlay,
  - scaling, scrolling, moving window analysis
  - map algebra

★ new analytics
  - spatial classification
  - flow analysis
  - trajectory analysis
  - ....

→ almost everything provided in R packages
Many things more

- RPyGEo
- rgdal : links to Geospatial Data Abstraction Library
- RSAGA : links to SAGA GIS
- RgoogleMaps
- stem
- spgrass6 : interface with GRASS 6 GIS

★ the competitive work of R. Bivand!
★ how to screen and organize everything?
● get geographical video
● load it in R (sequentially)
● do some analysis
● add some layers (analytic or hand)
● export the maps
● make a movie

★ wait for some new media tools in R
Application Sequence

Sat24.com
14:30 (12:30 UTC)

Sat24.com
Conclusions

- how to validate the interest of a R-centric GIS?
  - generic acquisition systems (multi-type and multi-formats)
  - a consistent storage system (till where to go?) (Spatial SQL or not)
    - with special emphasize on the equilibrium R-storage and data-base-storage
  - normalized (extended) visualization tools
  - extension and interoperability of analytics application
    - free jungle, eastern market, or soviet?
  - real need for integration strategy (and tools)!
  - find the best compromise between end user and keen developer