Provenance Support for Quality Assessment of Maps: A User Study

Nicholas Del Rio and Paulo Pinheiro da Silva

QUALITY RESULTS

Applications deployed on cyber-infrastructures often rely on multiple data sources and distributed compute resources to access, process, and derive results. However, it is possible that non-intentional imperfections get introduced into the generation processes because of several reasons including the use of:

- low quality datasets
- data filtering techniques incompatible for the kind of map being generated
- inappropriate mapping parameters (low-resolution gridding parameters)

How can we assess the quality of results generated in complex environments such as cyber-infrastructures?

ASSESSING QUALITY WITH PROVENANCE

Information about how a result was derived, known as provenance, may help scientists to assess the quality of results such as maps. Provenance information can be decomposed into two categories:

- process meta-information – related to the generation processes
- source meta-information – related to the sources of information

USER STUDY

Twenty active scientists participated in a study to determine whether provenance is needed in order to identify and explain the quality of maps generated on cyber-infrastructures.

Scientists were given a set of maps, each generated with a different kind of imperfection. Scientists were asked to identify and explain the quality of these maps sometimes using Probe-It!, a provenance visualization tool.

RESULTS

Provided a single tail t-test at 95% confidence, we can conclusively state that scientists with access to provenance are more accurate in identifying and explaining the quality of maps than scientist without access to provenance.

FUTURE WORK

- port Probe-It! to the Web
- extend user study to include trust

USAGE SCENARIO: BUGGY SENSOR

1. We need to know how resultant maps were generated to verify the observed anomalies in maps 2 and 3

   How can we access the justification, i.e., process meta-information, of map 2 to find reasons supporting our beliefs

   Scientists were given a set of maps, each generated with a different kind of imperfection. Scientists were asked to identify and explain the quality of these maps sometimes using Probe-It!, a provenance visualization tool.

   RESULTS

   Provided a single tail t-test at 95% confidence, we can conclusively state that scientists with access to provenance are more accurate in identifying and explaining the quality of maps than scientist without access to provenance.

   FUTURE WORK

   - port Probe-It! to the Web
   - extend user study to include trust

   USAGE SCENARIO: BUGGY SENSOR

   1. We need to know how resultant maps were generated to verify the observed anomalies in maps 2 and 3

   2. Using both final results and source summaries, we can form a belief regarding the quality of source da Silva Data

   3. We access the justification, i.e., process meta-information, of map 2 to find reasons supporting our beliefs

   4. We access intermediate results, i.e., source datasets, and overlay different visualizations (2D plot, XMDV) in hopes of identifying an error

   5. Having identified da Silva data as a bad source, we look at the associated source meta-information and learn about the sensor that collected this data