

Integration of Network Measurement Data

Mahmoud Rafea

CLAES

mahmoud@claes.sci.eg

Outline

- Ontology based data integration
- Data analysis for Internet Monitoring

Ontology Based Data Integration

- *Why?*
 - *The change in the meaning of data over time: Ontological heterogeneity*
 - *The integration of multiple sources: Representational heterogeneity*
 - The abstraction and representation of the time domain differs across systems: *Heterogeneity in temporal entity (duration, segment, point)*
- How ?
 - Develop semantic integration techniques to effectively resolve temporal semantic conflicts between data sources and receivers.

Dimes Data

- Nature:
 - Represent network measurement of AS nodes connectivity
- Semantic heterogeneity problem
 - Temporal because AS nodes are dynamic. The same node number can move from one country to another
- Solution
 - Appropriately timestamp the data
 - Integrate data using Ontology
 - Develop a suitable temporal query language
- Example queries
 - What are the AS nodes of Egypt during November 2005
 - What are the countries that have AS node 123 during 2005

Data Analysis For Internet Monitoring

- Goal:
 - Detect change in Internet properties
- Types:
 - Extract different connection (in-degree and out-degree) distribution
 - Group nodes with similar feature(s) (abstraction) and monitor their size over time:
 - Geographical: nodes of a country or group of countries
 - Metric: nodes having similar number of connections