Graph Databases: A Social Network Analysis Use Case

Xavier Lopez, Senior Director, Oracle
Mark Rittman, CTO, RittmanMead
Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle’s products remains at the sole discretion of Oracle.
Property Graph & RDF Graph

Property Graph Model
- Graph Search & Analysis
- Big Data analytics
- Entity analytics

RDF Data Model
- Data federation
- Knowledge representation
- Inferencing

Social Network Analysis
- National Intelligence
- Public Safety
- Social Media search
- Marketing - Sentiment

Linked Data / Semantic Mediation
- Life Sciences
- Health Care
- Publishing
- Finance

Application Area
Graph Model
Industry Domain
Big Data Spatial and Graph (Graph)
Early Adopters of Graph Databases
Property Graph : Usage Scenarios

• Finance
  – Fraud detection, cross marketing

• Telecommunications
  – Call records analysis

• Retail
  – Recommendation, sentiment analysis

• Social
  – Network analytics, influencers, clustering

• Health Care
  – Doctor, patient, diagnosis, treatment analysis;
Modeling Customers/Products

Name: Josh  Age: 45
ID: Infant Formula  Date: 11/10
ID: Ibuprofen  Date: 11/05

Name: Lucy  Age: 27
ID: Jumper  Date: 11/01
ID: Jumper  Date: 12/12

buys

?
Modeling Social Networks

- **Name:** Anna  
  **Age:** 29  
  **Knows:** Josh

- **Name:** Mary  
  **Age:** 38

- **Name:** Josh  
  **Age:** 45

- **Name:** Tony  
  **Age:** 47

- **Name:** Lucy  
  **Age:** 54

**Relationships:**
- Anna works with Mary
- Josh knows Anna
- Tony works with Lucy
Big Data Graph Architecture

Lightning-Fast In-Memory Analytics
- YARN Container
- Standalone Server
- Embedded

Massively Scalable Graph Store
- Oracle NoSQL
- HBase
Common Graph Analysis Use Cases

**Product Recommendation**
- Recommend the most similar item purchased by similar people

**Influencer Identification**
- Find out people that are central in the given network – e.g. influencer marketing

**Community Detection**
- Identify group of people that are close to each other – e.g. target group marketing

**Graph Pattern Matching**
- Find out all the sets of entities that match to the given pattern – e.g. fraud detection
35 Social Network Analysis Algorithms

- Closeness Centrality
- Degree Centrality
- Degree Distribution
- Eigenvector Centrality
- Fattest-Path
- Hyperlink-Induced Topic Search
- SSSP
- Triangle Counting
- Random Walk with Restart
- Personalized Pagerank
- Strongly Connected Components
- Vertex Betweenness Centrality
- Weakly Connected Components
Property Graph Workflow

• Graph Data Management
  – Raw business data is converted to a graph schema
  – Horizontally Scalable: Hadoop, NoSQL

• Analysis and Exploration (in-memory analysis engine)
  – Data scientists try different ideas (algorithms) on the data
  – Flexible, interactive, iterative, small-scale (sampled), ....
Graph Visualization

Tom Sawyer Perspectives
RDF Graph: A Specialized Graph Model

**RDF Data Model**
- Data federation
- Knowledge representation
- Inferencing

**Property Graph Model**
- Graph Search & Analysis
- Big Data analytics
- Entity analytics

**Social Network Analysis**
- National Intelligence
- Public Safety
- Social Media search
- Marketing - Sentiment

**Linked Data / Semantic Mediation**
- Life Sciences
- Health Care
- Publishing
- Finance

**Application Area**
**Graph Model**
**Industry Domain**
Graph-based Metadata Layer

- W3C standard, flexible model for sparse and evolving data
- Common vocabulary enables data integration & app development
- Relational data stays in place, apps don’t need to change
Harmonizing the Enterprise and Big Data Systems

Enterprise-wide, Patient-centric, longitudinal Record System

Domain Ontologies
(business metadata + Ontologies)

Data Servers

Lab/clinical Care  Research  Content Mgmt

Data Sources / Data Types
Social Media  Medical Devices  Lab Information Systems  Subscription Services  Legacy Patient Records
Enterprise Information Harmonization

Industries

• Life Sciences
• Health Care
• Finance
• Media
• Networks & Communications
• Defense & Intelligence
Consolidated Knowledge Layer

Business Challenge
• Link database information on genes, proteins, metabolic pathways, compounds, ligands, etc. to original sources.
• Increase productivity for accessing, sharing, searching, navigating, cross-linking, analyzing internal/external data

Solution
• Semantic integration layer using RDF graph
• Rich domain-specific terminology (biology, chemistry and medicine) 1.6 M terms
• Terminology Hub: 8 GB of referential data (ontologies) that cross-reference various data repositories.
Graph Sessions at BIWA: Today

• Dismantling Criminal Networks with Graph and Spatial Visualization
  - 3:25pm, room 103

• Deploying a *Linked Data* Service at Italian Statistics Agency
  - 3:50pm, room 103

• Gain Insight into Your Graph Data: Hands-on Lab
  - 4:30pm, Bldg. 500
Social Network (Twitter) Analysis Demo

Mark Rittman, CTO
Rittman Mead
The Spatial and Graph SIG

• The SIG promotes interaction and communication that can drive the market for spatial technology and data

• Members connect and exchange knowledge via online communities and at annual conferences and events

• Meet us at the Summit
  • Morning Receptions
    • Tuesday and Wednesday / 7:45 to 8:30 a.m. / Registration Area
  • Birds of a Feather Session
    • Tuesday / 12 to 1 p.m. / Lunch Room

• Join us online
  • LinkedIn (search for “LinkedIn Oracle Spatial”)
  • Google+ (search for “Google+ Oracle Spatial”)
  • IOUG SIG (sign up for free membership through www.ioug.org)
  • OTN Spatial – Communities (search for “Oracle Spatial and Graph Community”)

• Contact the Board at oraclespatialsig@gmail.com
Resources

• Oracle Technology Network: www.oracle.com/technetwork/database/options/spatialandgraph

• blogs.oracle.com ➔ oraclespatial ➔ oracle_maps_blog ➔ bigdataspaitalgraph
Resources on Big Data Spatial and Graph

• Oracle Big Data Spatial and Graph on Oracle.com: https://www.oracle.com/database/big-data-spatial-and-graph


• Blog (technical examples and tips): https://blogs.oracle.com/bigdataspatialgraph/

• Big Data Lite Virtual Machine (a free sandbox environment to get started): http://www.oracle.com/technetwork/database/bigdata-appliance/oracle-bigdatalite-2104726.html