ORACLE®

From targeted marketing to fraud detection

How graph databases can complement relational technology

Hans Viehmann
Product Manager EMEA

HOUG 2019, Siófok, April 9th , 2019

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Graph – an important growth area for data & analytics

Gartner Identifies Top 10 Data and Analytics Technology Trends for 2019

Gartner

Trend No. 5: Graph

Graph analytics is a set of analytic techniques that allows for the exploration of relationships between entities of interest such as organizations, people and transactions.

The application of graph processing and graph DBMSs will grow at 100 percent annually through 2022 to continuously accelerate data preparation and enable more complex and adaptive data science.

Graph data stores can efficiently model, explore and query data with complex interrelationships across data silos, but the need for specialized skills has limited their adoption to date, according to Gartner.

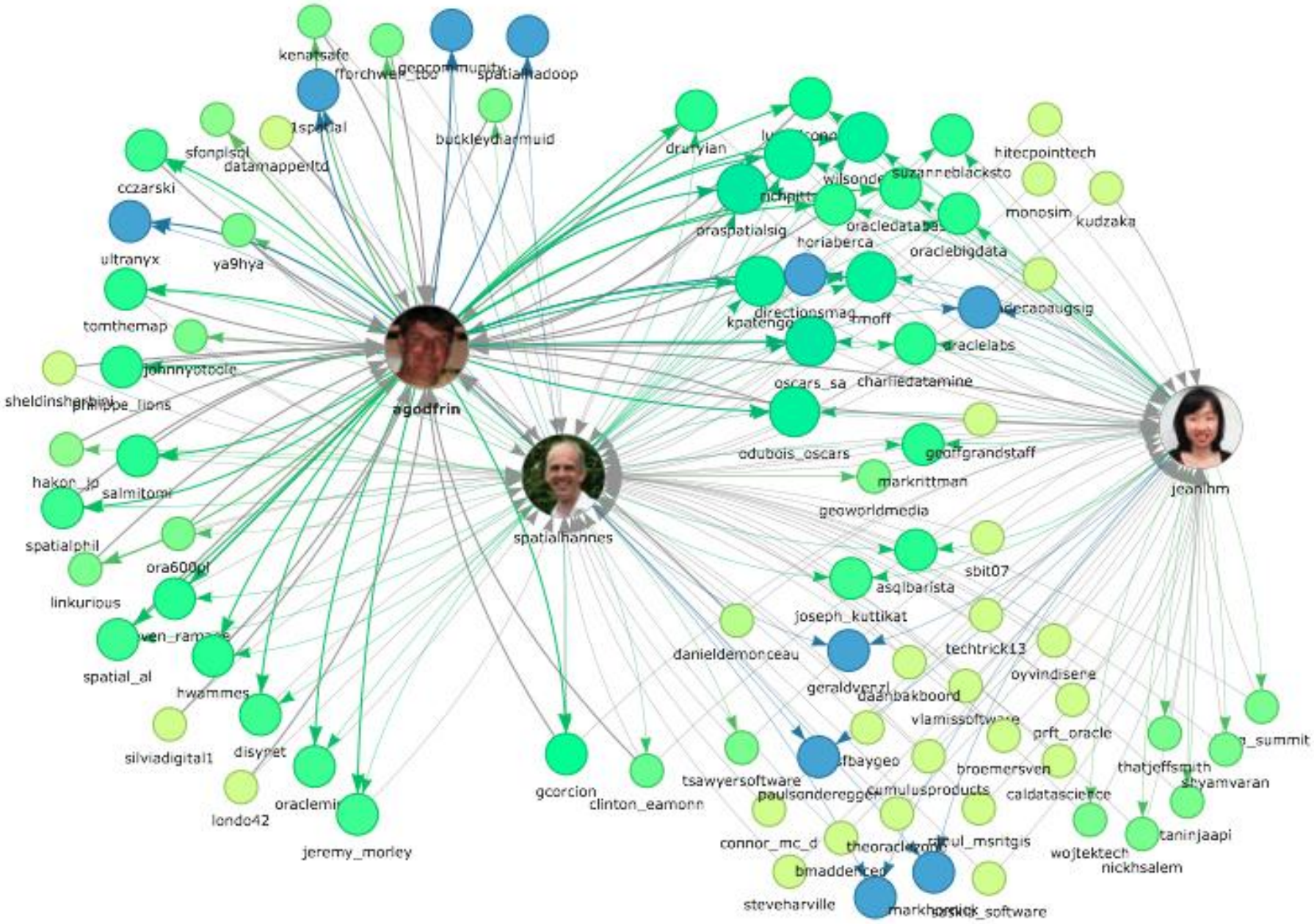
Graph analytics will grow in the next few years due to the need to ask complex questions across complex data, which is not always practical or even possible at scale using SQL queries.

Source: Gartner press release, 2/18/2019, www.gartner.com/en/newsroom/press-releases/2019-02-18-gartner-identifies-top-10-data-and-analytics-technolo

ORACLE

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. | Confidential – Oracle Internal/Restricted/Highly Restricted

- Following, no follow back
- Follower, no follow back
- Follow each other



https://twitter.jeffprod.com

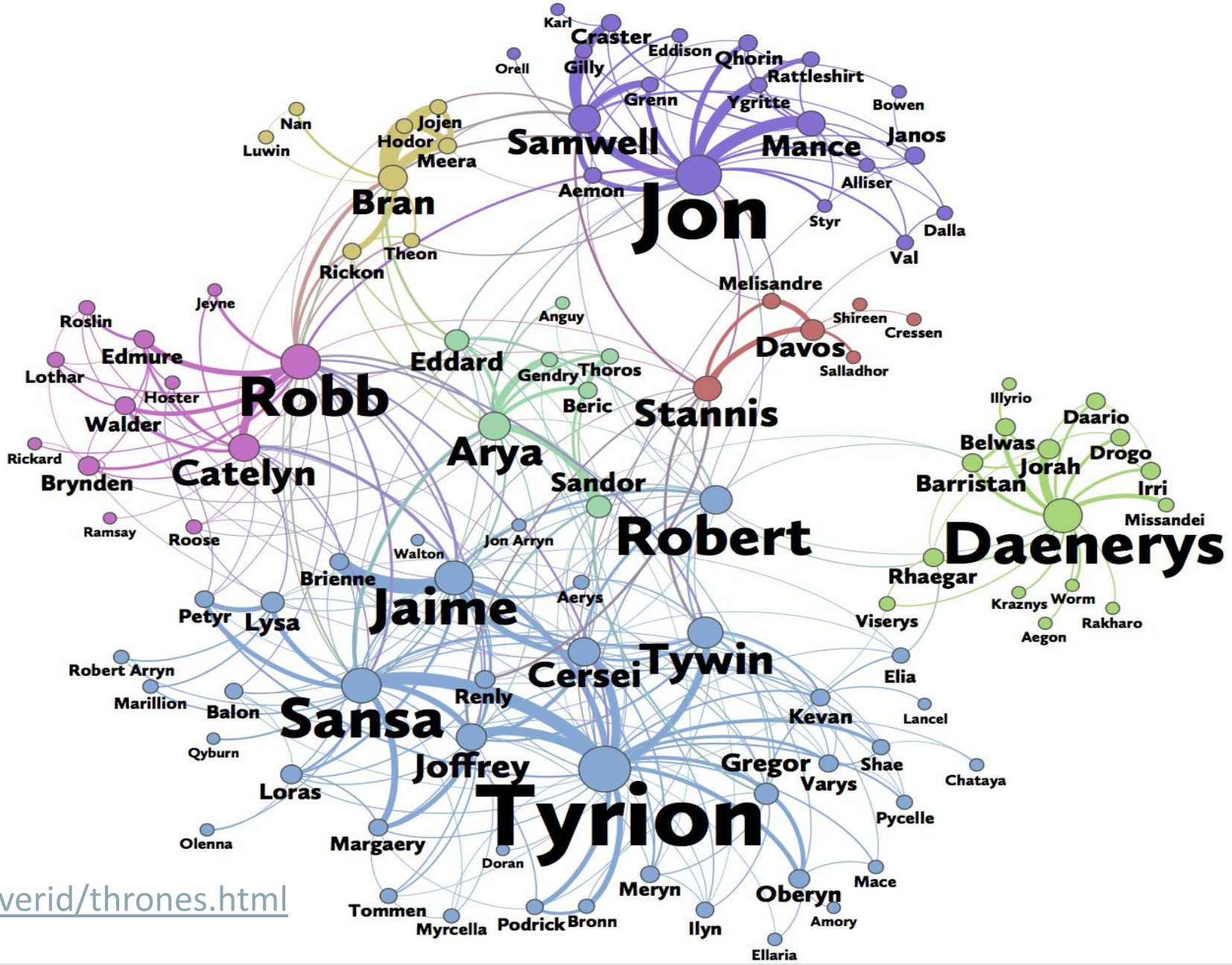
Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

A social network generated from Game of Thrones.

The color of a vertex indicates its community.

The size of a vertex corresponds to its PageRank value, and the size of its label corresponds to its betweenness centrality.

An edge's thickness represents its weight.

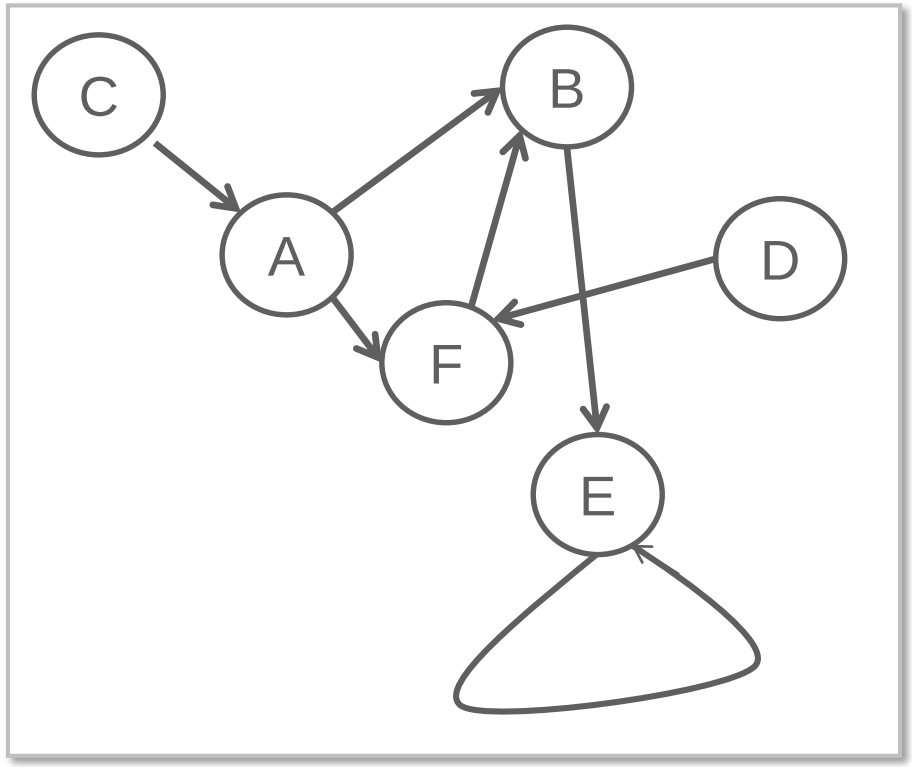


https://www.macalester.edu/~abeverid/thrones.html

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Graph Data Model

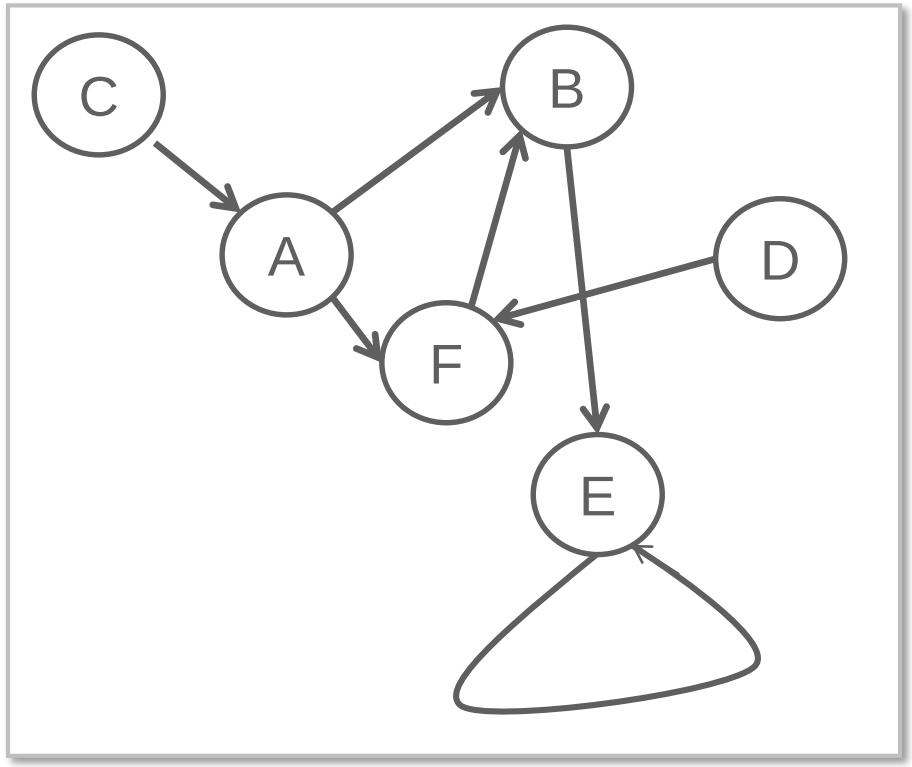
- What is a graph?
  - Data model representing entities as vertices and relationships as edges
  - Optionally including attributes
  - Also known as „linked data“
- What are typical graphs?
  - Social Networks
    - LinkedIn, facebook, Google+, ...
  - IP Networks, physical networks, ...
  - Knowledge Graphs
    - Apple SIRI, Google Knowledge Graph, ...



Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Graph Data Model

- Why are graphs popular?
  - Easy data modeling
    - „whiteboard friendly“
  - Flexible data model
    - No predefined schema, easily extensible
    - Particularly useful for sparse data
  - Insight from graphical representation
    - Intuitive visualization
  - **Enabling new kinds of analysis**
    - Overcoming some limitations in relational technology
    - Basis for Machine Learning (Neural Networks)



Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

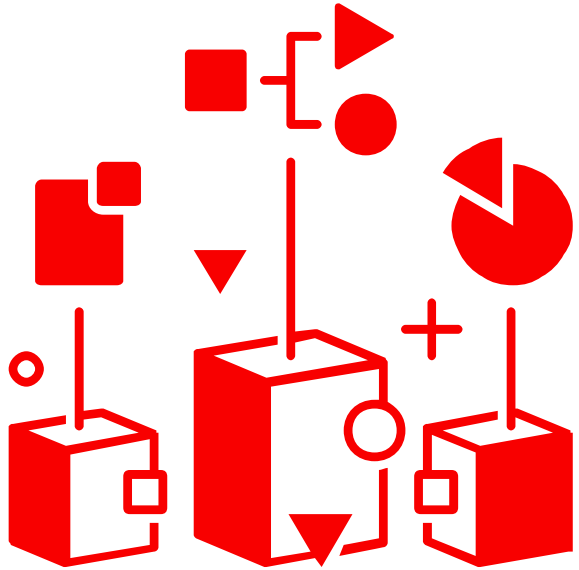
Oracle's Spatial and Graph Strategy

Enabling Spatial and Graph use cases on every platform

Oracle Database Spatial and Graph Option

Exadata
Non-Engineered Systems

Oracle Big Data Spatial and Graph



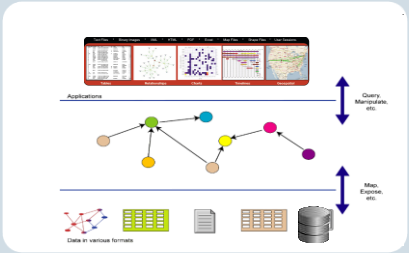
Big Data Appliance
Commodity Hadoop
Spark

Cloud Services

Database Cloud Service
Exadata Cloud Service

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

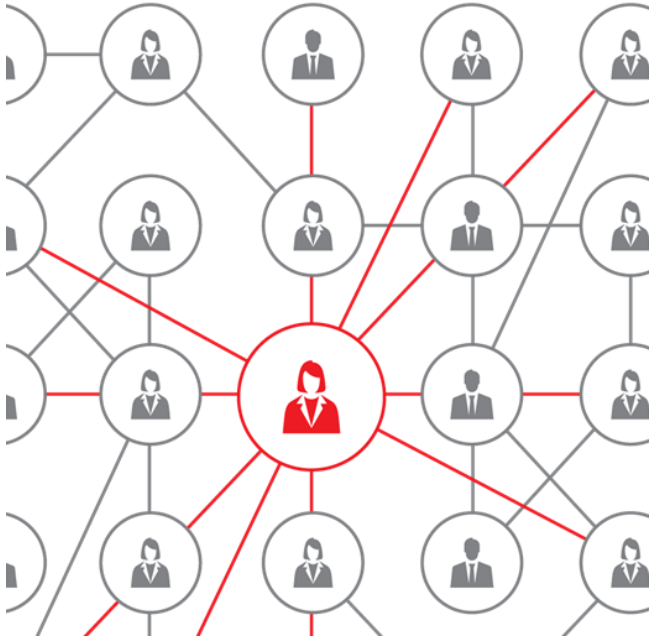
Two Graph Data Models

Use Case	Graph Model		Industry Domain
Social Network Analysis	**Property Graph Model** • Path Analytics • Social Network Analysis • Entity analytics		▪ Financial ▪ Retail, Marketing ▪ Social Media ▪ Smart Manufacturing
Linked Data Knowledge Graphs	**RDF Data Model** • Data federation • Knowledge representation		▪ Life Sciences ▪ Health Care ▪ Publishing ▪ Finance

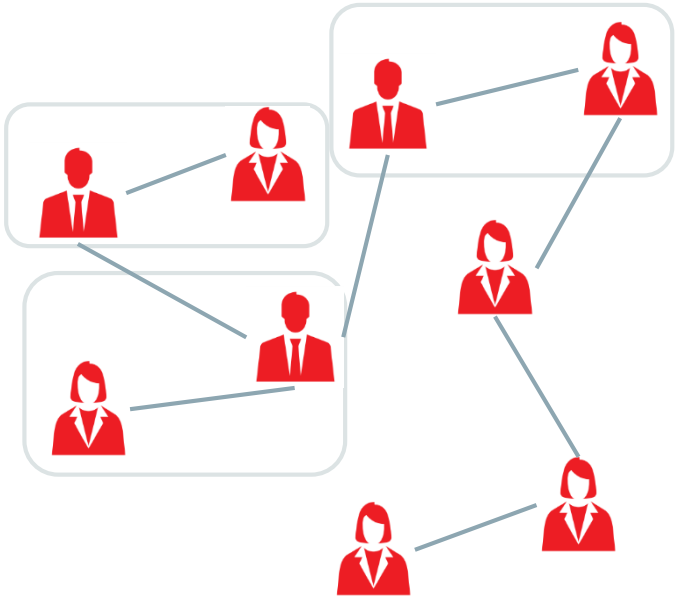
Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Graph Property Graph Analysis for Business Insight

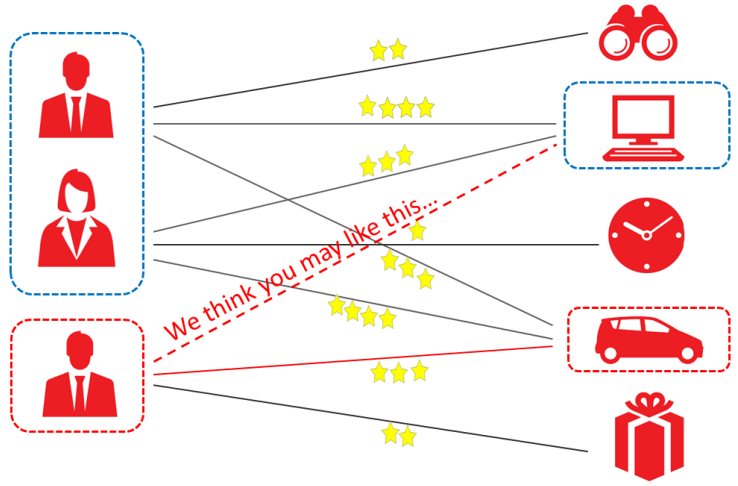
Identify Influencers



Discover Graph Patterns in Big Data

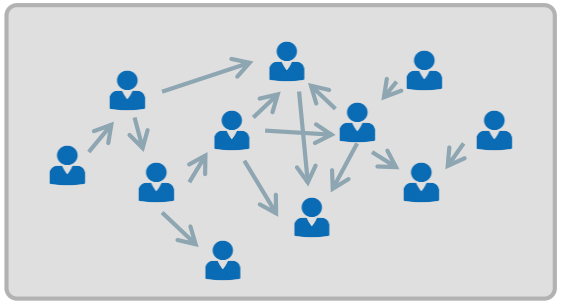


Generate Recommendations



Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Targeted Marketing in Telco

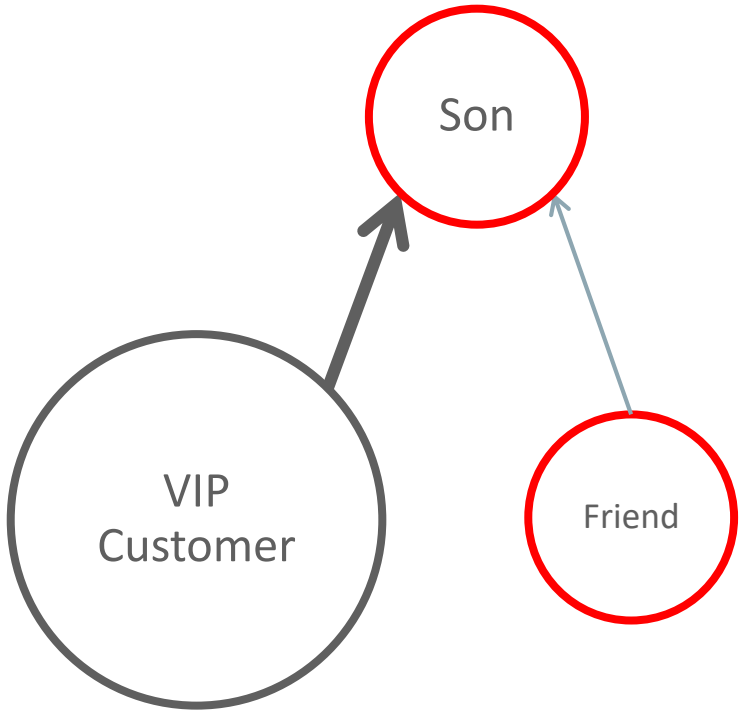


- Model each subscriber as a vertex in the graph
- Interactions between subscribers are represented by edges
  - Taking into account both on-net and off-net
- Based on call data records for voice, SMS, MMS
  - Usually combining all interactions in a property representing the strength of the edge
- Using centrality algorithms to determine important customers
- Target these customers with marketing campaigns for retention
  - Reducing churn risk for all additional customers he/she is connected with

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Graph Analysis: Influencer Identification

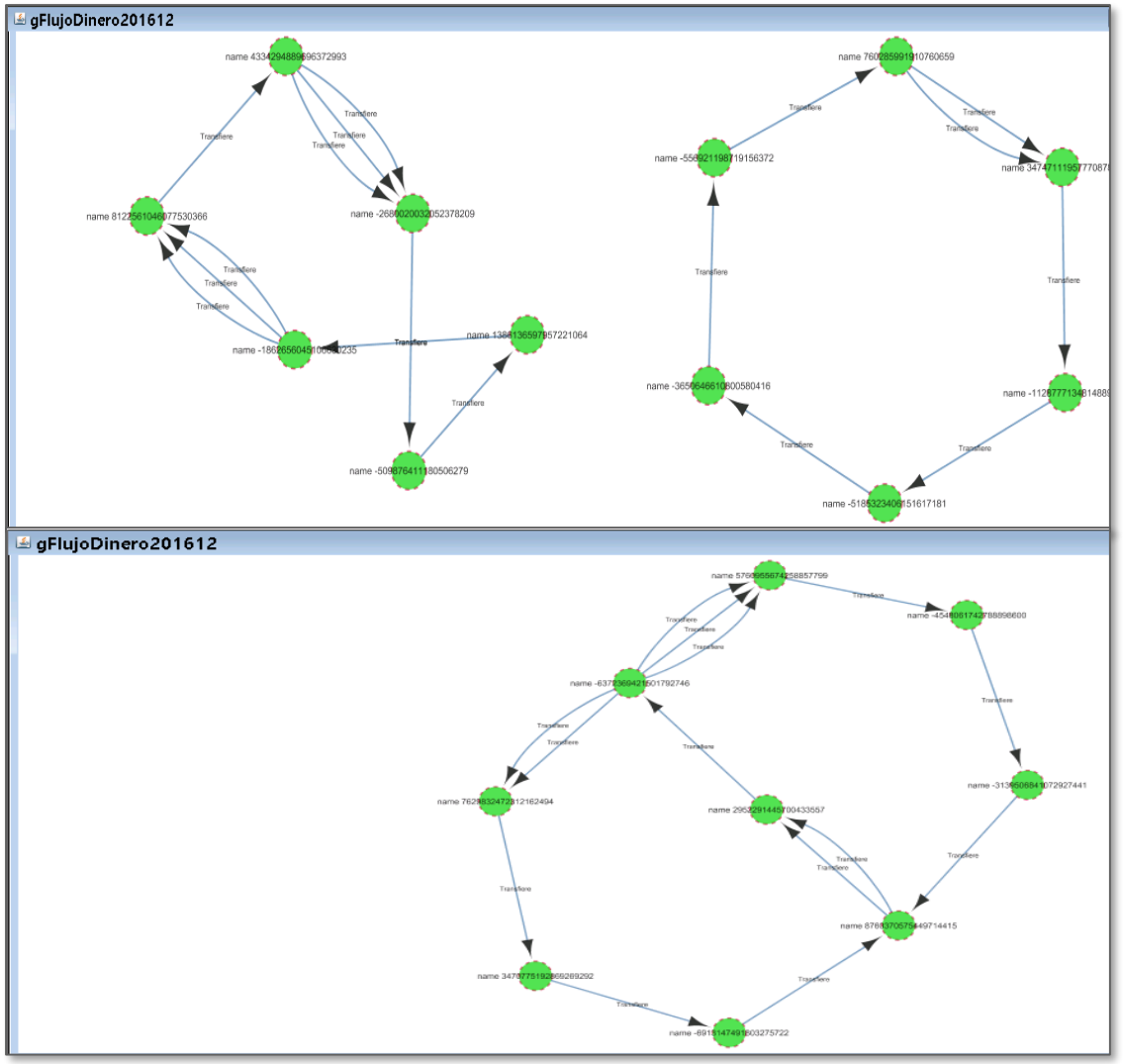
- Measuring importance using **Page Rank**
- Original algorithm developed by Larry Page for ranking in Google
- Making a node connected to by important nodes **also** important
- Can be measure of trust or prominence



Copyright © 2019 Oracle and/or its affiliates. All rights reserved. | Confidential – Oracle Internal/Restricted/Highly Restricted

Banco de Galicia

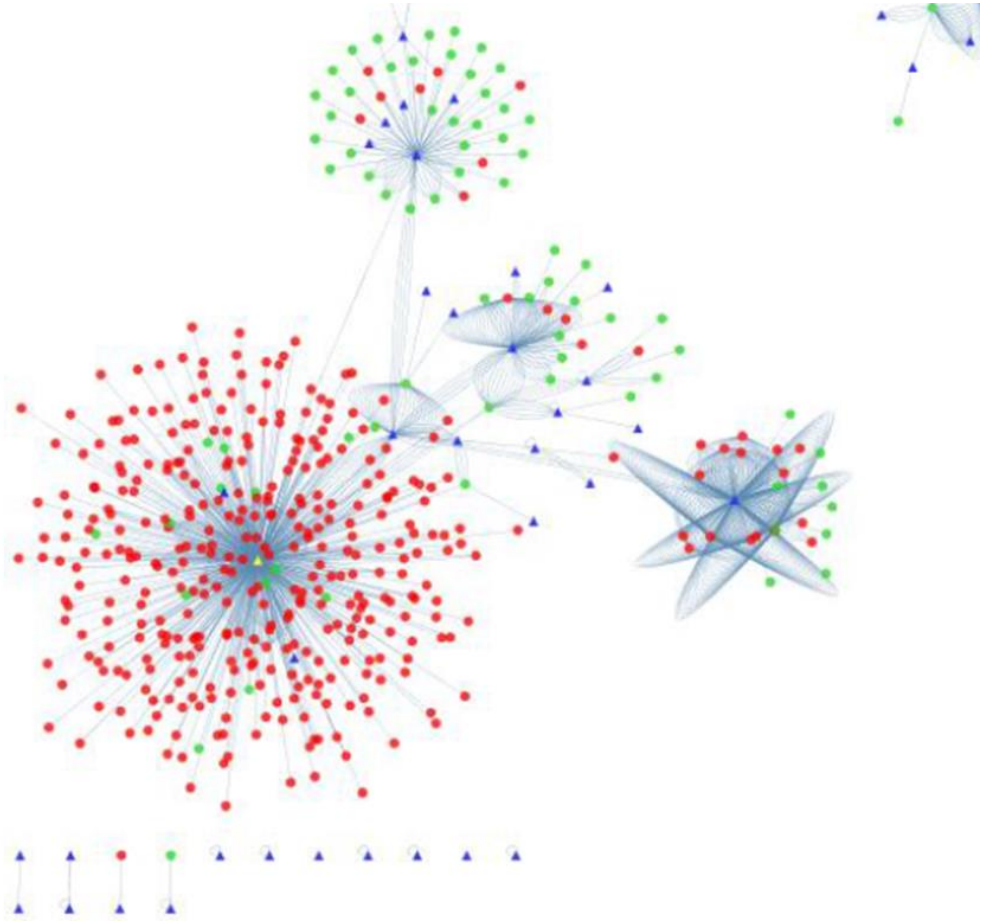
- Customer profitability analysis
  - Part of larger Hadoop/Big Data project
- Analysis of banking transactions
  - Focus on corporate customers
- Identification of undesired behavioural patterns, eg.
  - Customers using other banks to make large numbers of transactions
  - Many of which flow back to Banco Galicia
- Increase fees, terminate contracts, or move activities to Banco Galicia
- Implemented by Oracle Consulting



Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Standard Bank, South Africa

- Customer 360° analysis
  - Master data and transaction data
- Using customer ecosystem to identify potential new customers
- Evaluating money flows
  - Especially with non-banked contacts
  - Determining important entities
- Follow up with targeted marketing activities



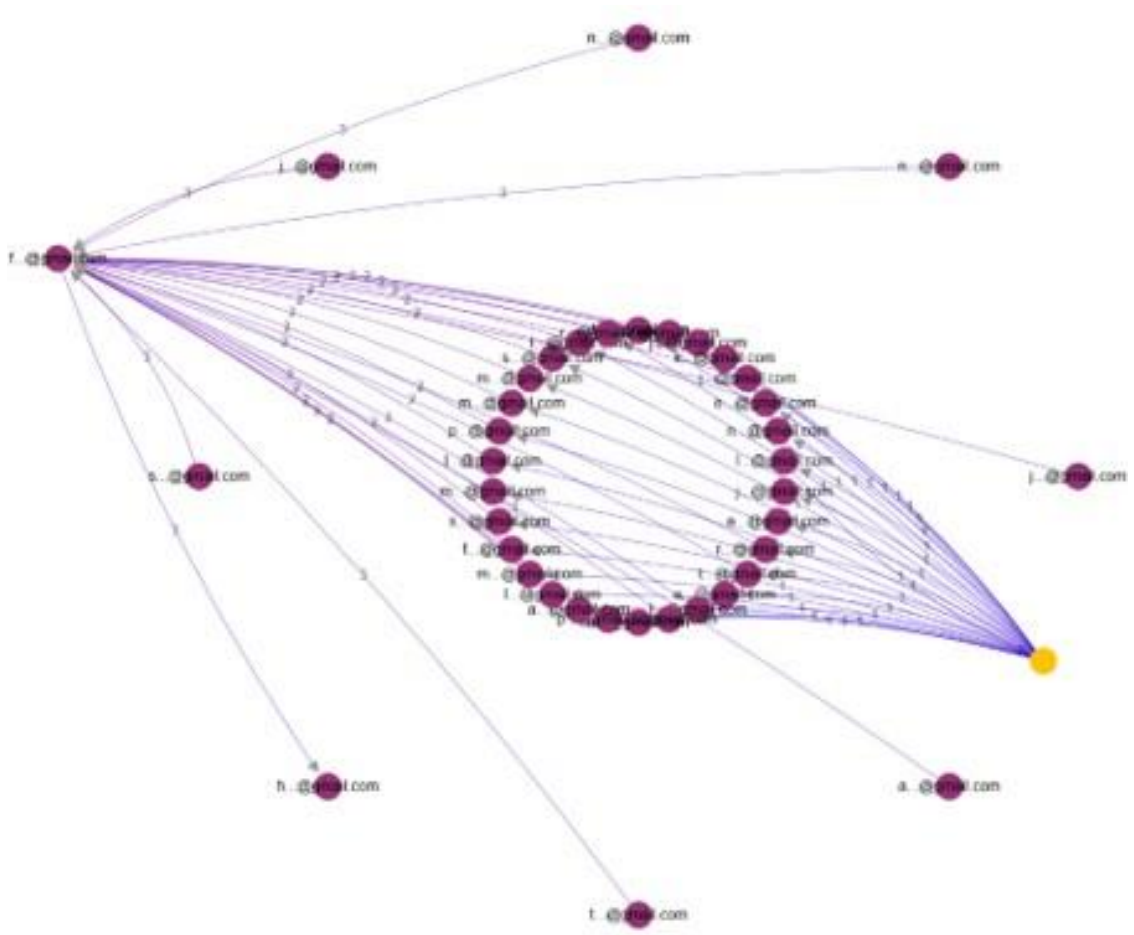
Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

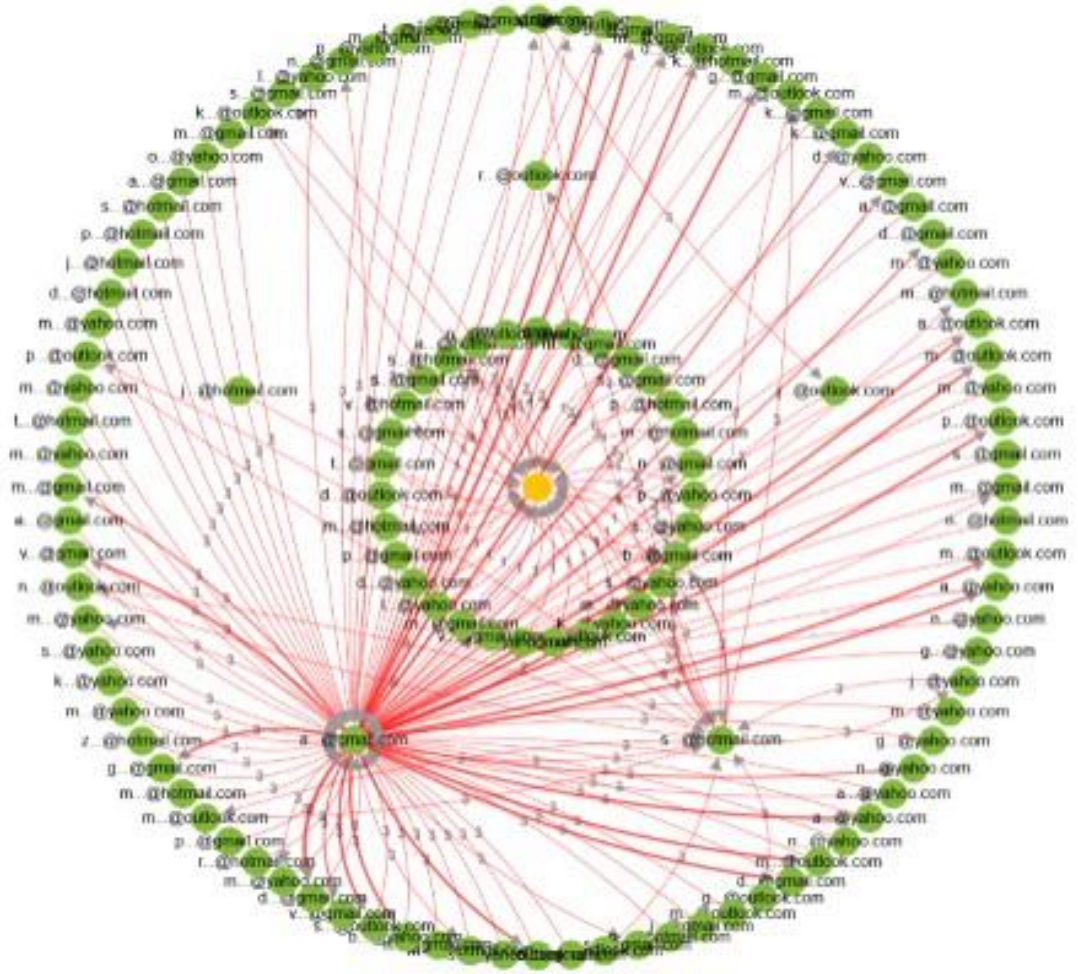
Paysafe

- Providing online payment solutions
  - Real-time payments, e-Wallets
  - 1bn revenue/yr
  - 500000 payments/day
- Strong demand for fraud detection
  - Only feasible with graph data
  - In real-time, upon money movement
  - During account creation
  - In investigation, visualizing payment flows
- Storing payments in database
  - Refreshing graph using delta update

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. | Oracle Confidential – Internal

Suspicious patterns in e-payments



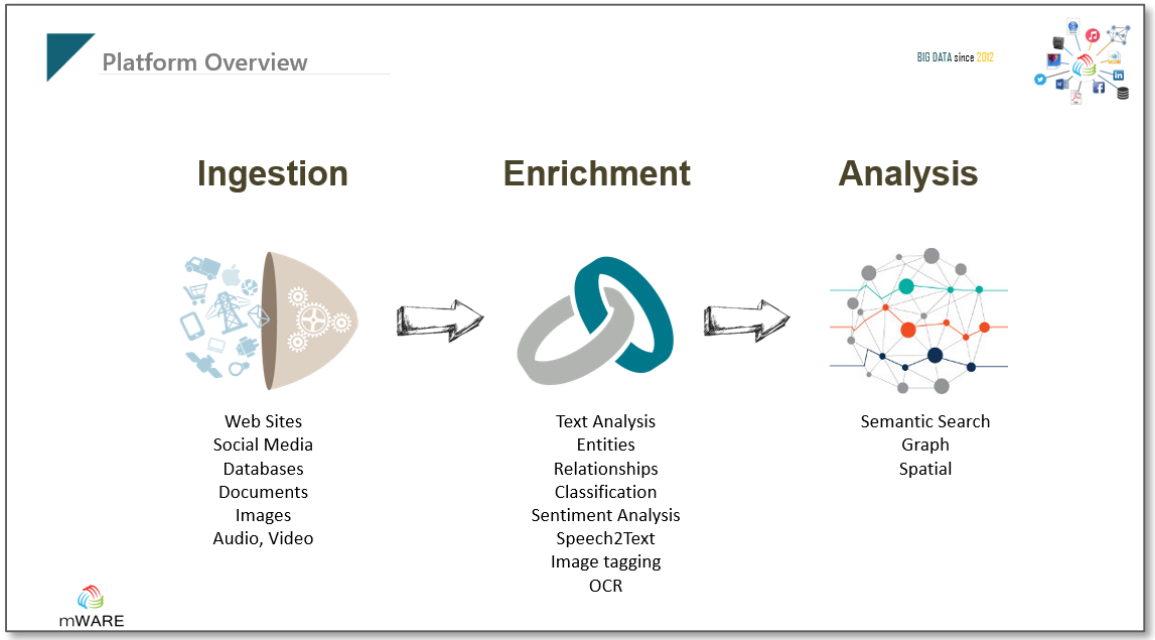


Copyright © 2019 Oracle and/or its affiliates. All rights reserved. | Oracle Confidential – Internal

Romanian Police Force

BIG DATA since 2012

- Creating Knowledge Graphs from all kinds of content
  - Social media networks, documents, images, audio, video, structured data
  - Using machine learning (text analysis, classification, entity extraction, face recognition, speech2text, ...)
- Enabling relationship analysis and semantic search
- bigCONNECT platform built by mWARE
  - Running on Big Data Applicance, Big Data Cloud Service or commodity Hadoop

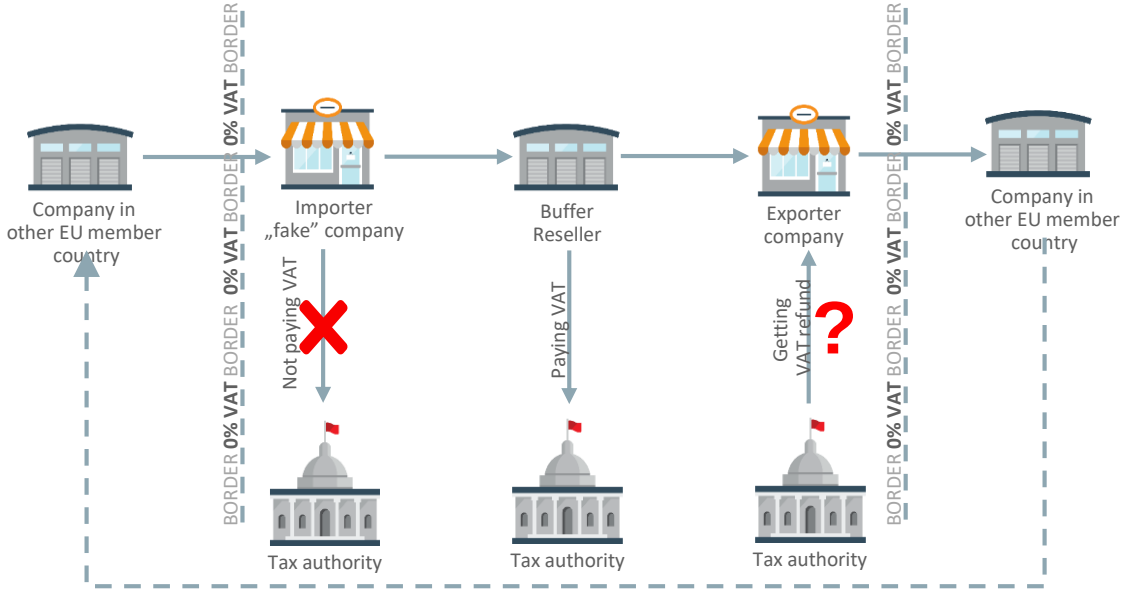


Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Ministry of Finance, Eastern Europe

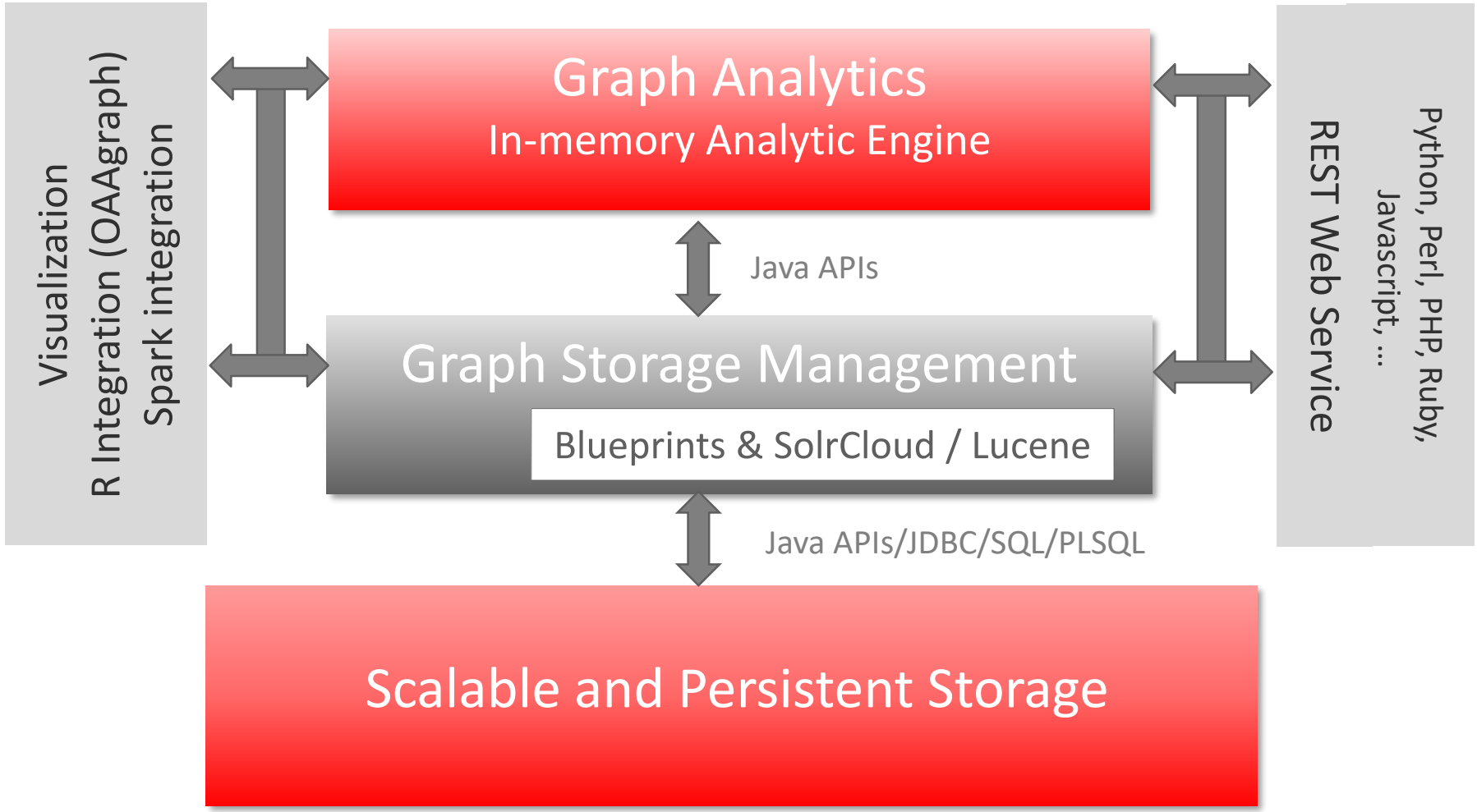
- Detecting relationships between people, accounts, companies
  - Similar to Paradise Papers
- Identifying suspicious patterns
  - Circular money transfers
  - Connections (existing path/shortest path) to companies in tax havens
- Ingesting accounting data in SAF-T format
  - Hadoop-based processing (Oozie, Spark, Hive)
  - Terabytes of data, rapidly growing
- Interactive graph analysis in Apex with Cytoscape.js

EU VAT fraud



Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

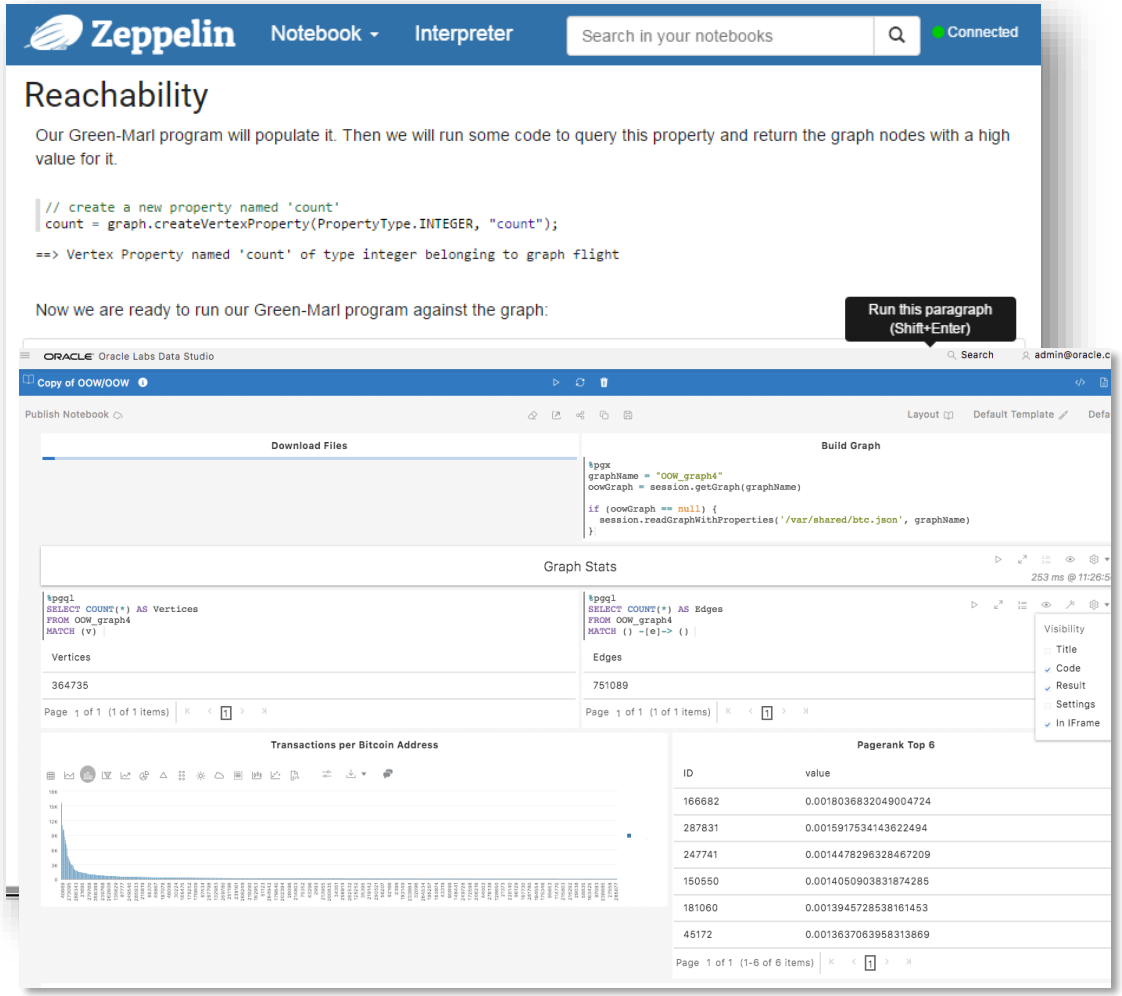
Oracle Graph Analytics Architecture



Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Notebook integration

- Multi-purpose notebook for data analysis and visualization
  - Browser-based script and query execution
- For documentation and interactive analysis
  - Typically used by Data Scientist
- Interpreters for graph analysis and graph pattern matching
  - PGX, PGQL, Markdown
- Graph visualization
- Integrated with Graph Cloud Service

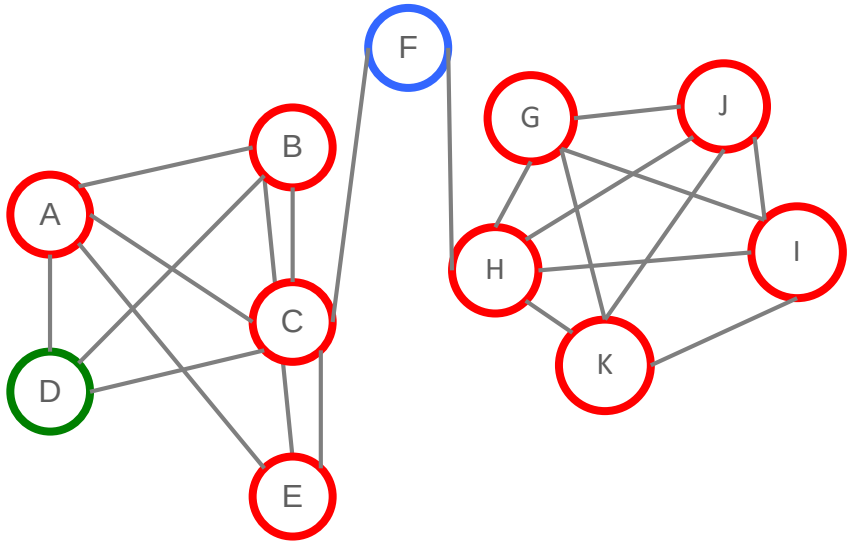


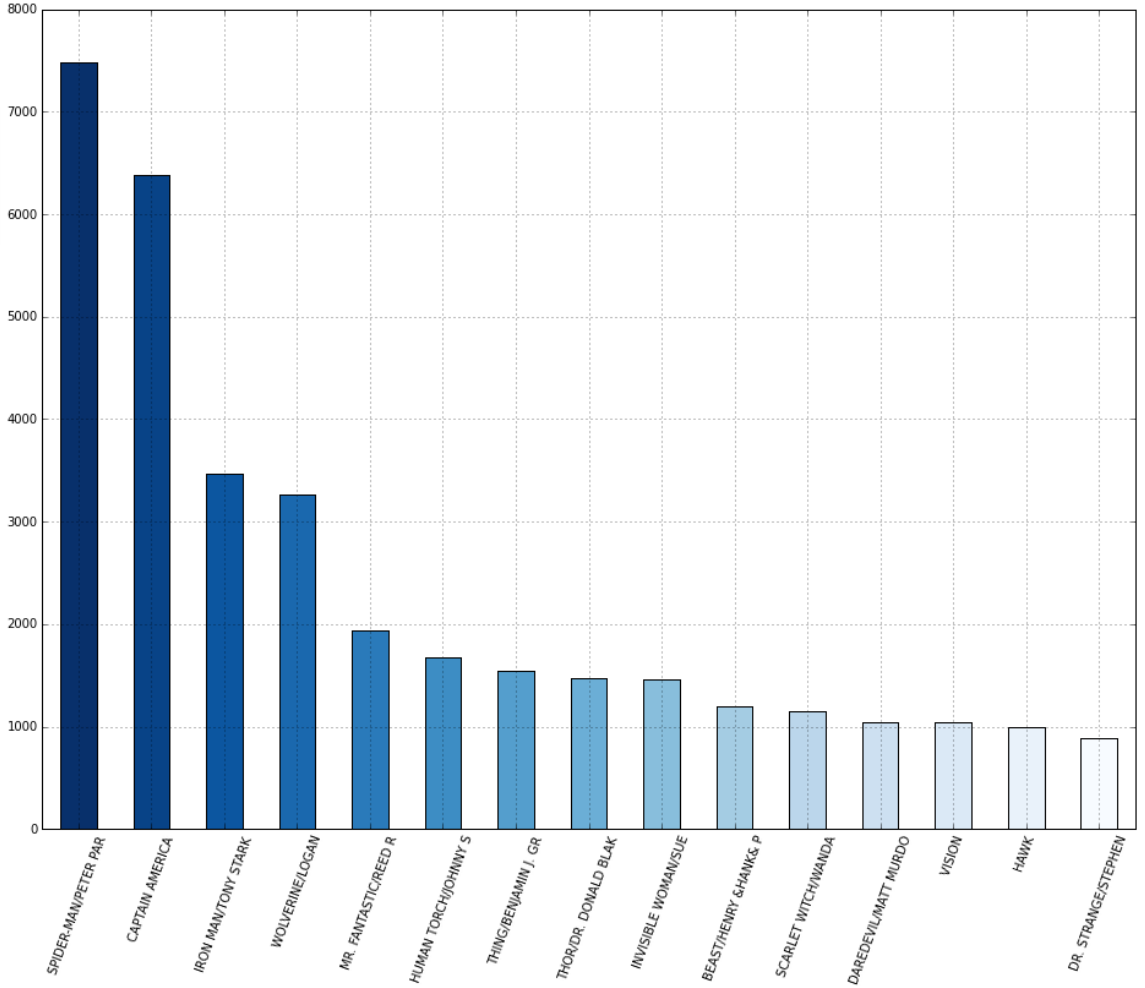
Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Example: Betweenness Centrality in Big Data Graph

Over 50 pre-built graph algorithms

```
analyst.vertexBetweennessCentrality(pg)
.getTopKValues(15)
```



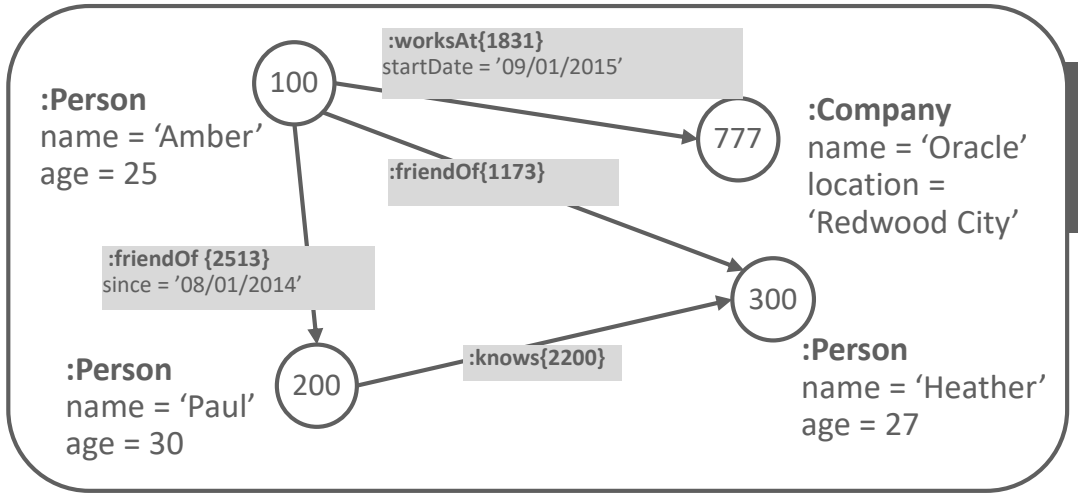


Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Basic graph pattern matching

- Find all instances of a given pattern/template in the data graph

```
SELECT v3.name, v3.age
  FROM socialNetworkGraph
 MATCH (v1:Person) -[:friendOf]-> (v2:Person) -[:knows]-> (v3:Person)
 WHERE v1.name = 'Amber'
```



socialNetwork Graph

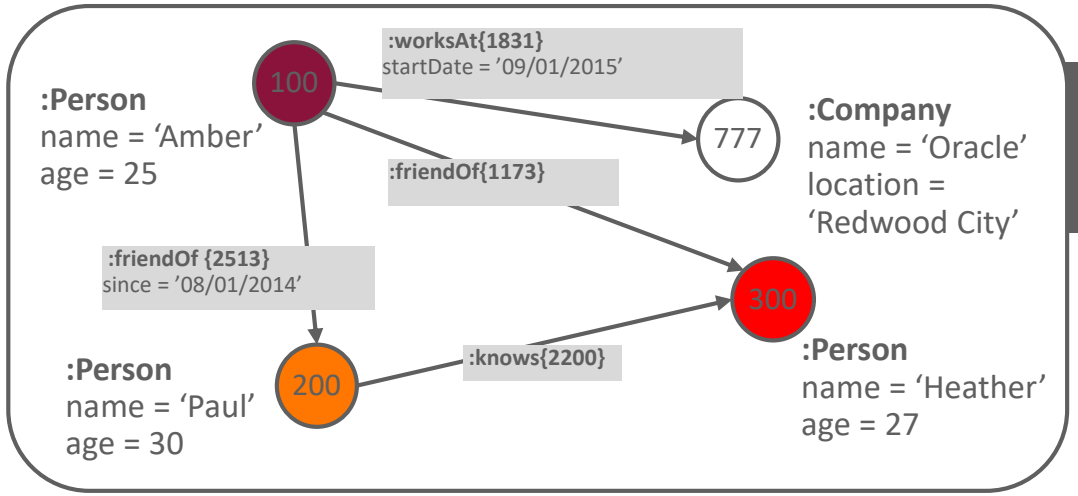
Query: Find all people who are known by friends of 'Amber'.

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Basic graph pattern matching

- Find all instances of a given pattern/template in the data graph

```
SELECT v3.name, v3.age
  FROM socialNetworkGraph
 MATCH (v1:Person) -[:friendOf]-> (v2:Person) -[:knows]-> (v3:Person)
 WHERE v1.name = 'Amber'
```



socialNetwork Graph

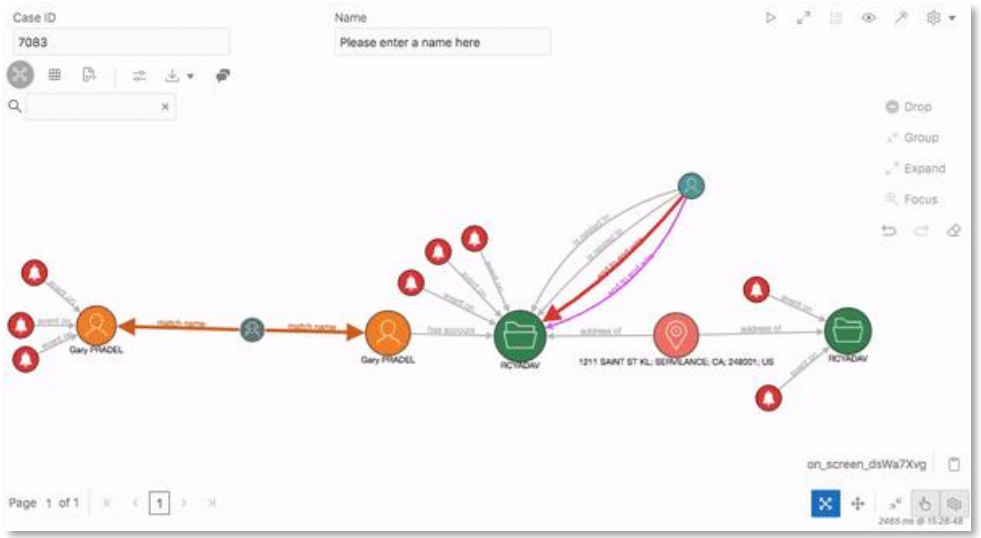
Query: Find all people who are known by friends of 'Amber'.

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Property Graph Visualization

New visualization component (PGViz), planned for CY20

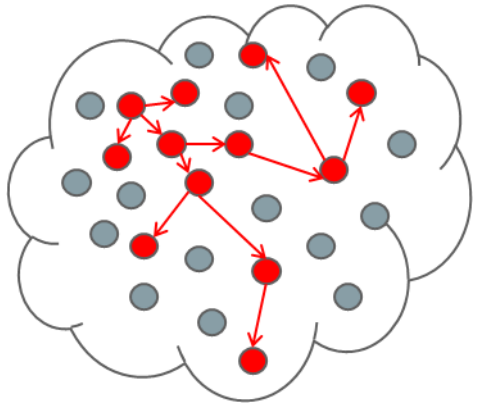
- Single-Page Web Application based on Oracle JET and D3.js
- Takes PGQL Query as input, renders result set visually
- Uses the same JavaScript visualization component Graph Studio (Graph Cloud service) is using
- Will support PGQL-to-PGX initially, but can work with anything that supports PGQL (including PGQL-to-SQL)



Copyright © 2019 Oracle and/or its affiliates. All rights reserved. | Oracle Confidential – Internal

Combining Graph Analytics and Machine Learning

Graph Analytics

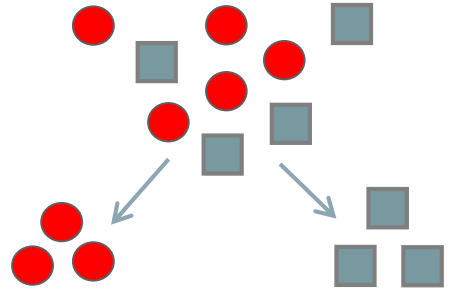


- Compute graph metric(s)
- Explore graph or compute new metrics using ML result

Add to structured data

Add to graph

Machine Learning



- Build predictive model using graph metric
- Build model(s) and score or classify data

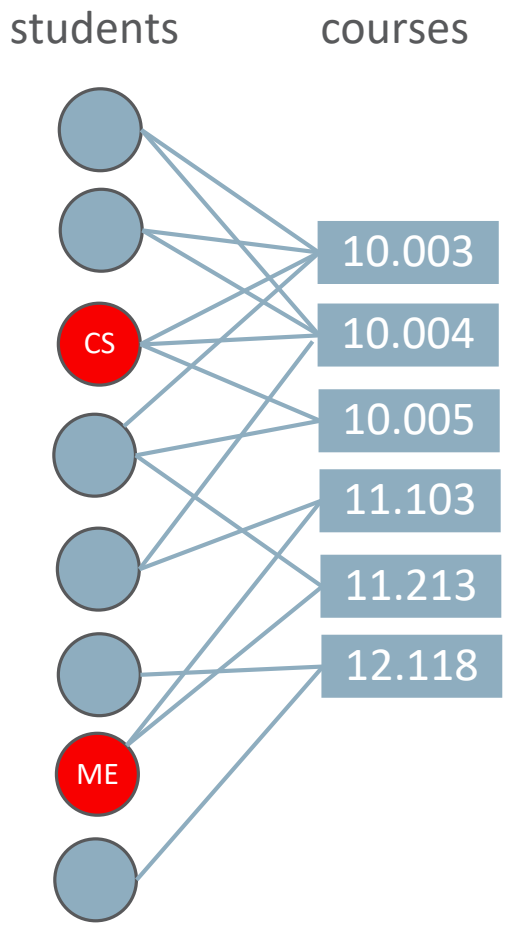
ORACLE

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Practical example – Student classification

- Can you predict a student's major or department just by looking at the classmates in the course that (s)he is taking?

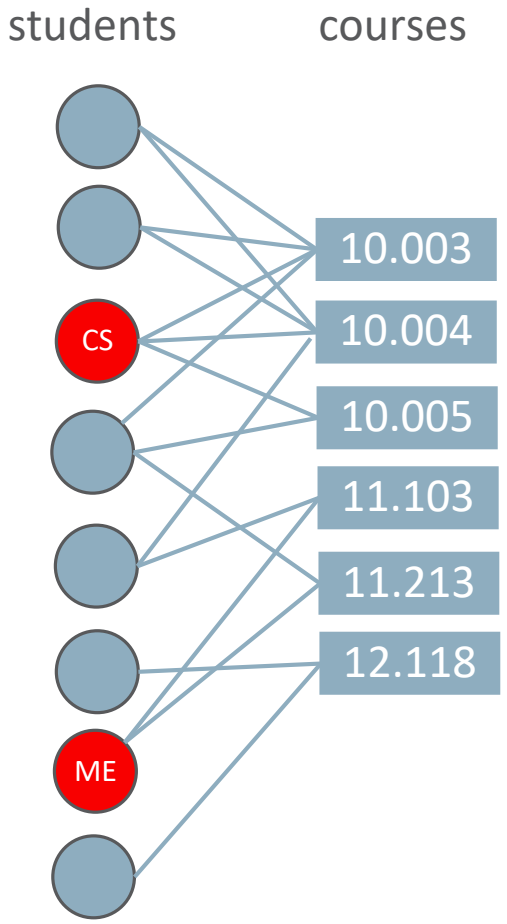
- Very similar to customer segmentation problem
  - Student => Customer
  - Course taking => Item or service purchase
  - Department => Segment label



Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Evaluation – Comparison

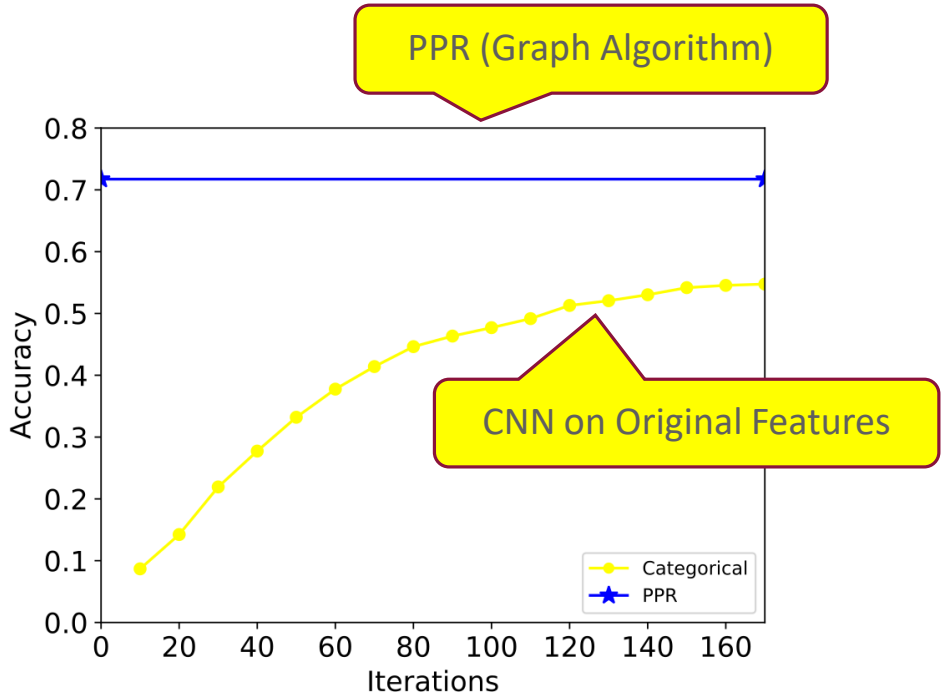
1. CNN trained on “standard” features (e.g., student age, courses taken, …)
2. Use PPR and predict the department of the highest-scoring vertex
3. Train a CNN on vertex embeddings extracted with DeepWalk
4. Add “standard” features beside graph embeddings



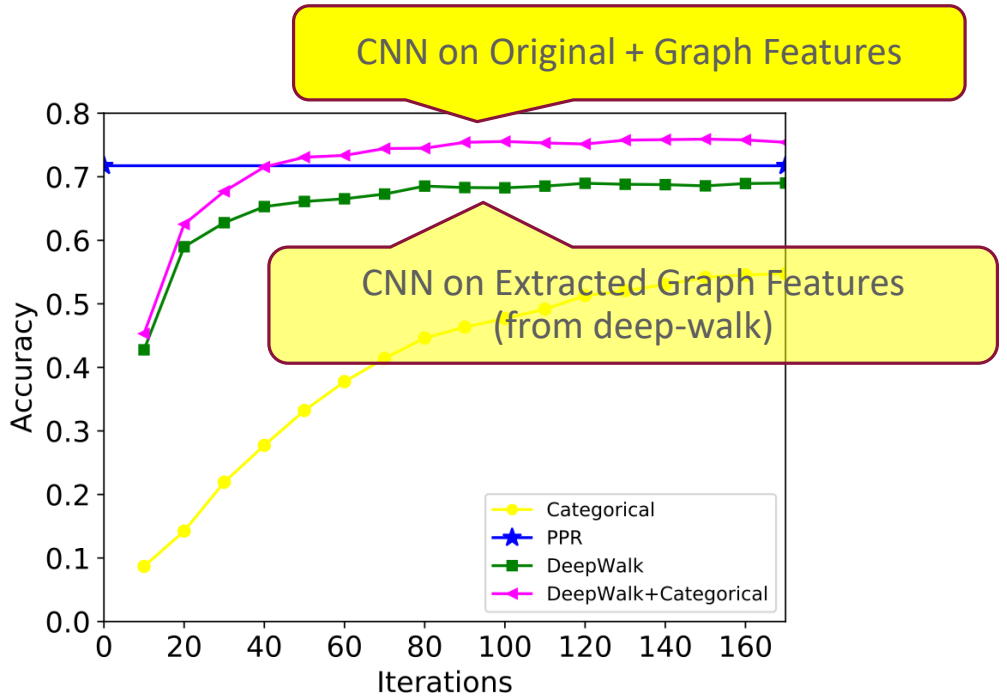
Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Results

- (Result #1) Graph-based prediction gives better result than naïve application of ML (e.g. CNN) on basic student features (e.g. age, gender, background, …)



- (Result #2) Deep-Walk preserves information from graph representation
- (Result #3) Deep-Walk allows to combined graph data with other features



Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Data Processing Workflow

- Convert raw dataset to graph
- Load graph into in-memory analytics engine
- Execute graph algorithm
  - Analyze result
- Export embedding
- Use for Machine Learning
  - eg. using Oracle Advanced Analytics

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Summary

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Summary

Graph capabilities in Oracle Big Data Spatial and Graph

- Graph databases are powerful tools, complementing relational databases
  - Especially strong for analysis of graph topology and connectedness
- Graph analytics offer new insight
  - Especially relationships, dependencies and behavioural patterns
- Oracle Property Graph technology offers
  - Comprehensive analytics through various APIs, integration with relational database
  - Scaleable, parallel in-memory processing
  - Secure and scaleable graph storage using Hadoop platform or Oracle Database
- Available both on-premise or in the Cloud already today

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Q&A

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

More information

- Oracle Big Data Spatial and Graph OTN product page: www.oracle.com/technetwork/database/database-technologies/bigdata-spatialandgraph
  - White papers, software downloads, documentation and videos
- Oracle Big Data Lite Virtual Machine - a free sandbox to get started: www.oracle.com/technetwork/database/bigdata-appliance/oracle-bigdatalite-2104726.html
- Hands On Lab included in /opt/oracle/oracle-spatial-graph/
  - Content also available on GITHub under http://github.com/oracle/BigDataLite/
- Blog – examples, tips & tricks: blogs.oracle.com/bigdataspatialgraph
- @OracleBigData, @SpatialHannes, @JeanIhm Oracle Spatial and Graph Group

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Introduction to Graph analytics

Youtube videos

- What is Oracle Big Data Spatial and Graph?
  https://youtu.be/t9pJJhzZKOE

  How can graph analytics help my business?
  https://youtu.be/0dJNzBi7B-k

  Detecting anomalies with Oracle Big Data Spatial and Graph
  https://youtu.be/nfP6HDOImjY

  Generating recommendations with Oracle Big Data Spatial and Graph
  https://youtu.be/9LRlF3of-Hs

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

Integrated Cloud

Applications & Platform Services

Copyright © 2019 Oracle and/or its affiliates. All rights reserved. |

ORACLE®