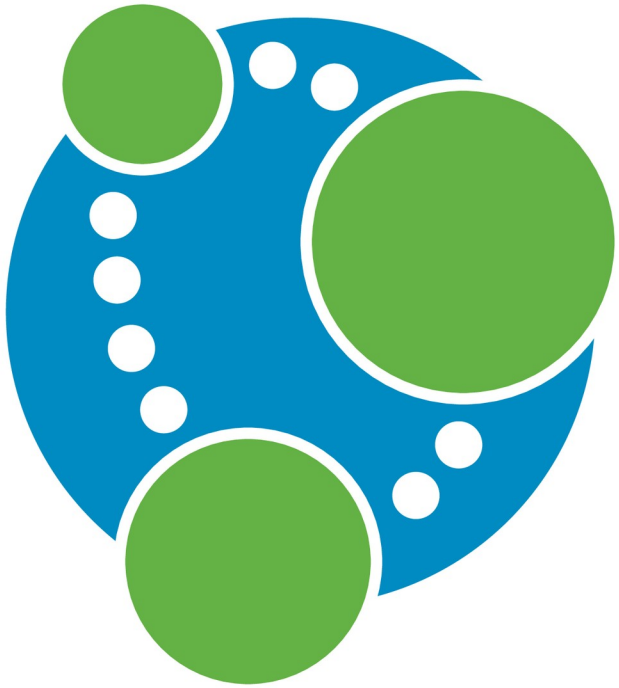
Introduction to



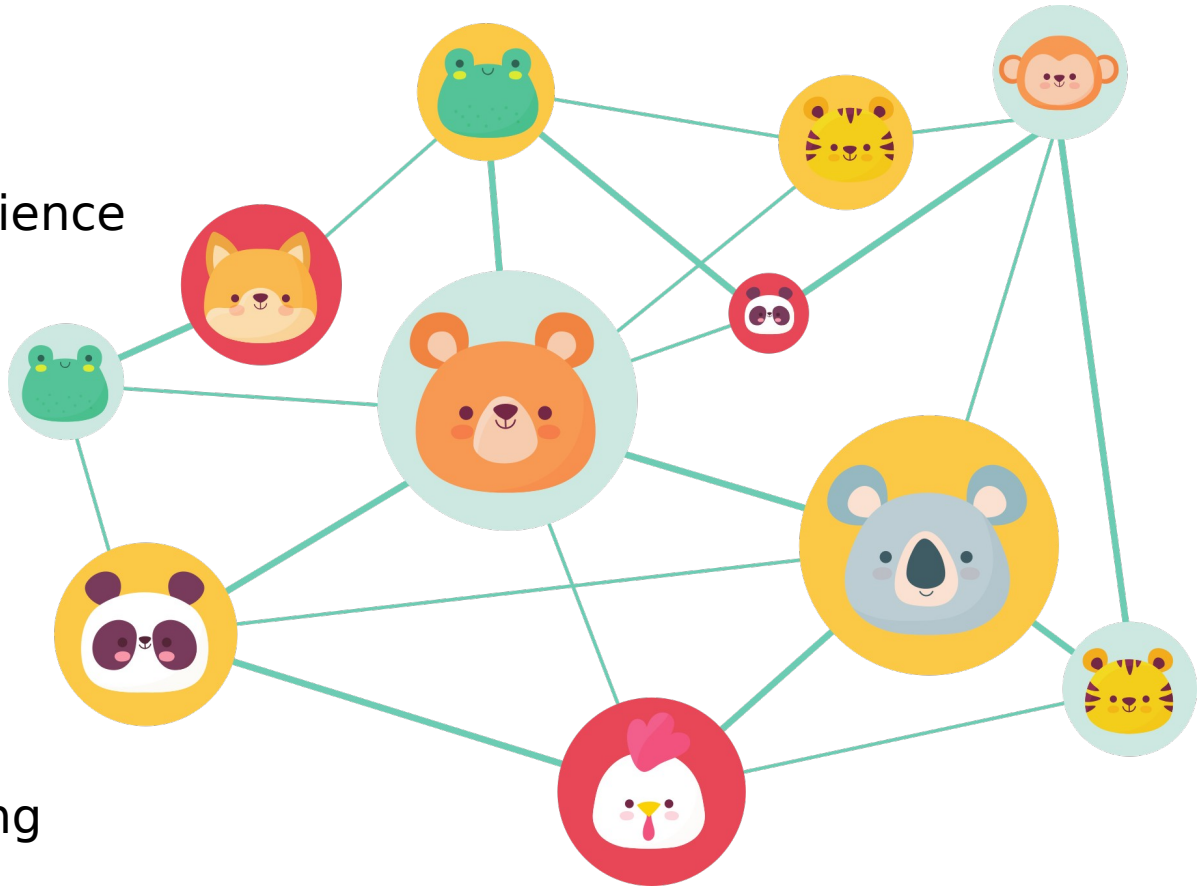
Social Network Data Analysis

Using Neo4j

Amirabbas Asadi

Workshop Topics

- Social Network Data Analysis
- Graph Theory Basic Definitions
- Neo4j, Cypher and Graph Data Science
- Diameter and Shortest Path
- Graph Centrality
- Community Detection
- Link Prediction
- Review on Graph Machine Learning



Graph created by freepik - www.freepik.com

Connected Data

Social Network Analysis

Why Social Network Analysis is Challenging?

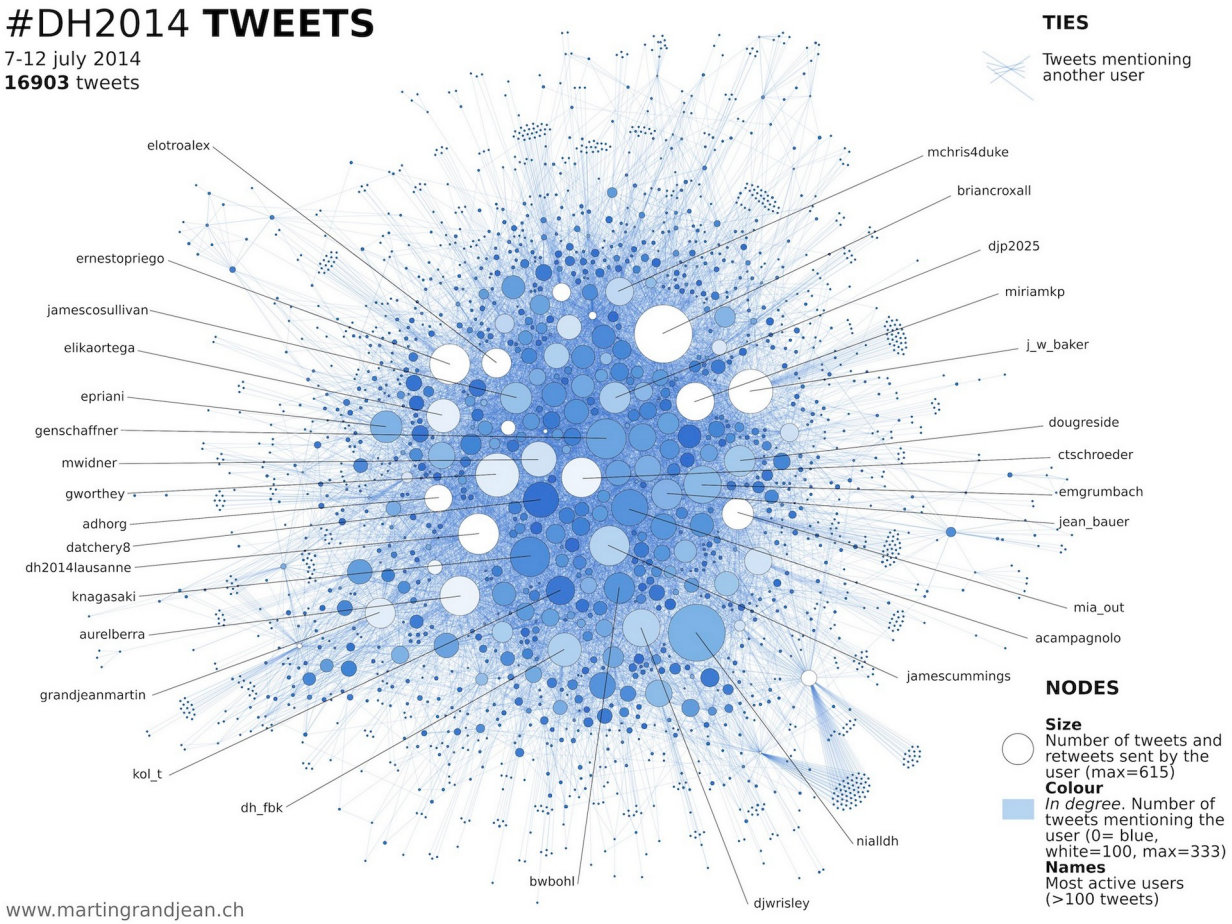
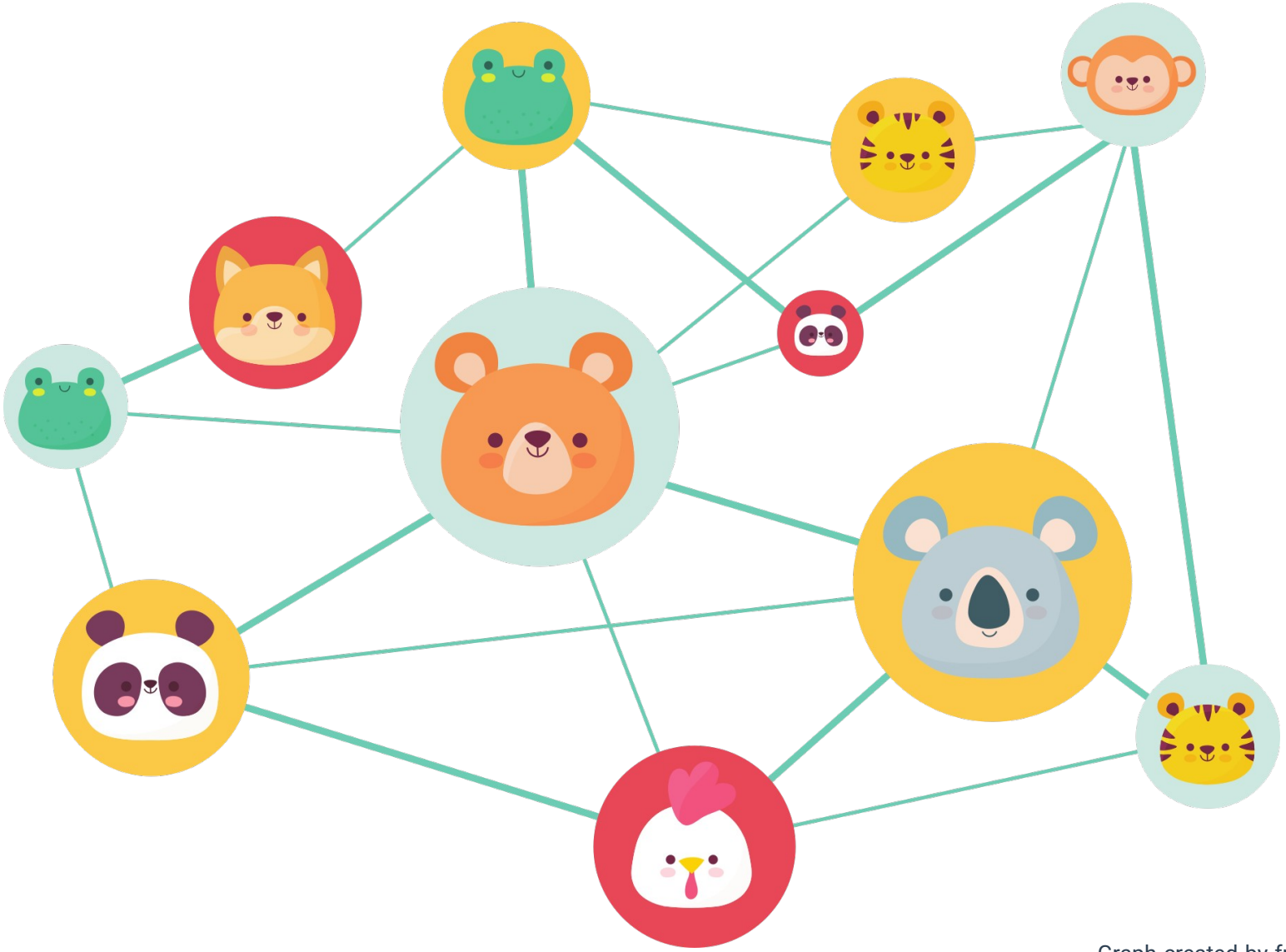


Image Source: martingrandjean.ch/

Graph Theory Basic Definitions

Consider This Graph



Graph created by freepik - www.freepik.com

Graph Databases

Why not to use Relational DBs for Connected data?

I hate the JOIN query !!!!

Students

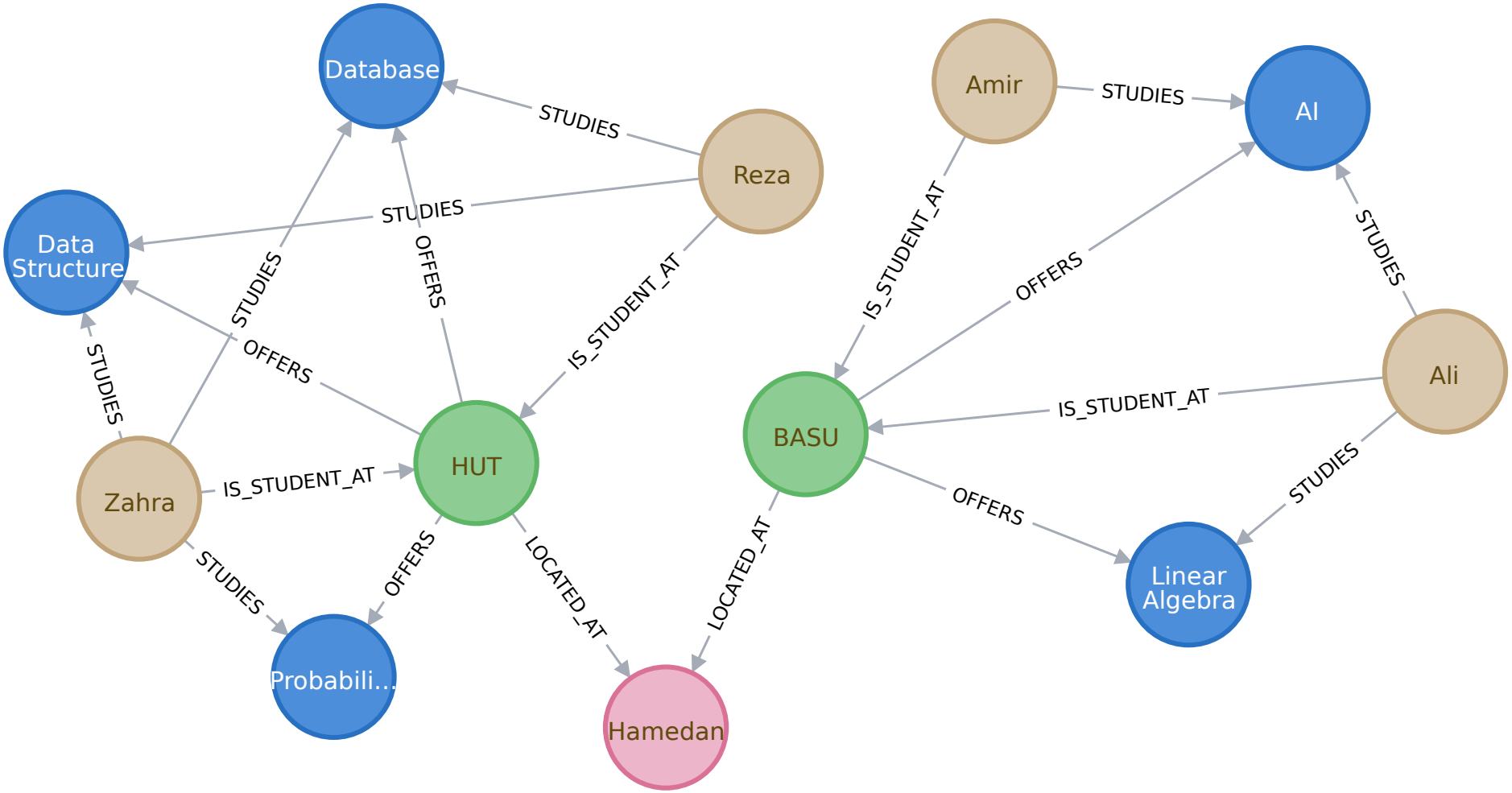
Courses

Universities

Cities

Student_Course

Graph Databases Suggest a more intuitive model


Database
Reza
Data Structure
HUT
Zahra
Probabili...
Hamedan
Amir
AI
BASU
Ali
Linear Algebra
STUDIES
OFFERS
IS_STUDENT_AT
LOCATED_AT

Graph Databases and Tools

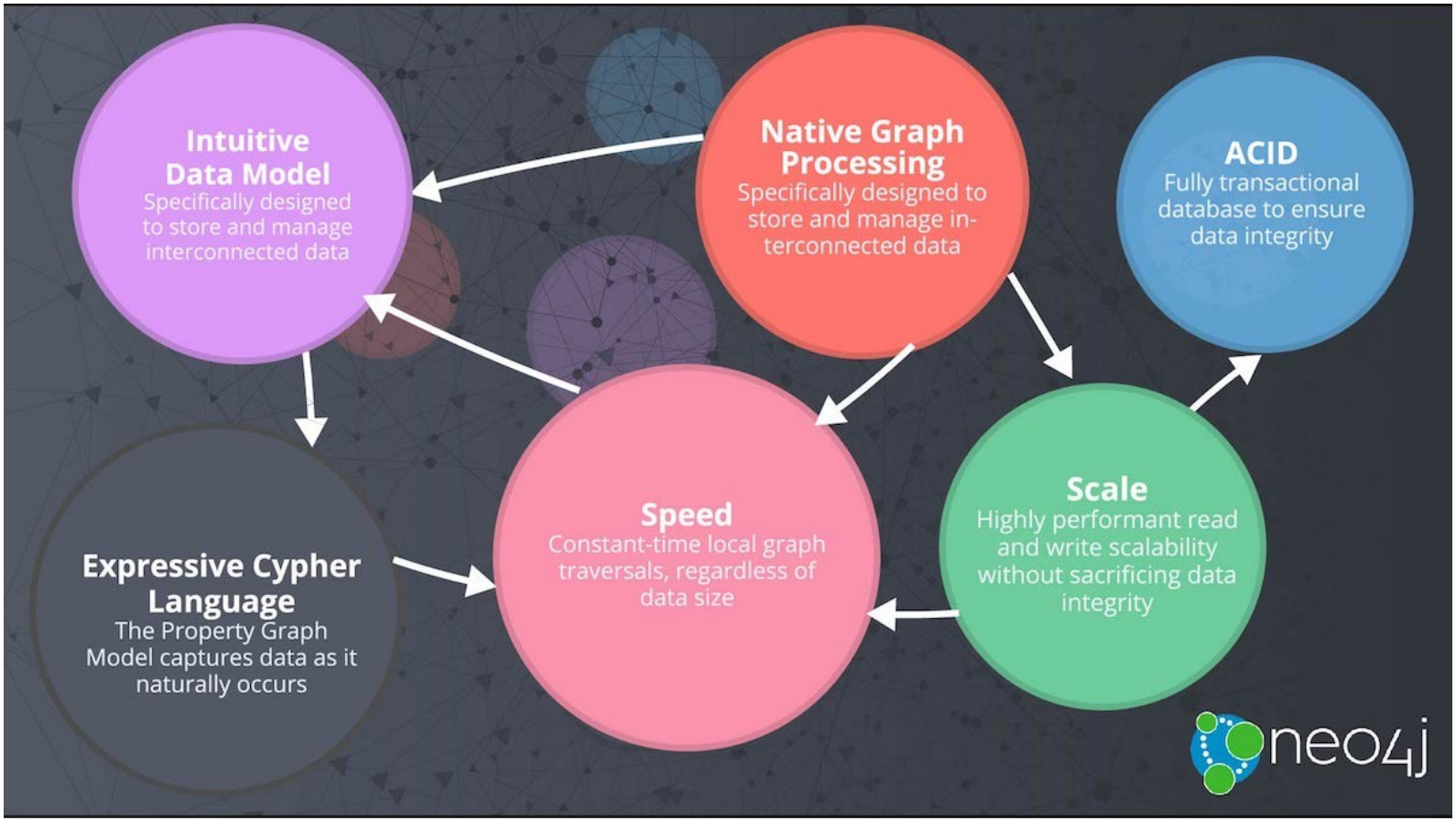


Image Source: Neo4j Docs

What do we need?

- Neo4j Desktop
- Graph Data Science Package

https://neo4j.com/

Graph Data Model In Neo4j

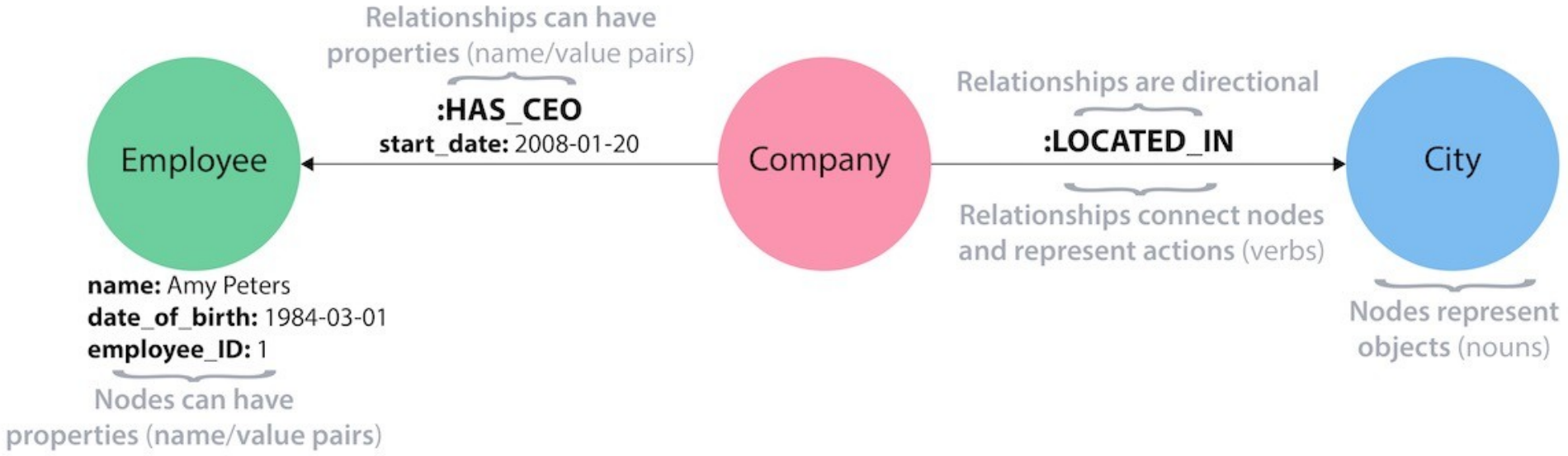


Image Source: Neo4j Docs

Say Hello To Cypher

Cypher Basics

CREATE, MATCH, SET, DELETE

Import Data From File

Example Data: https://snap.stanford.edu/data/wiki-Vote.html

Somehow

We know each other :)

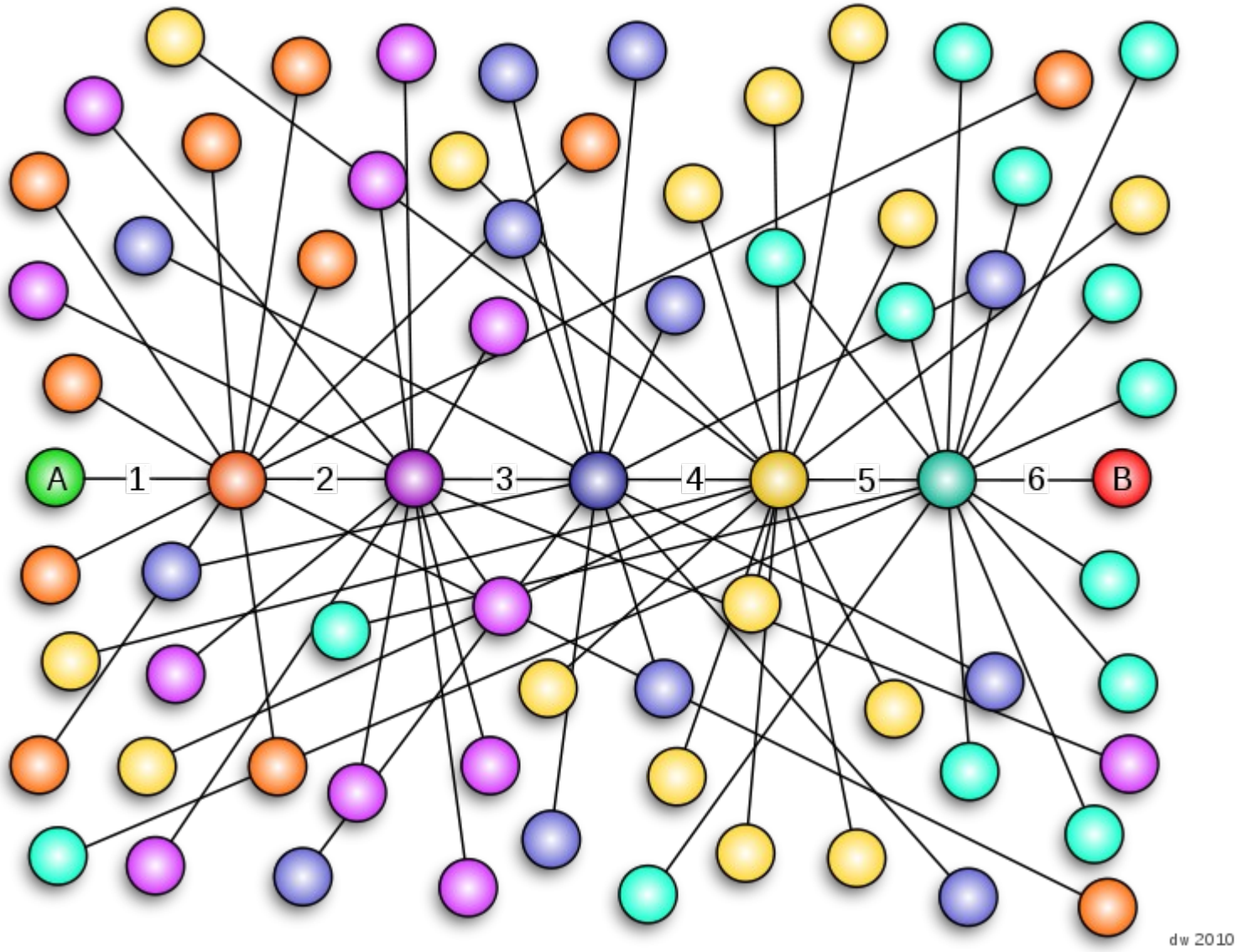


Image Source: Wikipedia

How to find Diameter and Shortest Path??

Finding Shortest Path using Neo4j

Barack Obama

Katy Perry

Justin Bieber

Rihanna

Taylor Swift

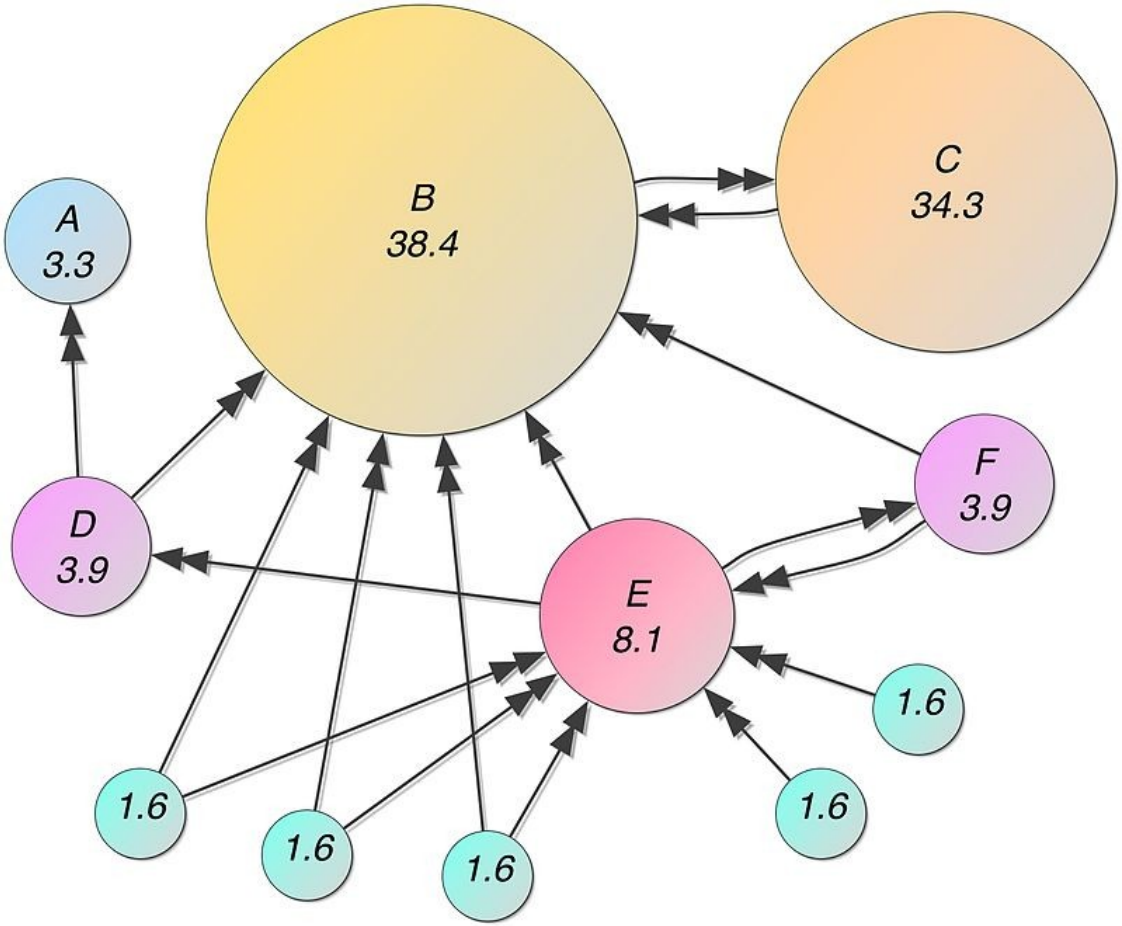
Source: https://www.tweetbinder.com/blog/top-twitter-accounts/

Centrality Detection Algorithms

Degree Centrality

Eigenvector Centrality

PageRank Algorithm



$$PR(u) = \sum_{v \in B_u} \frac{PR(v)}{L(v)}$$

Image Source: Wikipedia

Community Detection, Clustering

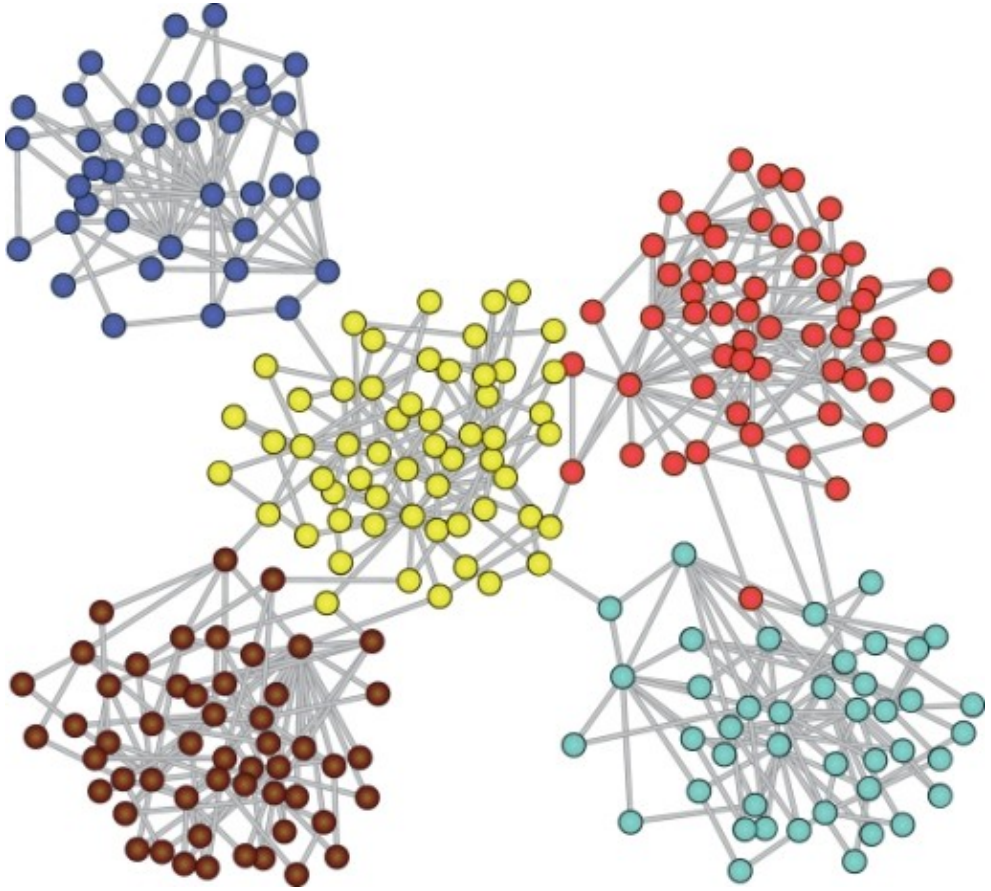


Image Source: github.com/PrestonEn/alien-pineapple/

What’s wrong with this?

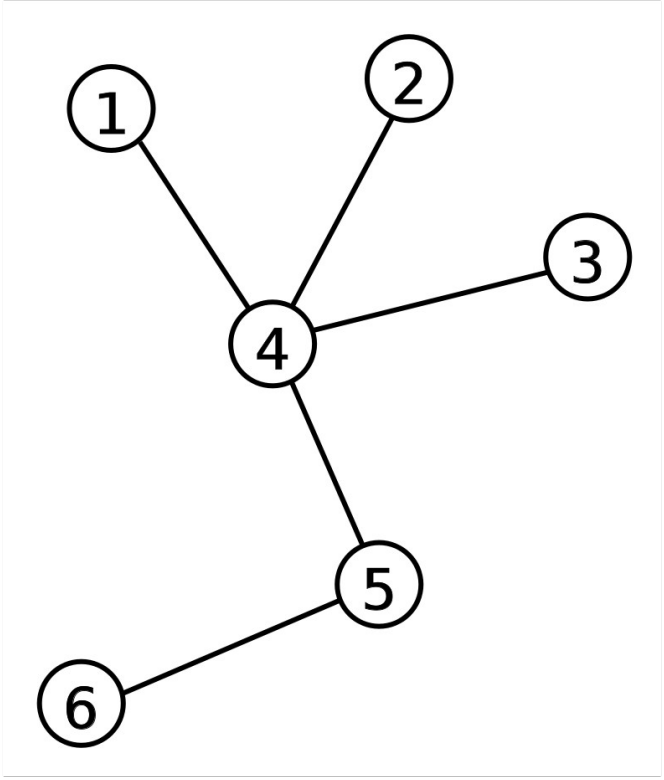


Image Source: Wikipedia

Triangles & Clustering Coefficients

How to find a community?

Label Propagation Algorithm

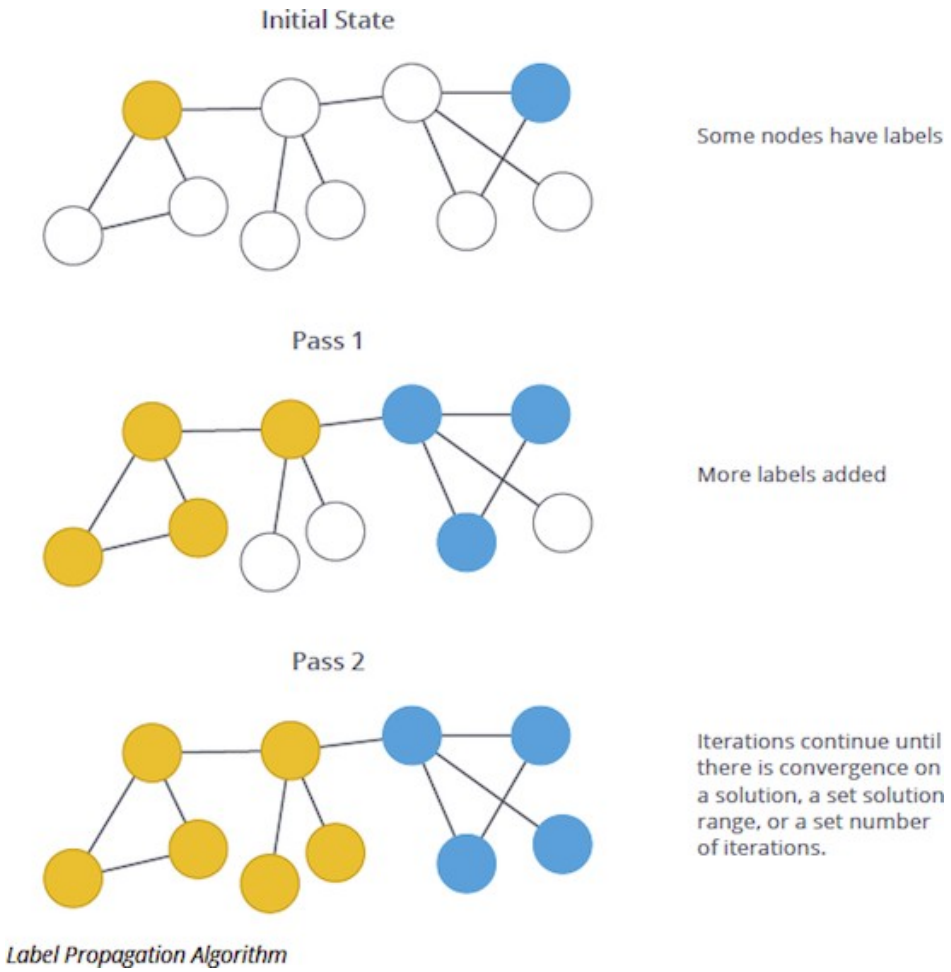


Image Source: Neo4j Blog Post by Mark Needham & Amy E. Hodler

The people you may know ...

Link Prediction

Adamic Adar

Preferential Attachment

Common Neighbors

...

We can do better ...

Machine Learning on Graphs

Why Learning on Graphs is hard?

Node Embedding

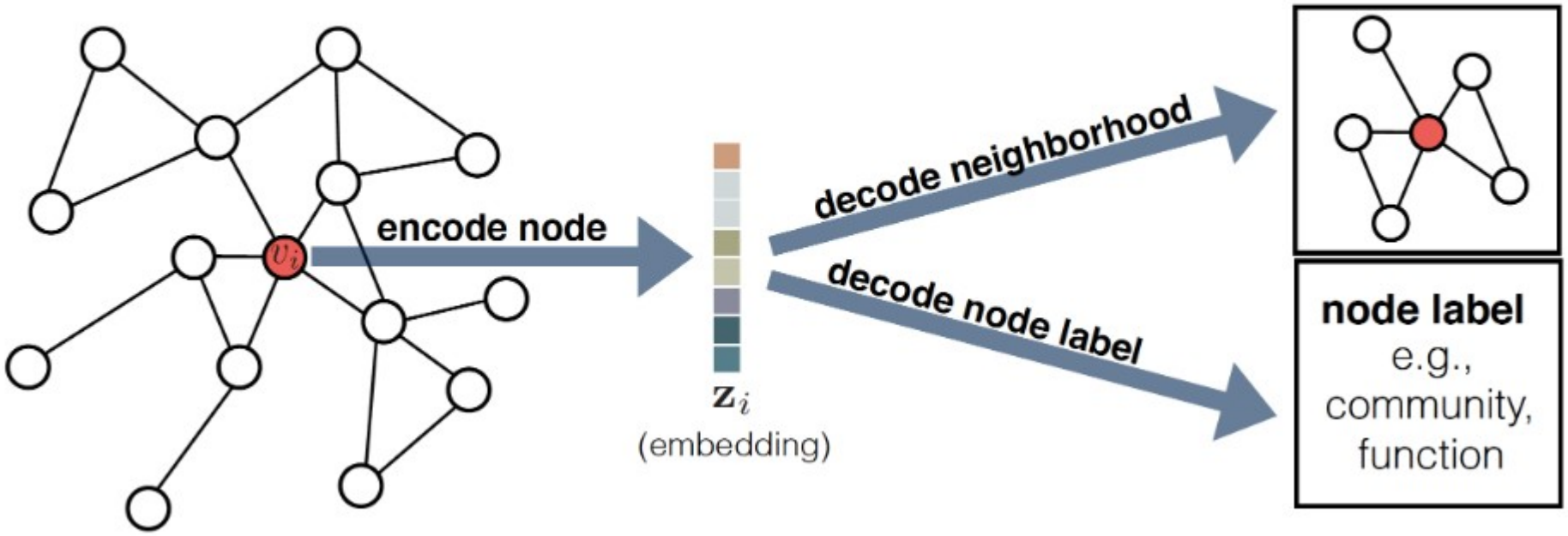
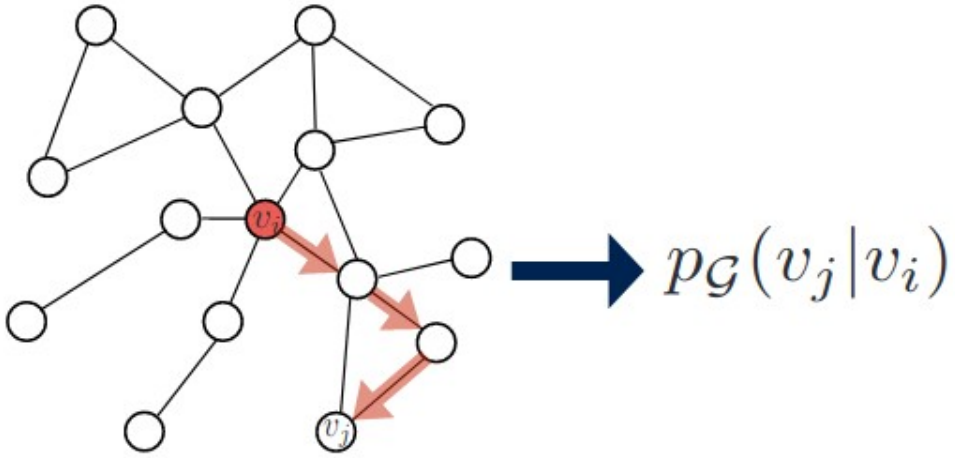
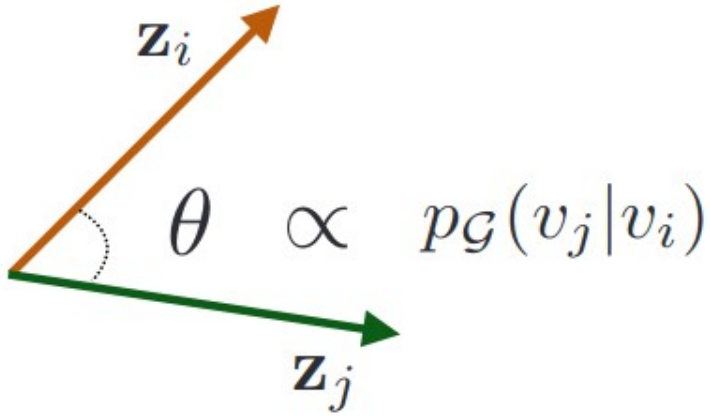


Image from: Representation Learning on Graphs: Methods and Applications, arXiv:1709.05584 [cs.SI]

Random Walks & Co-occurrence Statistics



1. Run random walks to obtain co-occurrence statistics.



2. Optimize embeddings based on co-occurrence statistics.

Image from: Representation Learning on Graphs: Methods and Applications, arXiv:1709.05584 [cs.SI]

Using Deep Autoencoder

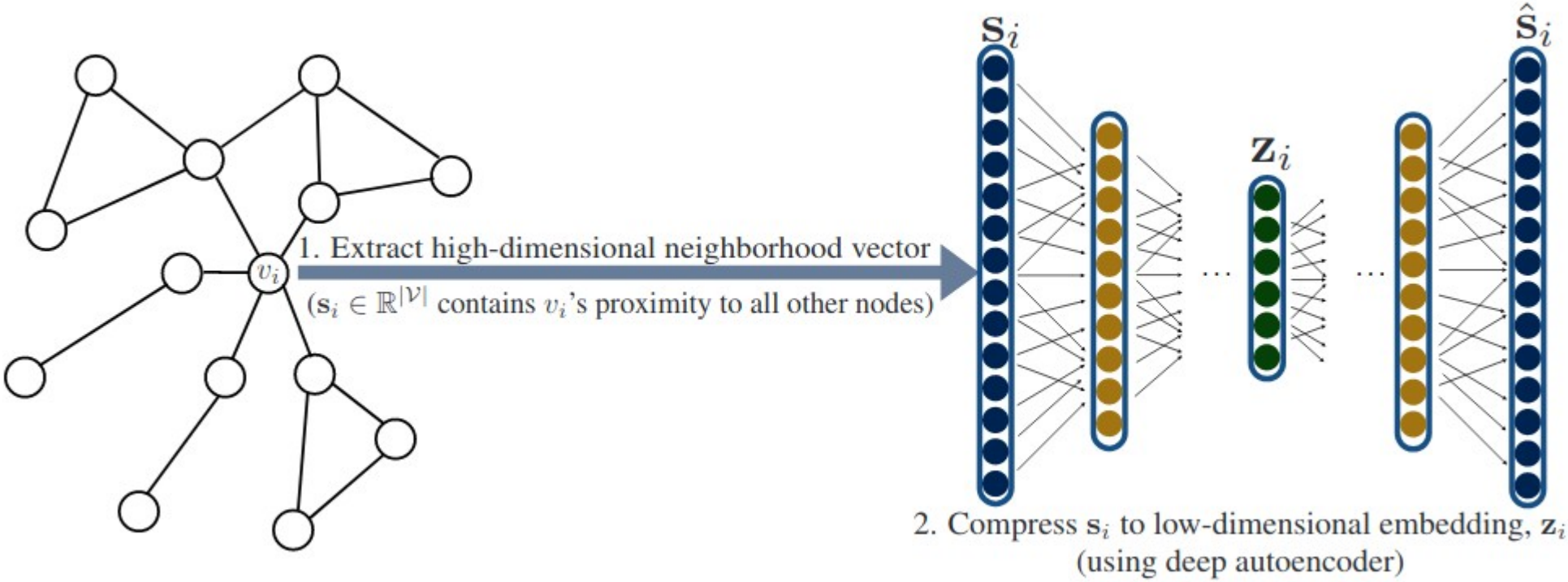


Image from: Representation Learning on Graphs: Methods and Applications, arXiv:1709.05584 [cs.SI]

Graph Convolutional Neural Network

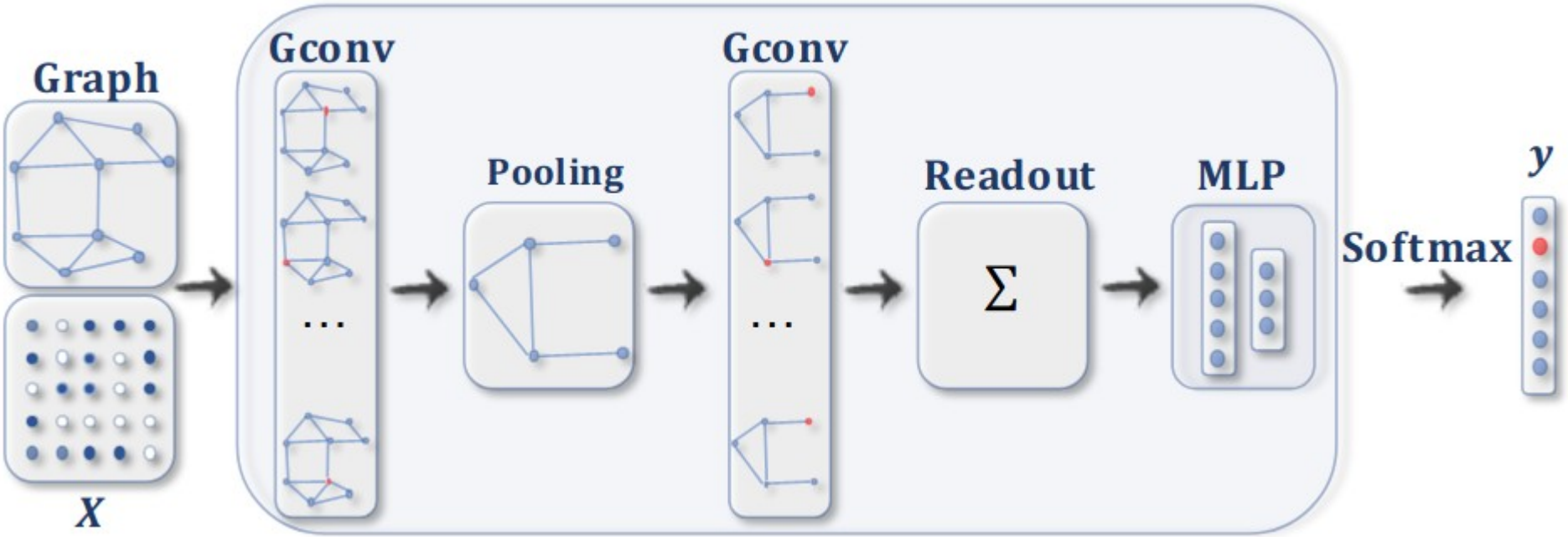


Image from: A Comprehensive Survey on Graph Neural Networks, arXiv:1901.00596 [cs.LG]

Conclusion

References

- A Comprehensive Survey on Graph Neural Networks, Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, Philip S. Yu, arXiv:1901.00596 [cs.LG]
- Representation Learning on Graphs: Methods and Applications,William L. Hamilton, Rex Ying, Jure Leskovec, arXiv:1709.05584 [cs.SI]
- J. Leskovec, J. Kleinberg and C. Faloutsos. Graph Evolution: Densification and Shrinking Diameters. ACM Transactions on Knowledge Discovery from Data (ACM TKDD), 1(1), 2007.
- Yin, Austin R. Benson, Jure Leskovec, and David F. Gleich. "Local Higher-order Graph Clustering." In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2017.
- J. Leskovec, D. Huttenlocher, J. Kleinberg. Signed Networks in Social Media. CHI 2010.
- J. Leskovec, D. Huttenlocher, J. Kleinberg. Predicting Positive and Negative Links in Online Social Networks. WWW 2010.

References

- https://snap.stanford.edu/data/email-Eu-core.html
- https://neo4j.com/blog/graph-algorithms-neo4j-label-propagation/
- https://neo4j.com/developer/
- https://neo4j.com/docs/graph-data-science/1.0/
- https://www.tweetbinder.com/blog/top-twitter-accounts/
- https://neo4j.com/blog/graph-algorithms-neo4j-label-propagation/
- https://en.wikipedia.org/wiki/Six_degrees_of_separation

Thank You!

a.asadikorom@eng.basu.ac.ir