Introduction

CE642: Social and Economic Networks

Maryam Ramezani
Sharif University of Technology
maryam.ramezani@sharif.edu

Welcome!

- Introduction
  - My name: Maryam Ramezani
    - I graduated from Sharif University of Technology
      - B.Sc. Thesis: Social Networks Analysis and Modeling Tool
      - M.Sc. Thesis: Community Detection in Social Networks by Using Information from Diffusion Network
      - PhD. Thesis: Diffusion in Social Networks Based on Partial Information

Course overview

- The course goal
  - To read some recent and interesting papers on complex networks
  - Understand the underlying techniques
  - Think about interesting problems
- Prerequisites:
  - Mathematical background on algorithm, graph theory, probabilities, linear algebra
  - The course will be more “theoretical”, but your project may be more “practical”.
  - You should be able to write non-trivial programs (in Python), you will work with Gephi and NetworkX too.
- Style
  - Both slides and blackboard

Course overview

- The course is self-contained.
- No single topic is too hard by itself.
- But we will cover and touch upon **many topics** and this is what makes the course hard.
- Website: https://sut-ce-courses.github.io/socialnetwork/
  - Slides will be posted after the class
  - Papers and pointers to additional literature
- Quera Page: https://quera.org/course/20942
  - Password: 8031404

Topics

- Graph Theory
- Graph-searching Algorithms
- Network metrics
- Motif and Graph Clustering
- Community Detection Algorithms
- Random walk
- Graph models
- Game on Network
- Diffusion and Epidemics on Complex Networks

Topics

- Link Prediction
- Recommendation Systems
- Influence Maximization
- Outbreak Detection
- Representation Learning on Graphs
- Network Inference
- Message Passing and Node Classification
- Graph Convolutional Networks
- Knowledge Graphs and Meta Paths
- Fairness on Graph

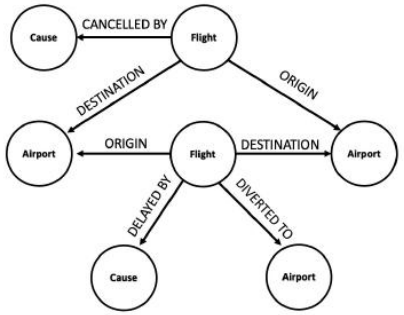
Grading

- Final grade (23 Points) will be composed of:
  - Homework: 7 Points
    - 4 written and coding assignments- Penalty: 10 days delayed then per hour - 0.5%
  - Review Quiz: 0.5 Points (1403/12/07)
  - Research Study: 3 Points
    - 4 series (1403/12/19 ,1404/01/26, 1404/02/23, 1404/03/13)
  - Midterm Exam: 5 Points (1404/02/02)
  - Final Exam: 5 Points (1404/04/03)
  - Research Project (Groups of 3): 2.5 Points (1404/04/13)

Why Graphs?

Graphs are a general language for describing and analyzing entities with relations/interactions

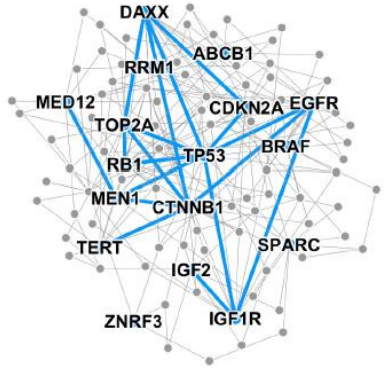
Many Type of Data are Graphs (1)



Event Graphs

Image credit: SalientNetworks

Computer Networks



Disease Pathways

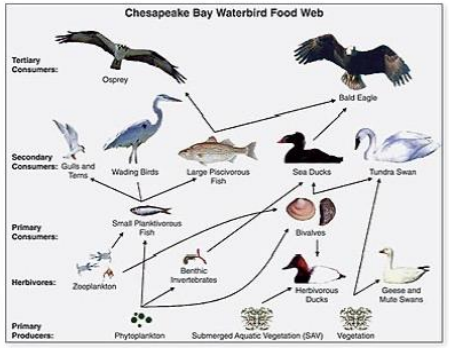


Image credit: Wikipedia

Food Webs

Image credit: Pinterest

Particle Networks

Image credit: visitlondon.com

Underground Networks

Many Type of Data are Graphs (2)

Image credit: Medium

Social Networks

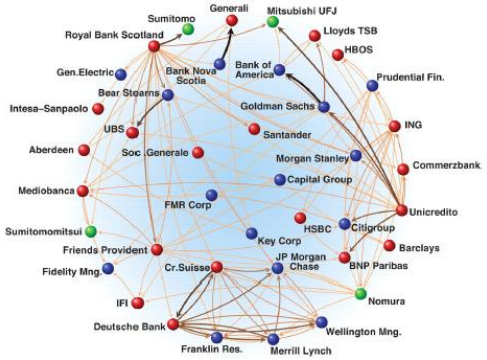


Image credit: Science

Economic Networks

Image credit: Lumen Learning

Communication Networks

Citation Networks

Image credit: Missoula Current News

Internet

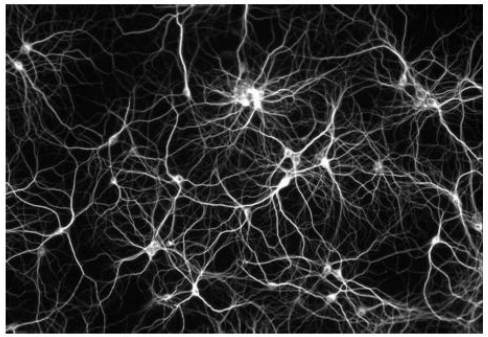


Image credit: The Conversation

Networks of Neurons

Many Type of Data are Graphs (3)

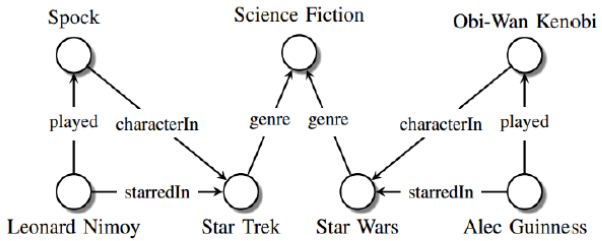


Image credit: Maximilian Nickel et al

Knowledge Graphs

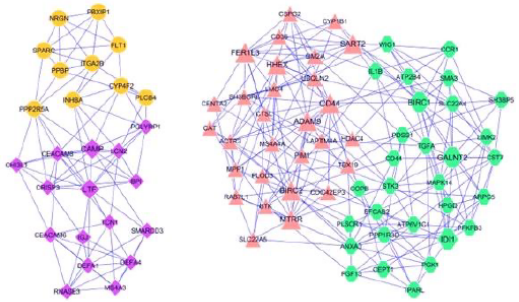


Image credit: ese.wustl.edu

Regulatory Networks

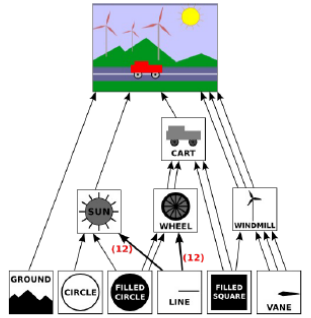


Image credit: math.hws.edu

Scene Graphs

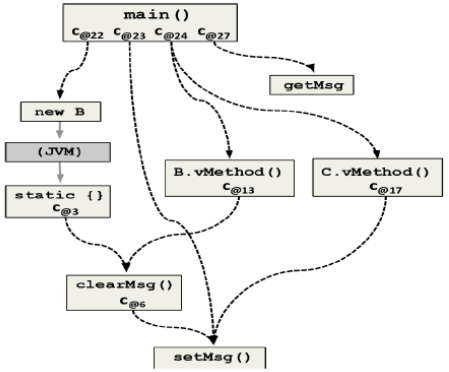


Image credit: ResearchGate

Code Graphs

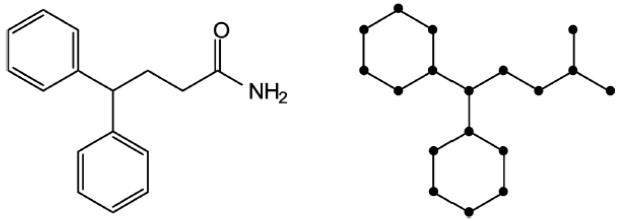


Image credit: MDPI

Molecules

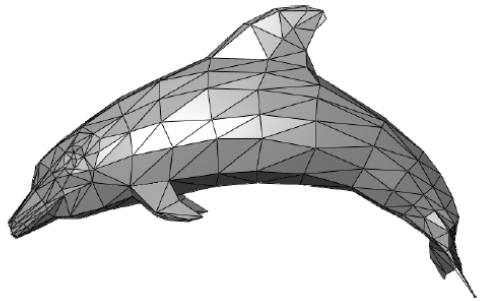
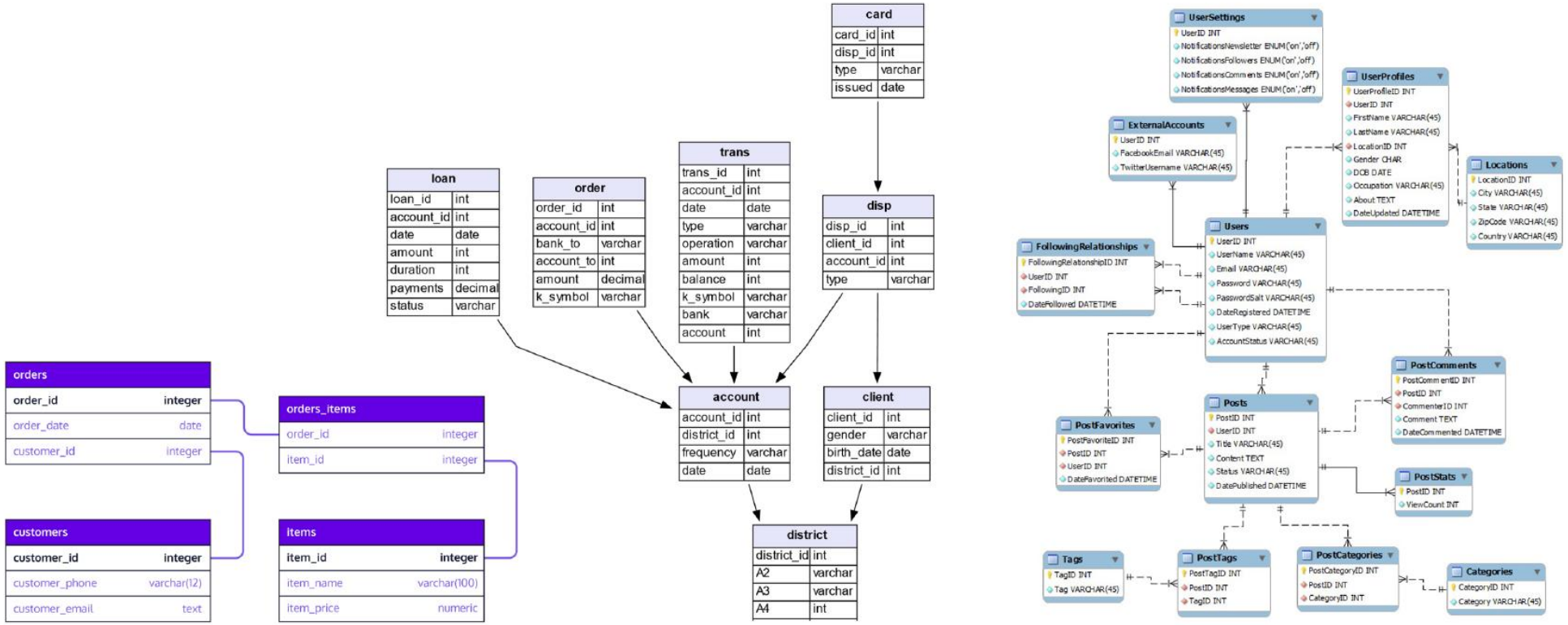


Image credit: Wikipedia

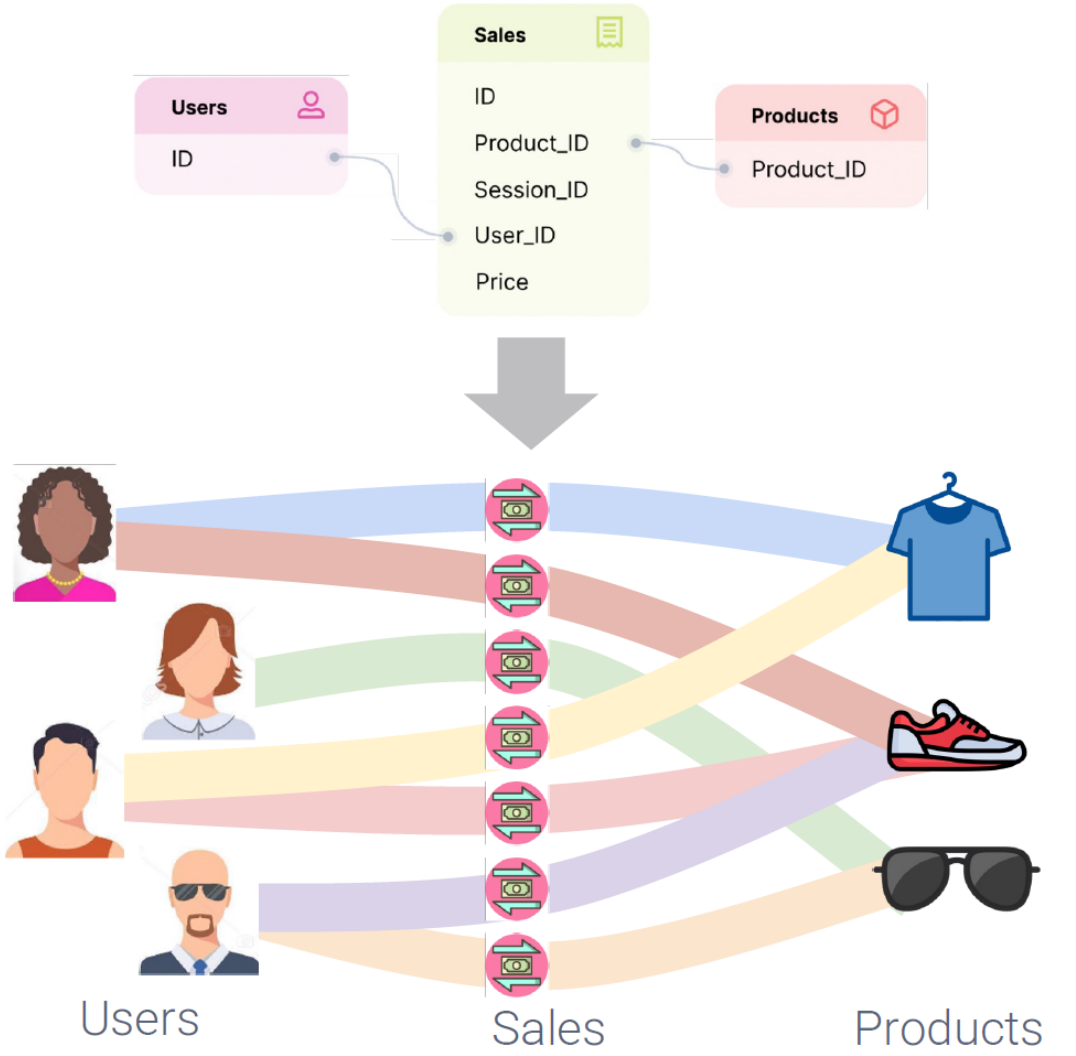
3D Shapes

Databases are Graphs!



Commerce Finance Social Media

Relational Deep Learning



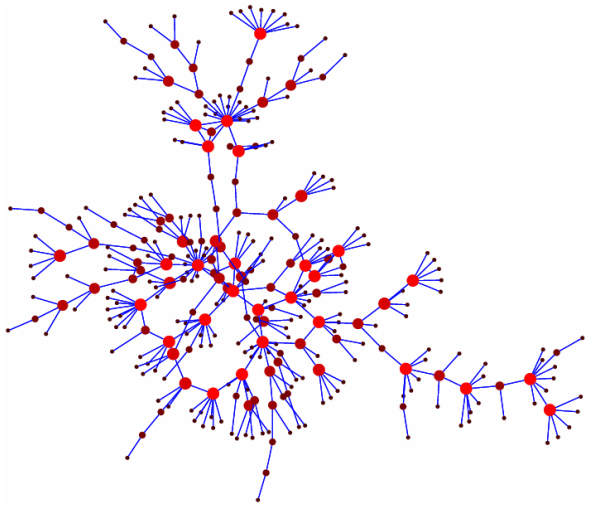
What is a network?

- Network: a collection of entities that are interconnected with links.
  - people that are friends
  - computers that are interconnected
  - web pages that point to each other
  - proteins that interact

- In mathematics, networks are called graphs, the entities are nodes, and the links are edges

Networks in the past

- Graphs have been used in the past to model existing networks (e.g., networks of highways, social networks)
  - Usually these networks were small
  - Network can be studied visual inspection can reveal a lot of information



Networks now

- More and larger networks appear
  - Products of technological advancement
    - e.g., Internet, Web
  - Result of our ability to collect more, better, and more complex data
    - e.g., gene regulatory networks
- Networks of thousands, millions, or billions of nodes
  - impossible to visualize

Understanding large graphs

- What are the statistics of real life networks?
- Can we explain how the networks were generated?

Measuring network properties

- Around 1999
  - Watts and Strogatz, Dynamics and small-world phenomenon
  - Faloutsos, On power-law relationships of the Internet Topology
  - Kleinberg et al., The Web as a graph
  - Barabasi and Albert, The emergence of scaling in real networks

Real network properties

- Most nodes have only a small number of neighbors (degree), but there are some nodes with very high degree (power-law degree distribution)
  - scale-free networks
- If a node x is connected to y and z, then y and z are likely to be connected
  - high clustering coefficient
- Most nodes are just a few edges away on average.
  - small world networks
- Networks from very diverse areas (from internet to biological networks) have similar properties
  - Is it possible that there is a unifying underlying generative process?

Generating random graphs

- Classic graph theory model (Erdös-Renyi)
  - each edge is generated independently with probability p

- Very well studied model but:
  - most vertices have about the same degree
  - the probability of two nodes being linked is independent of whether they share a neighbor
  - the average paths are short

Modeling real networks

- Real life networks are not “random”
- Can we define a model that generates graphs with statistical properties similar to those in real life?
  - a flurry of models for random graphs

Processes on networks

- Why is it important to understand the structure of networks?

- Epidemiology: Viruses propagate much faster in scale-free networks
- Vaccination of random nodes does not work, but targeted vaccination is very effective

Web search

- First generation search engines: the Web as a collection of documents
  - Suffered from spammers, poor, unstructured, unsupervised content, increase in Web size
- Second generation search engines: the Web as a network
  - use the anchor text of links for annotation
  - good pages should be pointed to by many pages
  - good pages should be pointed to by many good pages
    - PageRank algorithm, Google!

The future of networks

- Networks seem to be here to stay
  - More and more systems are modeled as networks
  - Scientists from various disciplines are working on networks (physicists, computer scientists, mathematicians, biologists, sociologist, economists)
  - There are many questions to understand.

Mathematical Tools

- Graph theory
- Probability theory
- Linear Algebra

Any Question?