Deep Learning for Network Biology

Marinka Zitnik and Jure Leskovec

Stanford University

This Tutorial

snap.stanford.edu/deepnetbio-ismb

ISMB 2018

July 6, 2018, 2:00 pm - 6:00 pm

This Tutorial

1) Node embeddings

- Map nodes to low-dimensional embeddings
- *Applications:* PPIs, Disease pathways

2) Graph neural networks

- Deep learning approaches for graphs
- *Applications:* Gene functions

3) Heterogeneous networks

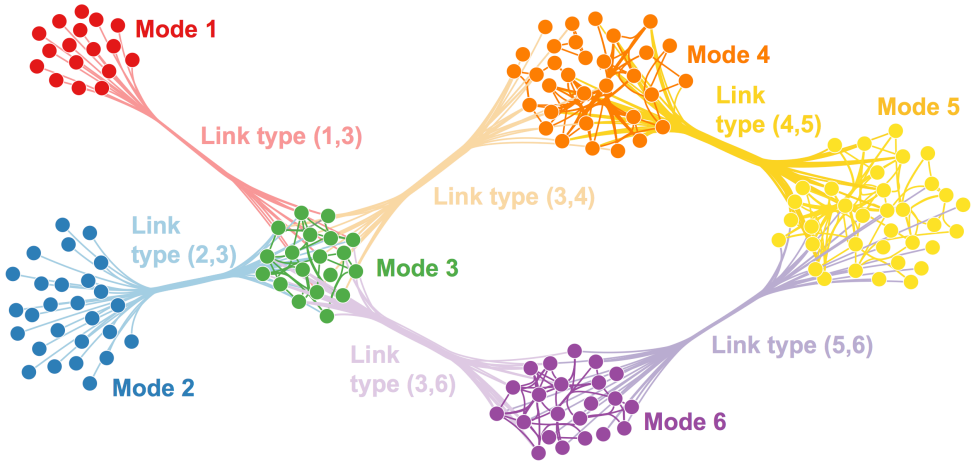
- Embedding heterogeneous networks
- *Applications:* Human tissues, Drug side effects

Part 3: Heterogeneous Networks

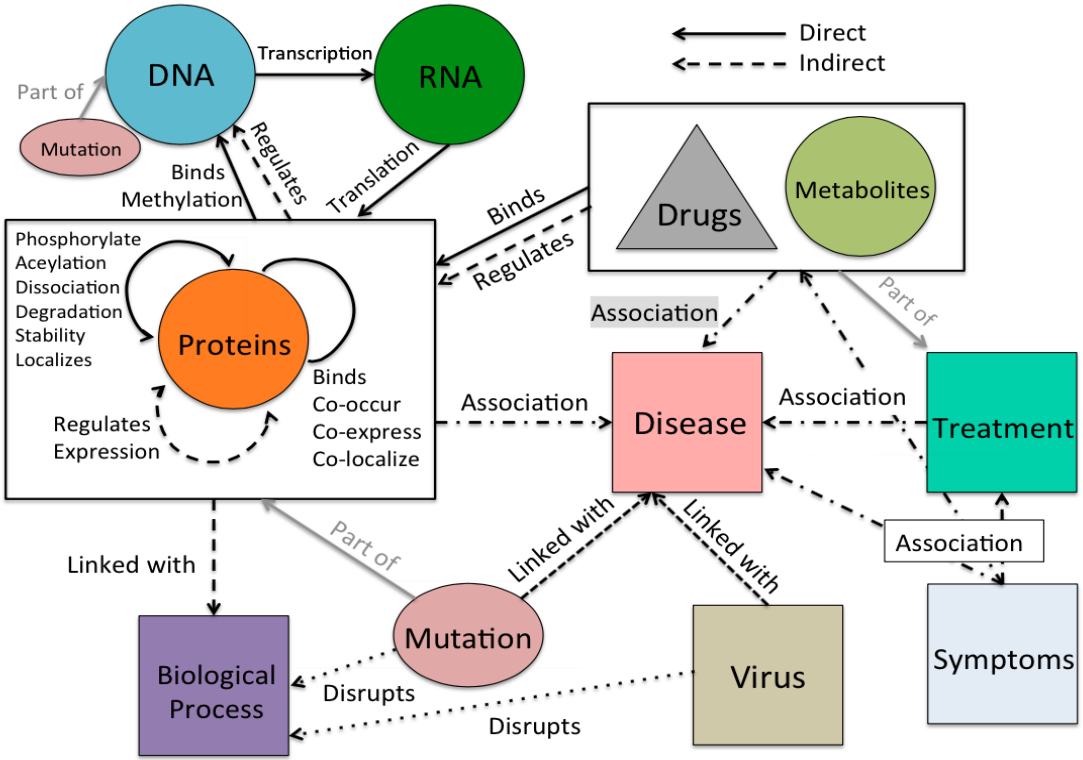
Homogeneous Nets

So far we have embedded homogeneous networks

Can we embed heterogeneous networks (i.e., het nets)?



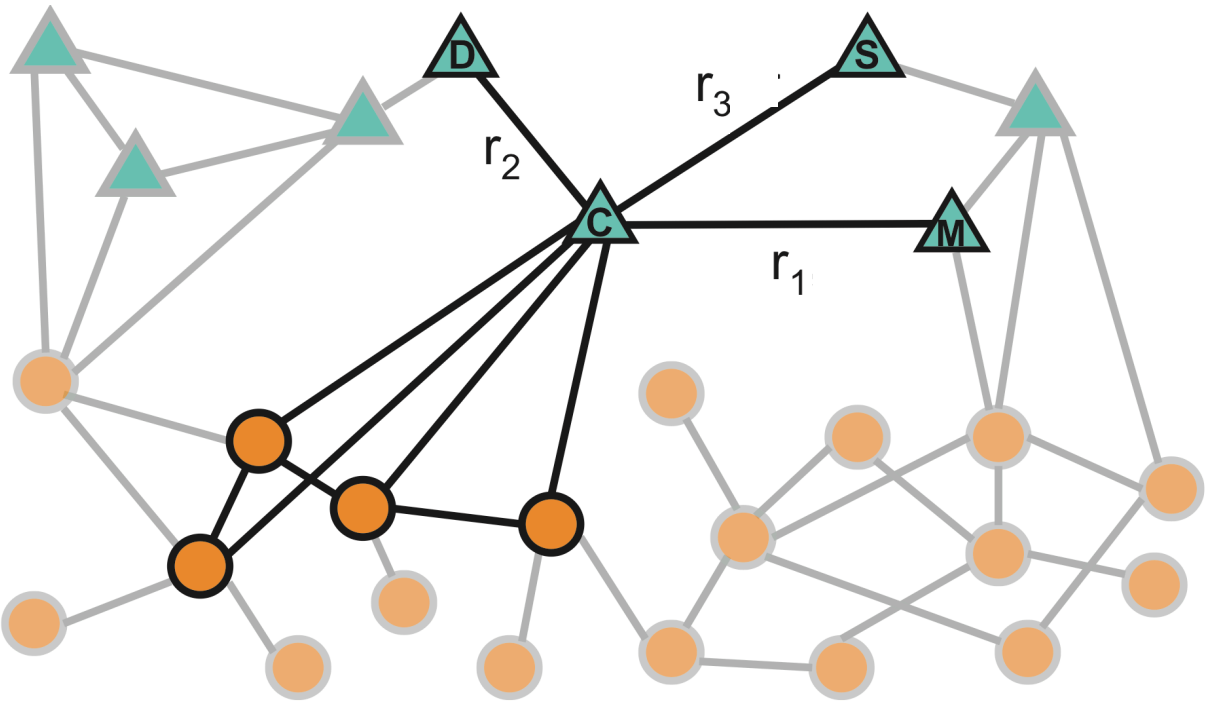
Many Het Nets in Biology



Setup

- Assume we have a graph G:
  - V_t is the vertex set for node type t
  - $\boldsymbol{A}_r$ is the adjacency matrix for edge type r
  - $\mathbf{X}_t \in \mathbb{R}^{m \times |V|}$ is a matrix of features for nodes of type t
    - Biologically meaningful node features:
      - E.g., immunological signatures, gene expression profiles, gene functional information
    - No features:
      - Indicator vectors (one-hot encoding of a node)

Example: Het Net



r_1 Gastrointestinal bleed side effect
r_2 Bradycardia side effect
r_3 Nausea side effect
r_4 Mumps side effect
Drug-protein interaction
Protein-protein interaction

Tutorial Resource

MAMBO: Multimodal biomedical networks

- Tool for construction, representation and analysis of large multimodal networks:
  - Nets with millions of nodes and billions of edges
  - Nets with thousands of modes (i.e., entity types) and links (i.e., relationship types)
- Network analytics through **SNAP**

http://snap.stanford.edu/mambo

Outline of This Section

1. **Shallow embeddings** for het nets:
   - OhmNet
   - Metapath2vec
2. **Deep embeddings** for het nets:
   - Decagon

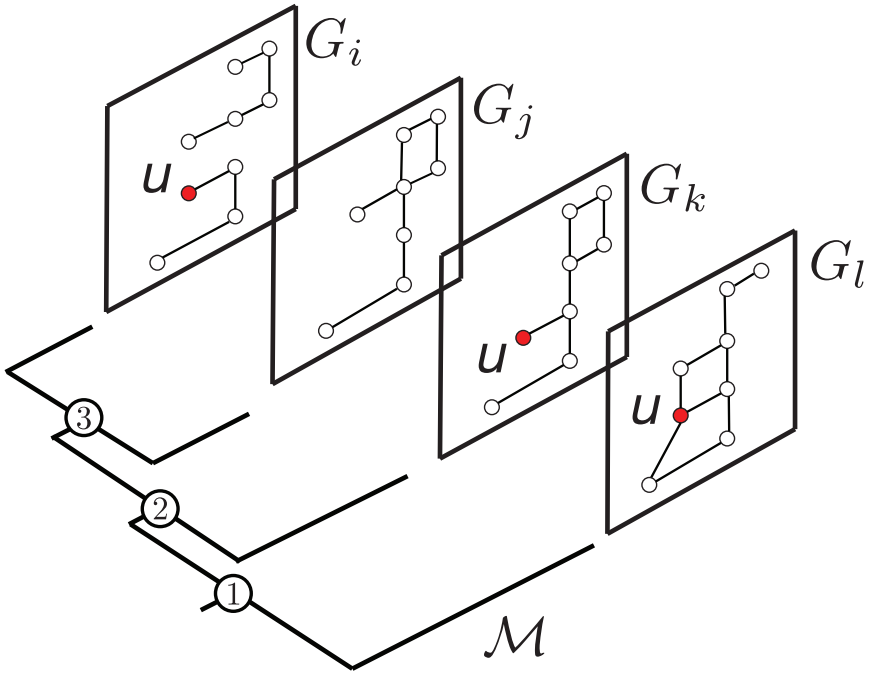
OhmNet

Based on material from:

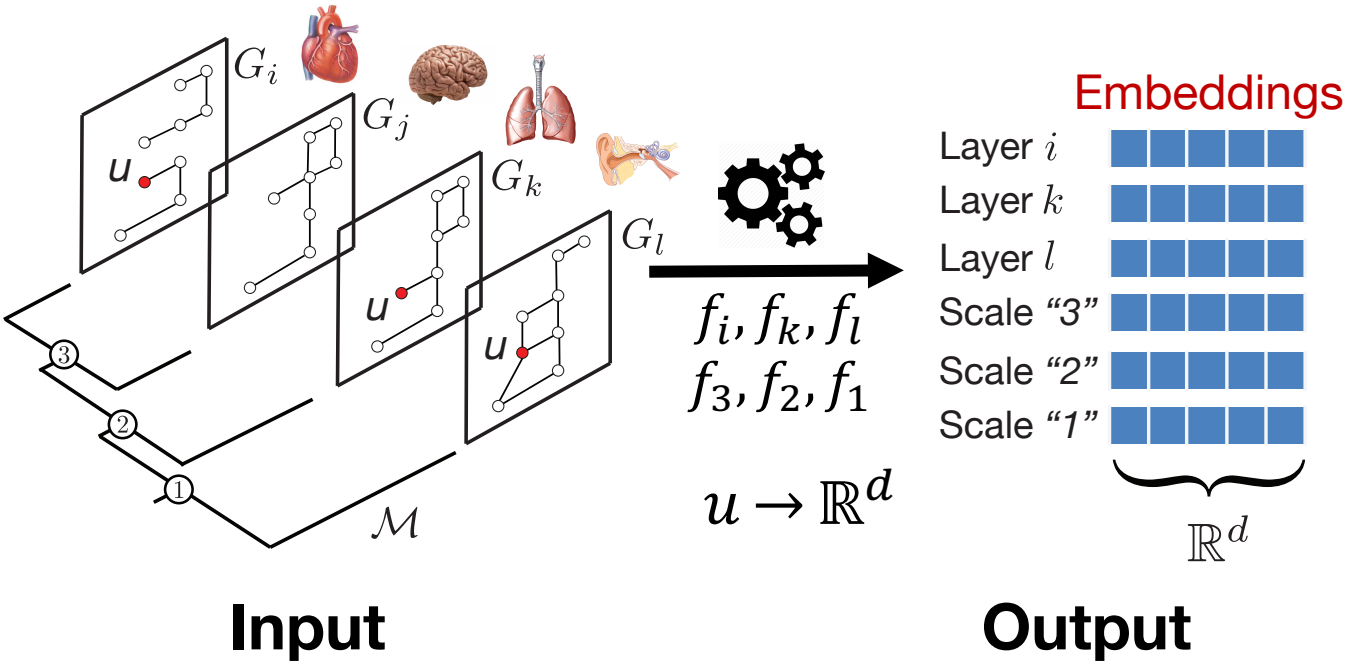
- Zitnik et al., 2017. Predicting multicellular function through multi-layer tissue networks. *ISMB & Bioinformatics.*

Embedding Layered Graphs

Extending node2vec to **multi-layer graphs**



OhmNet: Multi-Layer Graphs

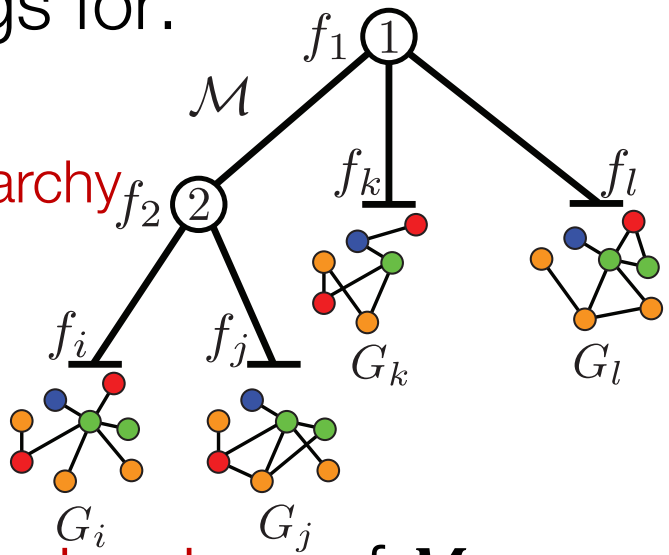


Input **Output**

How to learn mapping functions f_i?

Multi-Layer Graphs

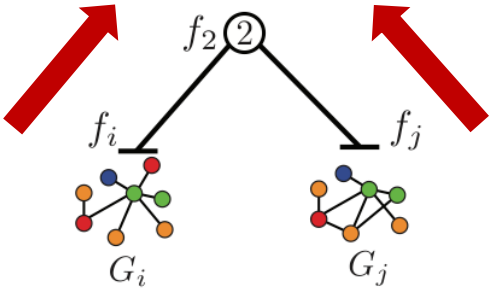
- **Input:** Given graphs G_i and hierarchy M
- **Output:** Embeddings for:
  - Nodes in each graph
  - Nodes in each sub-hierarchy



- Capture hierarchical structure of M

Multi-Layer Graphs

- For graphs G_i:
  - Use node2vec's biased walks (**see Part T1**)
- For hierarchy M:
  - Encode dependencies between graphs
  - **Recursive regularization:** embeddings at level i are encouraged to be similar to embeddings in i's parent in the hierarchy

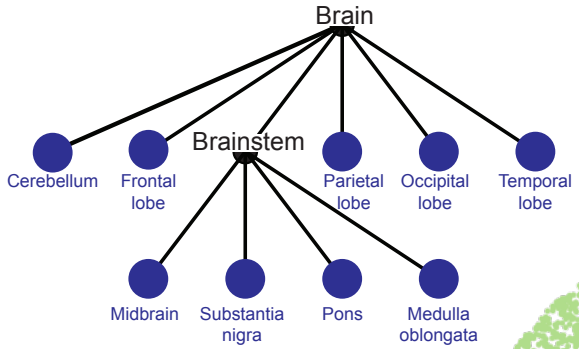


Random Walk Optimization

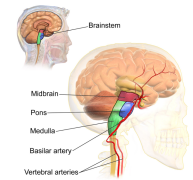
- Given simulated random walks for each graph:
  - Optimize node embeddings as described in **Part T1**
  - **Extra:** Include terms for recursive regularization in the loss function

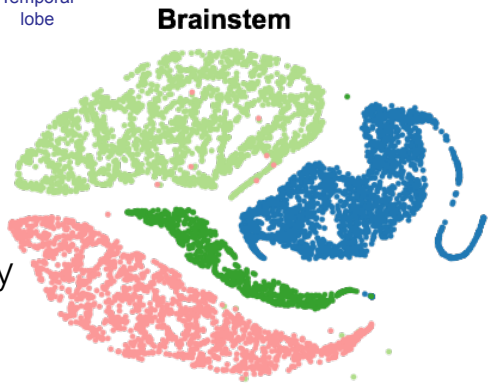
Example: Brain Networks

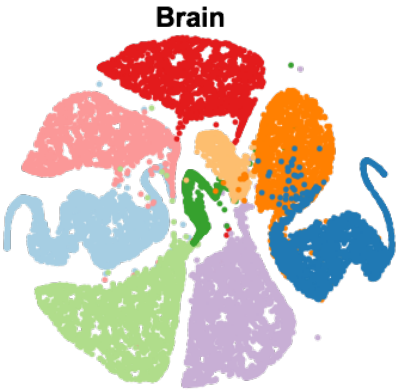
Do embeddings match human anatomy?



9 brain tissue PPI networks in a two-level tissue hierarchy







Cerebellum
Medulla oblongata
Substantia nigra
Frontal lobe
Temporal lobe
Pons
Parietal lobe
Occipital lobe
Midbrain

Metapath2vec

Based on material from:

- Dong et al., 2017. metapath2vec: Scalable representation learning for heterogeneous networks. *KDD.*

Metapaths

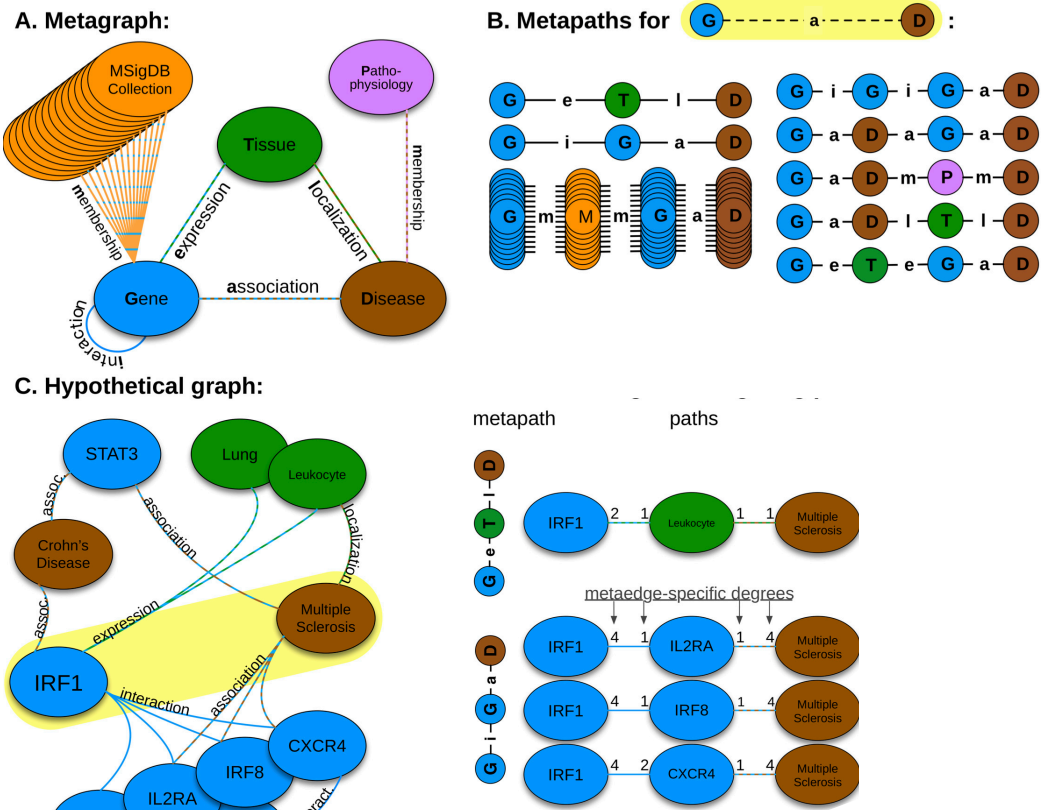


Image from: Himmelstein et al. 2015. Heterogeneous network edge prediction: A data integration approach to prioritize disease-associated genes. *PLoS Comp Bio*.

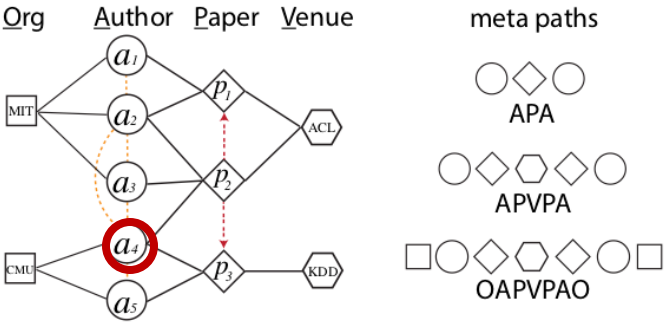
Metapath2vec: Two Main Steps

Extending node2vec to **het nets:**

1. **Metapath-based random walks**
   - Specify a metapath of interest
   - Run random walks that capture structural correlations between different node types
2. **Random walk optimization**
   - Given the random walks, optimize node embeddings (similar to **Part T1)**

Step 1: Run Random Walks

- Given a metapath:
  - E.g., **OAP**VPAO



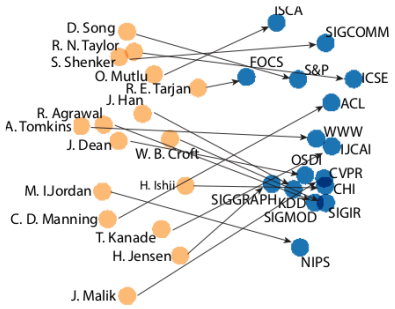
- What is the next step of a walker on node a_4 that transitioned from node CMU?
  - **Standard random walk:** The next step can be all types of nodes surrounding it:
    - a_2, a_3, a_5, p_2, p_3, and CMU
  - **Metapath-based random walk:** The next step can only be a paper node (P), given that its current node is an author node a_4 (A) and its previous step was an organization node CMU (O):
    - Follow the semantics of this metapath

Step 2: Optimize

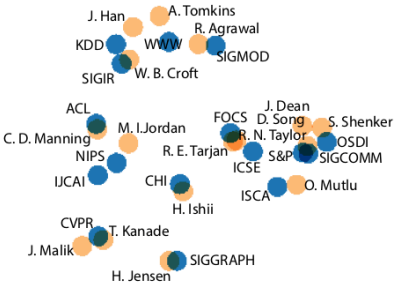
1. Simulate many **metapath-based random walks** starting from each node
2. For each node u, get $\mathbf{N}_t(u)$ as a nodes of type t that are visited by random walks starting at u
3. **For each node u,** learn its embedding by **predicting which nodes are in $\mathbf{N}_t(u)$:**

$$\mathcal{L} = \sum_{u \in V} \sum_{t \in V_t} \sum_{v \in N_t(u)} -\log(P(v|\mathbf{z}_u))$$

Metapath2vec: Example



DeepWalk / node2vec



metapath2vec

- 2D projections of the learned embeddings for:
  - 16 CS conferences and corresponding high-profile researchers in each field
- Metapath2vec:
  - Groups author-conference pairs closely
  - Automatically organizes these two types of nodes
  - Learns internal relationships between them:
    - E.g., J. Dean → OSDI
    - E.g., C. D. Manning → ACL
- Not possible using methods for homogeneous networks

Outline of This Section

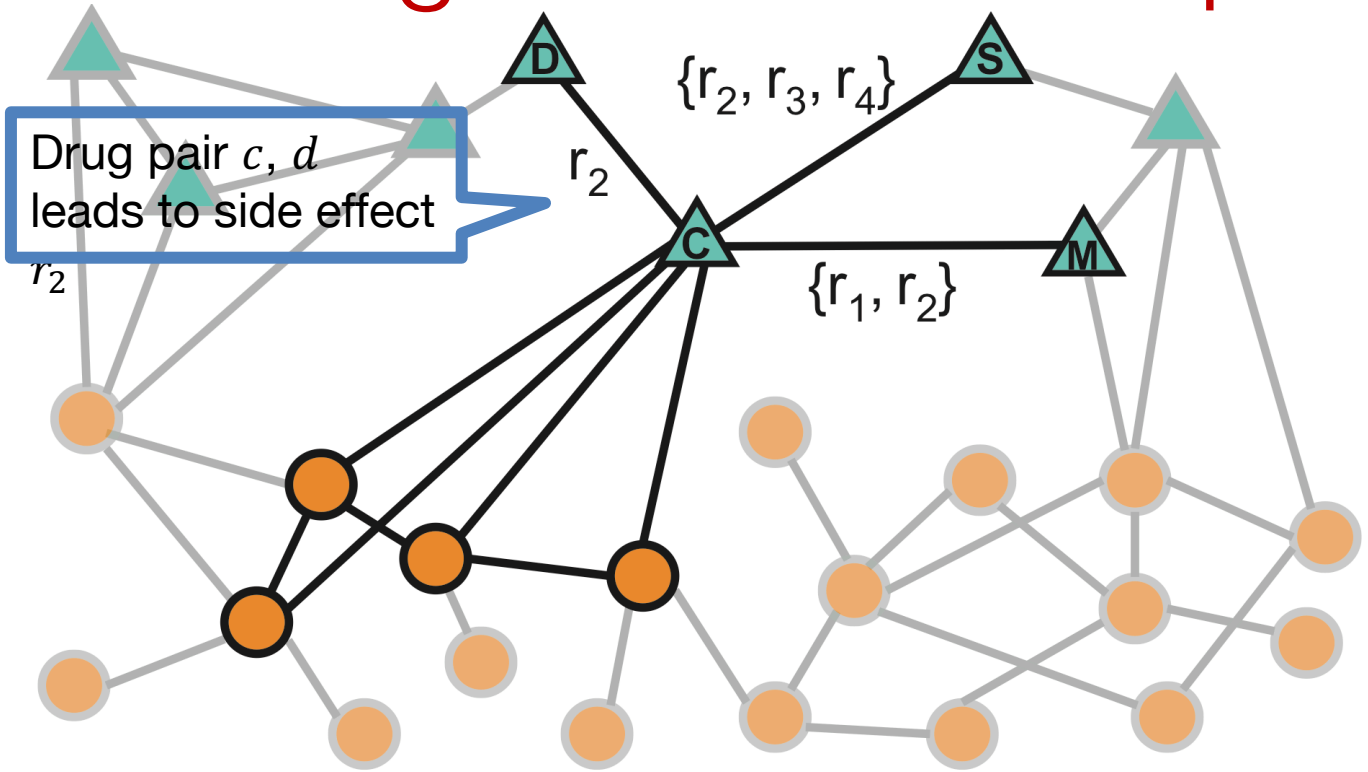
1. **Shallow embeddings** for het nets:
   - OhmNet
   - Metapath2vec
2. **Deep embeddings** for het nets:
   - Decagon

Deep Embeddings for Heterogeneous Graphs

Based on material from:

- Zitnik et al., 2018. Modeling polypharmacy side effects with graph convolutional networks. *ISMB & Bioinformatics.*

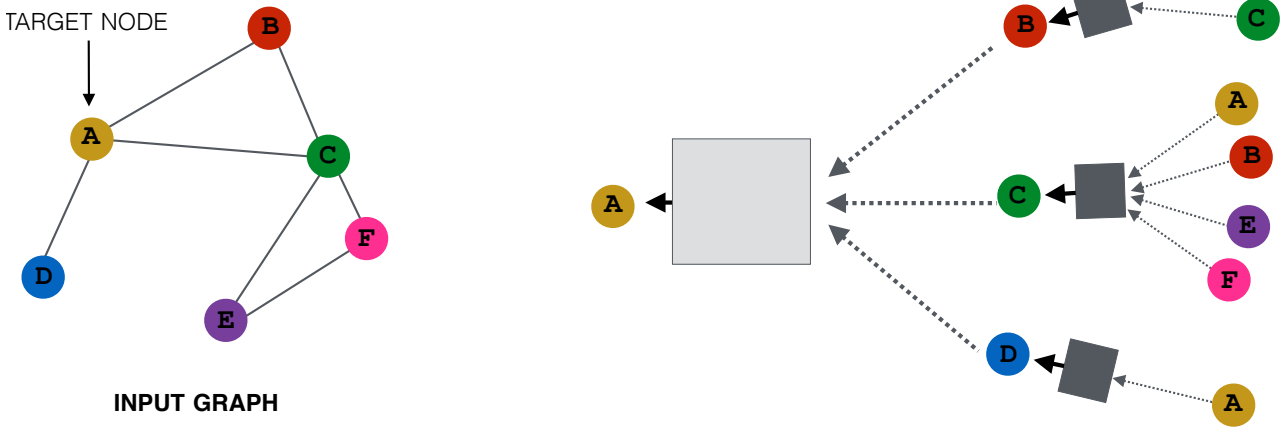
Running Het Net Example



r_1 Gastrointestinal bleed side effect
r_2 Bradycardia side effect
r_3 Nausea side effect
r_4 Mumps side effect
Drug-protein interaction
Protein-protein interaction

Idea: Aggregate Neighbors

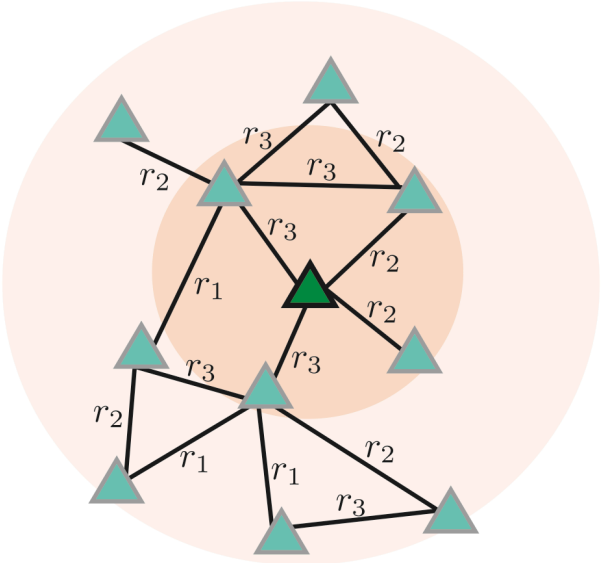
Key idea: Generate node embeddings based on **network neighborhoods separated by edge type**



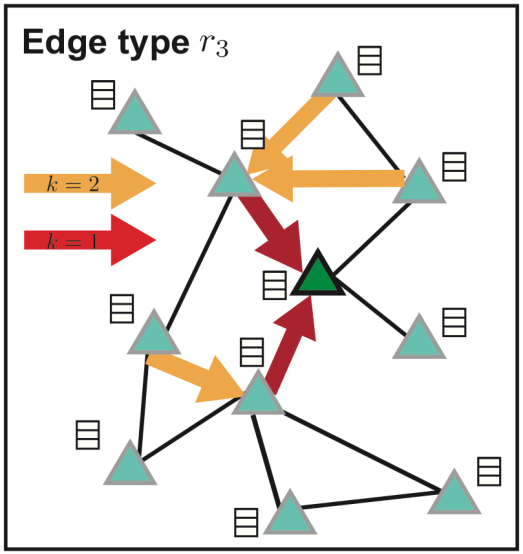
Idea: Aggregate Neighbors

Each edge type is **modeled separately**

A node's neighborhood defines a **computation graph**

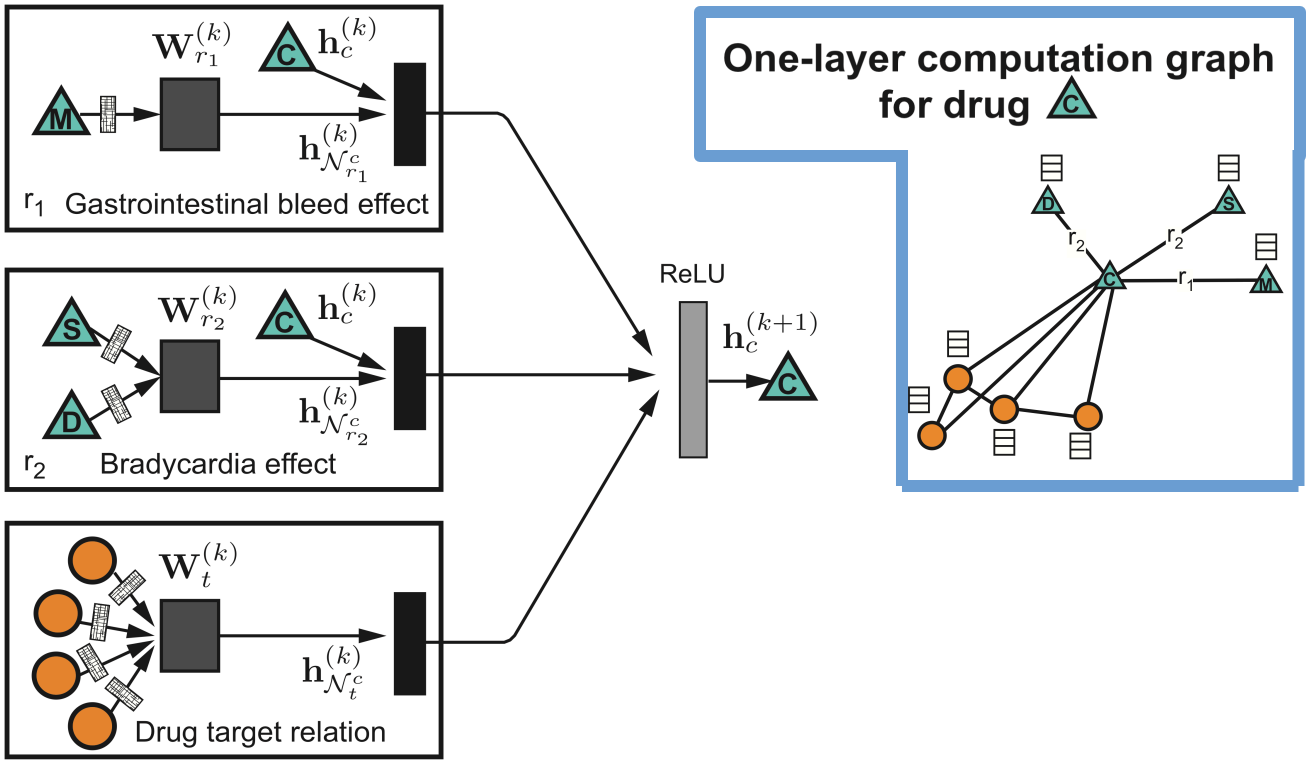


Determine a node's computation graph



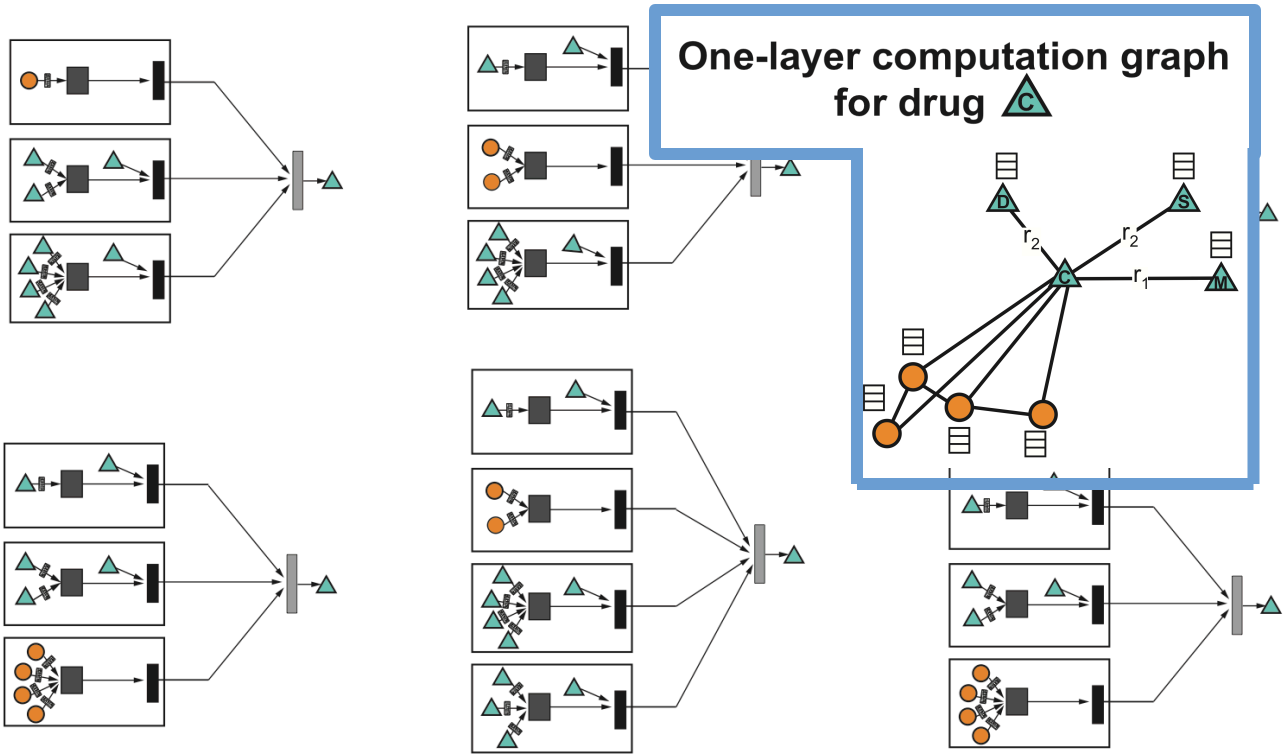
Learn how to transform and propagate information across the graph

Example: Aggregation



$\mathbf{W}_{r_i}^{(k)}$ Neural network weight matrices

Example: Aggregation



$\mathbf{W}_{r_i}^{(k)}$ Neural network weight matrices

An example batch of computation graphs

The Math: Deep Encoder

- **Approach:** Average **neighbor messages for each edge type** and apply a neural network

$$\mathbf{h}_v^0 = \mathbf{x}_v$$

Initial 0-th layer embeddings are equal to node features

Previous layer embedding of v

$$\mathbf{h}_v^k = \phi\left(\sum_r \left(\mathbf{W}_r^{k-1} \sum_{u \in N_r(v)} \frac{\mathbf{h}_u^{k-1}}{\sqrt{|N_r(v)||N_r(u)|}} + \frac{\mathbf{h}_v^{k-1}}{|N_r(v)|}\right)\right)$$

Aggregate neighbor's previous layer embeddings

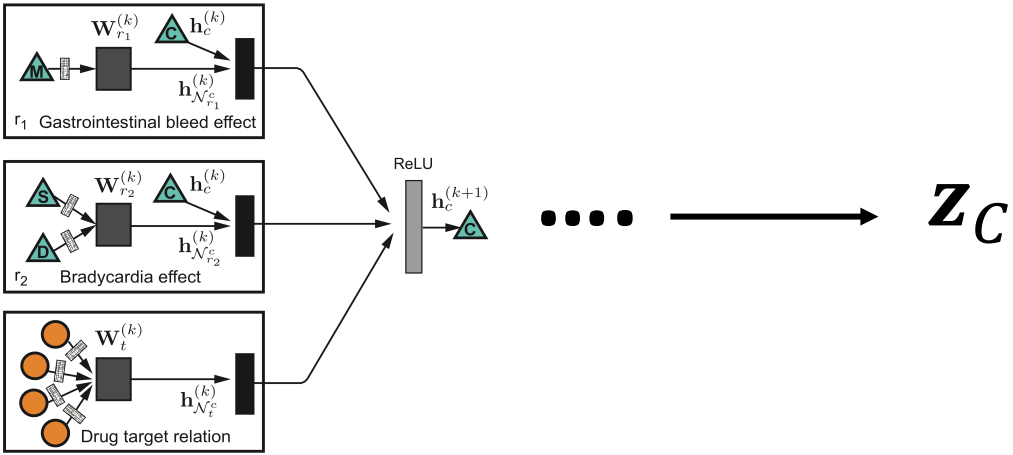
$$\mathbf{z}_v = \mathbf{h}_v^K$$

Embedding after K layers of neighborhood aggregation

Non-linearity (e.g., ReLU)

Training the Model

How do we train the model to generate embeddings?

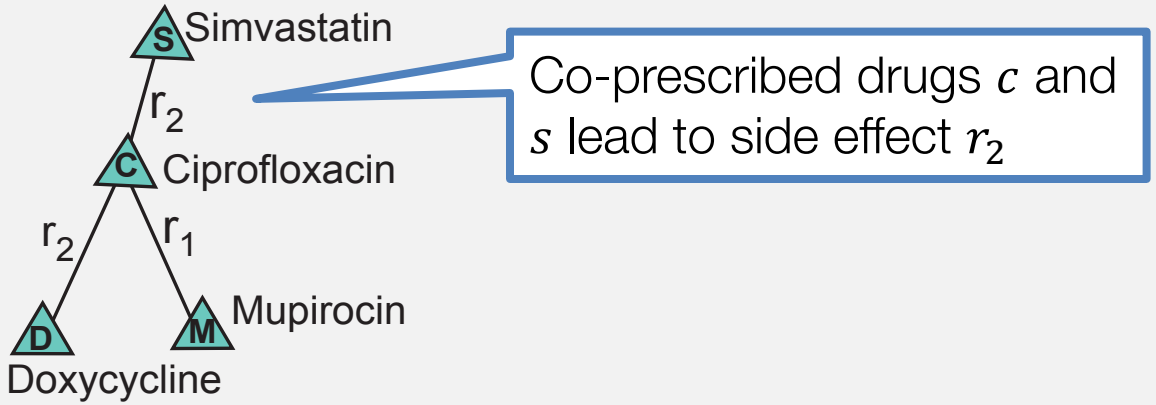


Need to define a loss function on the embeddings!

Example: Drug Side Effects

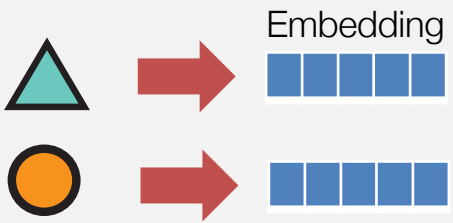
Goal: Predict labeled edges between drug nodes

Query: Given a drug pair c, s, how likely does an edge (c, r_2, s) exist?

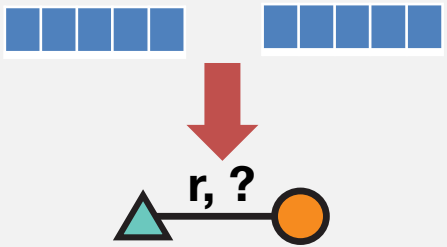


Example: Drug Side Effects

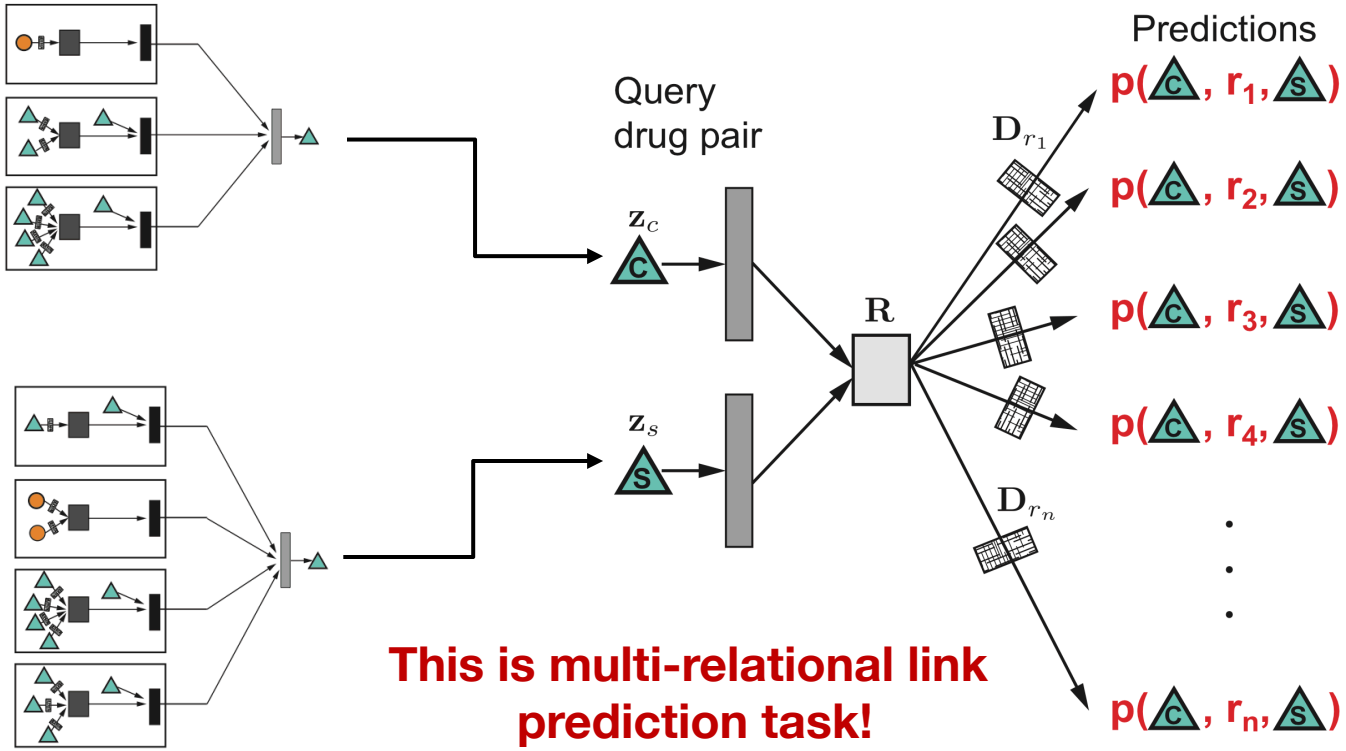
1) Take the graph and learn a d-dimensional vector (*embedding*) for every node



2) Use the learned embeddings to predict side effects of drug pairs



Example: Drug Side Effects



Modeling Polypharmacy Side Effects with Graph Convolutional Networks

July 10, 2018 at 12:20 pm

http://snap.stanford.edu/decagon

Outline of This Section

1. **Shallow embeddings** for het nets:
   - OhmNet ✓
   - Metapath2vec
2. **Deep embeddings** for het nets:
   - Decagon ✓

This Tutorial

1) Node embeddings

- Map nodes to low-dimensional embeddings
- *Applications:* PPIs, Disease pathways

2) Graph neural networks

- Deep learning approaches for graphs
- *Applications:* Gene functions

3) Heterogeneous networks

- Embedding heterogeneous networks
- *Applications:* Human tissues, Drug side effects

PhD Students


Claire Donnat


Mitchell Gordon


David Hallac


Emma Pierson


Geet Sethi


Himabindu Lakkaraju


Rex Ying


Tim Althoff


Will Hamilton


Alex Porter

Post-Doctoral Fellows


Baharan Mirzasoleiman

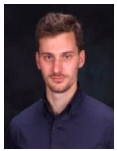

Marinka Zitnik


Michele Catasta


Srijan Kumar

Research Staff


Stephen Bach


Adrijan Bradaschia


Rok Sosic

Industry Partnerships

Funding

Collaborators

Dan Jurafsky, Linguistics, Stanford University
Christian Danescu-Miculescu-Mizil, Information Science, Cornell University
Stephen Boyd, Electrical Engineering, Stanford University
David Gleich, Computer Science, Purdue University
VS Subrahmanian, Computer Science, University of Maryland
Sarah Kunz, Medicine, Harvard University
Russ Altman, Medicine, Stanford University
Jochen Profit, Medicine, Stanford University
Eric Horvitz, Microsoft Research
Jon Kleinberg, Computer Science, Cornell University
Sendhill Mullainathan, Economics, Harvard University
Scott Delp, Bioengineering, Stanford University
Jens Ludwig, Harris Public Policy, University of Chicago

Many interesting high-impact projects
in Machine Learning and Large Biomedical Data

Applications: Precision Medicine & Health, Drug Repurposing,
Drug Side Effect modeling, Network Biology, and many more