Module V: Knowledge Graph Inference and Applications

2:05 pm - 3:10 pm

Module 5 Overview

KG Inference and Applications

Knowledge Graph Inference

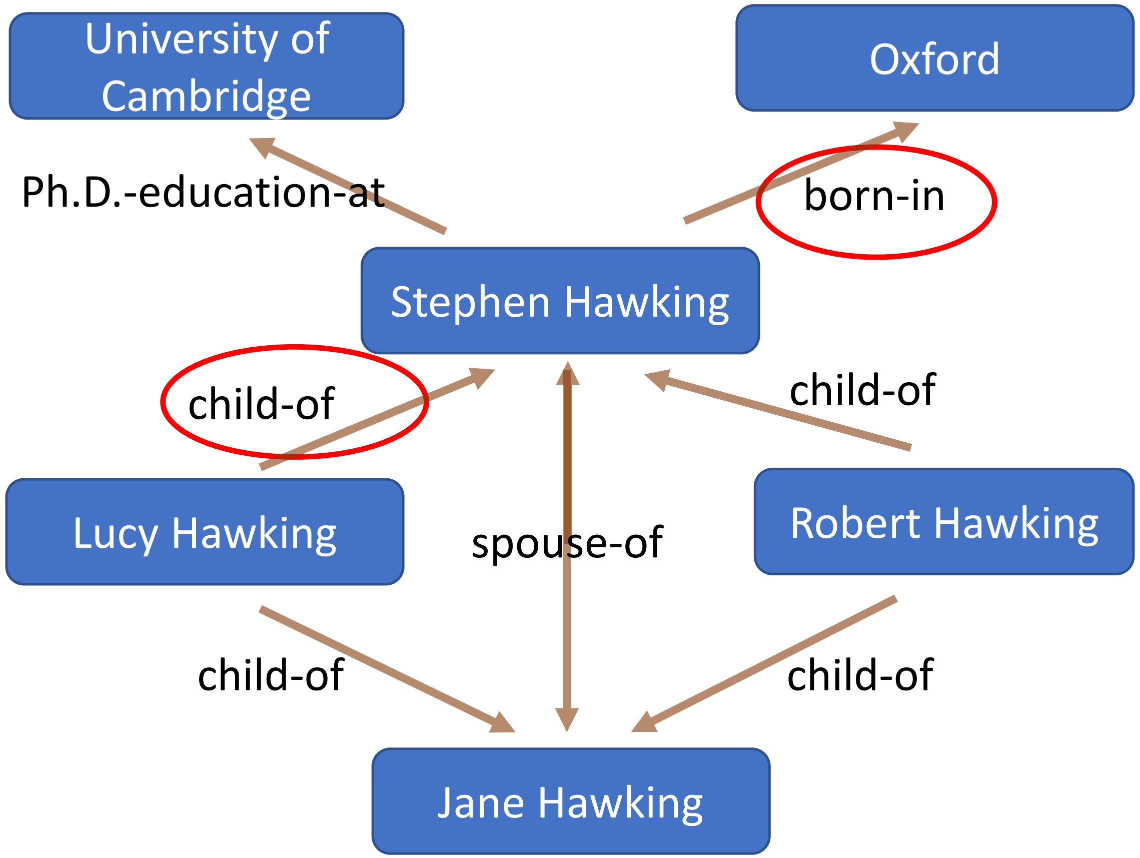
- What & Why
- How
  - Problem formulation & Overview
  - Knowledge Graph Embedding

Knowledge Graph Applications

- Entity Recommendation

Lab 5 – Paper Recommendation in ***MAG***

Knowledge Graph Inference -- WHAT



Knowledge Graph Inference -- WHY

- Knowledge Base is largely incomplete
- Need systematic and scalable approaches to complete knowledge graph

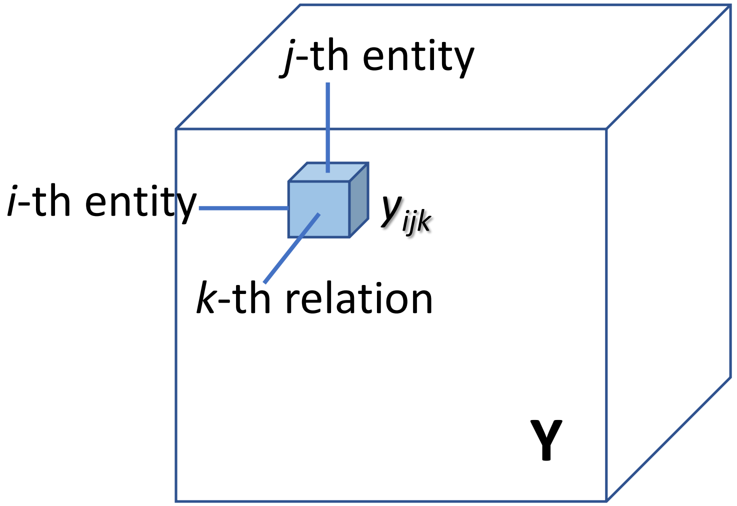
Freebase

Relation	Percentage unknown	
	All 3M	*Top 100K*
PROFESSION	68%	24%
PLACE OF BIRTH	71%	13%
NATIONALITY	75%	21%
EDUCATION	91%	63%
SPOUSES	92%	68%
PARENTS	94%	77%
CHILDREN	94%	80%
SIBLINGS	96%	83%
ETHNICITY	99%	86%

Incompleteness of Freebase for some relations that apply to entities of type PERSON. Left: all 3M Freebase PERSON entities. Right: only the 100K most frequent PERSON entities.

West, et al., WWW'14, Knowledge Base Completion via Search-Based Question Answering

Knowledge Graph Inference -- HOW



$$Y_{ijk} = \begin{cases} 1, & \text{if the triple } (e_i, r_k, e_j) \text{ exists;} \\ 0, & \text{otherwise.} \end{cases}$$

Element-wise

$$Y \in \{0, 1\}^{Ne \times Ne \times Nr}$$

adjacency tensor
(adjacency matrix)

$$P(Y)$$

Estimate the joint-distribution

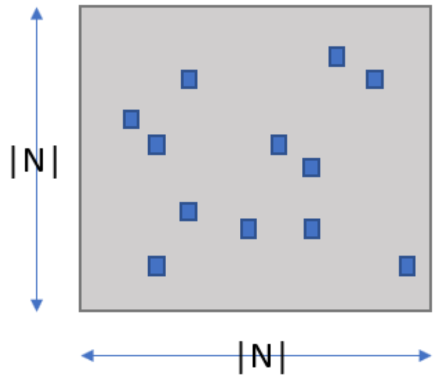
$$P(yijk)$$

Predict *unobserved* triples

PROBLEM FORMULATION -- *STATISTICAL RELATIONAL LEARNING*

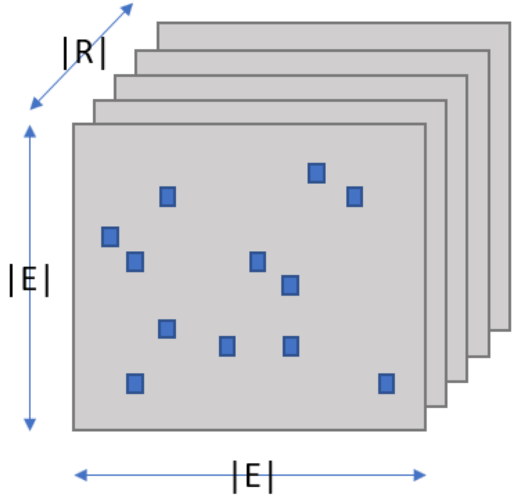
Knowledge Graph Inference -- HOW

- Within existing KG
  - Graph feature model
    - “Similar” entities
      - Local – common neighbors
      - Global – random walk
      - Quasi-local – random walk with bounded length
  - Latent feature model
- Use external sources / information
  - QnA system



Graph

|N|: Number of Nodes in Gra



Knowledge Graph

|E|: Number of Entities in Knowledge Graph
|R|: Number of Relations in Knowledge Graph

Knowledge Graph Inference -- HOW

LATENT FEATURE MODELS

- ***Entity*** Representation
  - Low dimensional vector: e_i
  - Initialization
    - Random
    - Average word vector with pre-trained vectors (V_{word}), e.g.

$$e_{homo\ sapiens} = 0.5 \times (V_{homo} + V_{sapiens})$$

- ***Relation type*** representation
  - Each relation type as ***matrix***:
    - W_k : bilinear weight matrix
    - A_k : linear feature map
  - Each relation type as ***vector***: r_k

- ***Entity-Relation*** interaction
  - Linear : $A_k\, e_i$
  - Bilinear: $e_i^T W_k\, e_j$

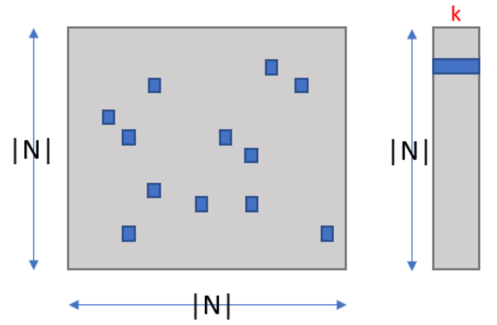
- ***Scoring function***
  - Margin-based ranking loss *Maximize the margin btw existing & non-existing triples*

$$\sum_{(s,r,o)\in T} \sum_{(s',r,o')\in T'(s,r,o)} max(0, 1 + f(s',r,o') - f(s,r,o))$$

  - Negative sampling loss *Negative log-likelihood of the correct triples & sampled corrupted triples*

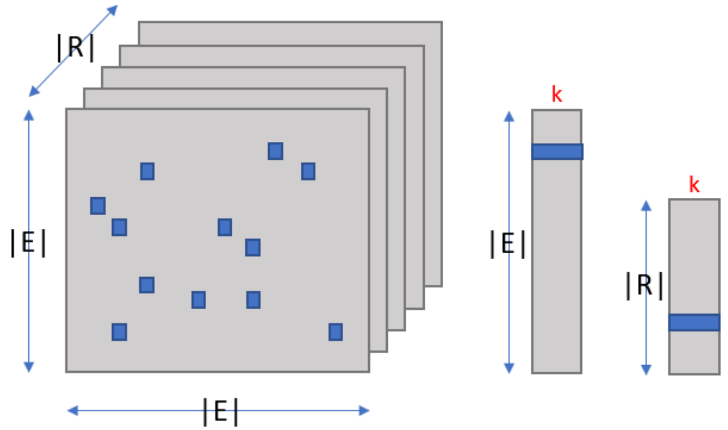
$$-\sum_{(s,r,o)\in T} \left(\log \sigma(f(s,r,o)) + \sum_{(s',r,o')\in T'(s,r,o)} \log \sigma(-f(s',r,o')) \right)$$

Graph



|N|: Number of Nodes in Graph
k: Dimensionality

Knowledge Graph



|E|: Number of Entities in Knowledge Graph
|R|: Number of Relations in Knowledge Graph
k: Dimensionality

Knowledge Graph Inference -- HOW

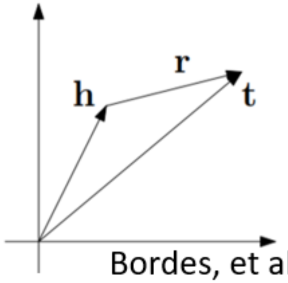
LATENT FEATURE MODELS - MILESTONES

Time

DISTMULT

2015

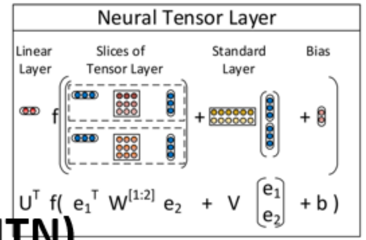
Yang, et al. ***ICLR***, Embedding entities and relations for learning and inference in knowledge bases



TransE

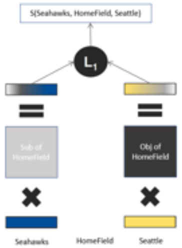
2013

Bordes, et al. ***NIPS***, Translating embeddings for modeling multi-relational data



Neural Tensor Networks (NTN)

Socher, et al. ***NIPS***, Reasoning with neural tensor networks for knowledge base completion

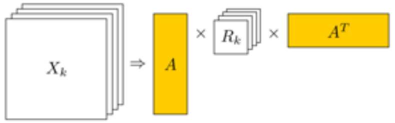


Structured Embedding (SE)

2011

Bordes, et al. ***AAAI***, Learning structured embedding of knowledge base

Tensor Factorization



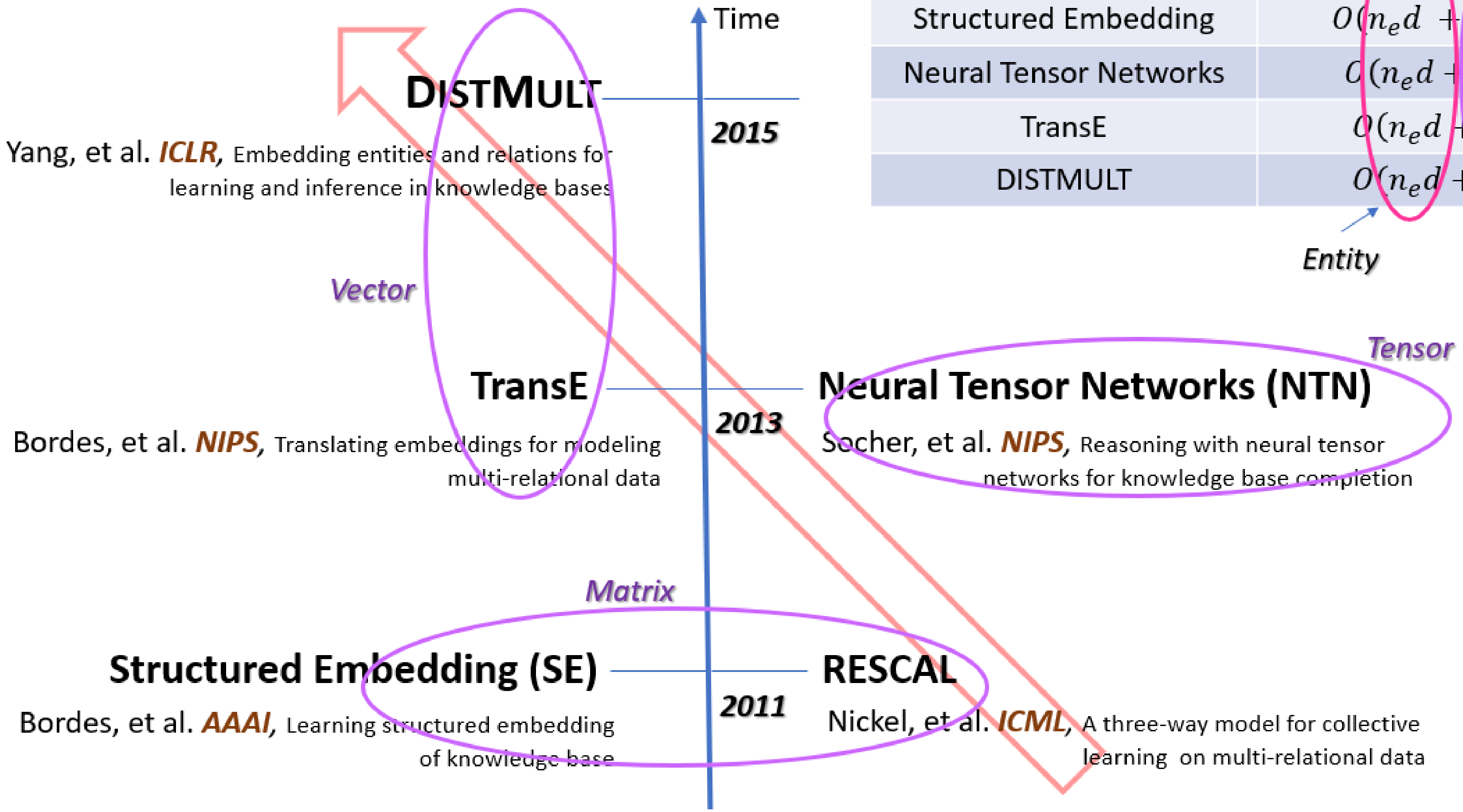
RESCAL

Nickel, et al. ***ICML***, A three-way model for collective learning on multi-relational data

Knowledge Graph Inference -- HOW

LATENT FEATURE MODELS - SCALABILITY

Scalability [# of params]



Method	# of Parameters
RESCAL	$O(n_e d + n_r d^2)$
Structured Embedding	$O(n_e d + 2 n_r d^2)$
Neural Tensor Networks	$O(n_e d + n_r d^3)$
TransE	$O(n_e d + n_r d)$
DISTMULT	$O(n_e d + n_r d)$

Matrix

Tensor

Vector

Entity *Relation*

Knowledge Graph Applications

Entity Recommendations

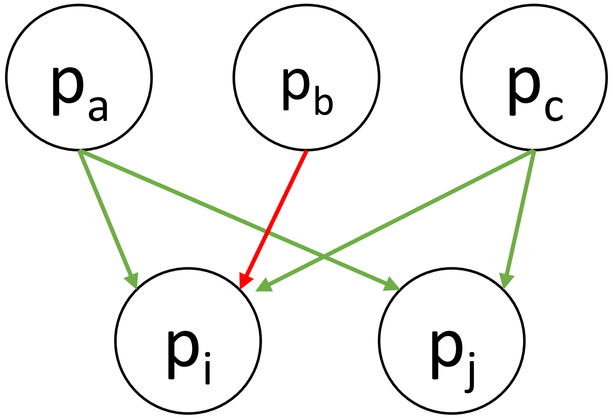
Network based features for paper recommendation

Advantages:

- Uses citation network (co-citation count)
- Similar to human behavior
- High user satisfaction

Disadvantages:

- Low Coverage due to incomplete citation information



$$cc_{i,j} = \sum_{k=1}^{n} c_{k,i} c_{k,j}$$

$c_{x,y} = 1$ denotes paper-x cites paper-y

Kanakia, A., Shen, Z., Eide, D., & Wang, K. (2019). A Scalable Hybrid Research Paper Recommender System for Microsoft Academic. In The World Wide Web Conference on (pp. 2893–2899).

Entity Recommendations

Content based features for paper recommendation

Advantages:

- Uses available paper metadata (titles, abstracts, etc)
- Very high coverage; Embeddings for all English papers
- Highly scalable – leverages topic hierarchy for classification

Disadvantages:

- Low precision compared to co-citation based approach

Kanakia, A., Shen, Z., Eide, D., & Wang, K. (2019). A Scalable Hybrid Research Paper Recommender System for Microsoft Academic. In The World Wide Web Conference on (pp. 2893–2899).

Entity Recommendations

…Combining both approaches

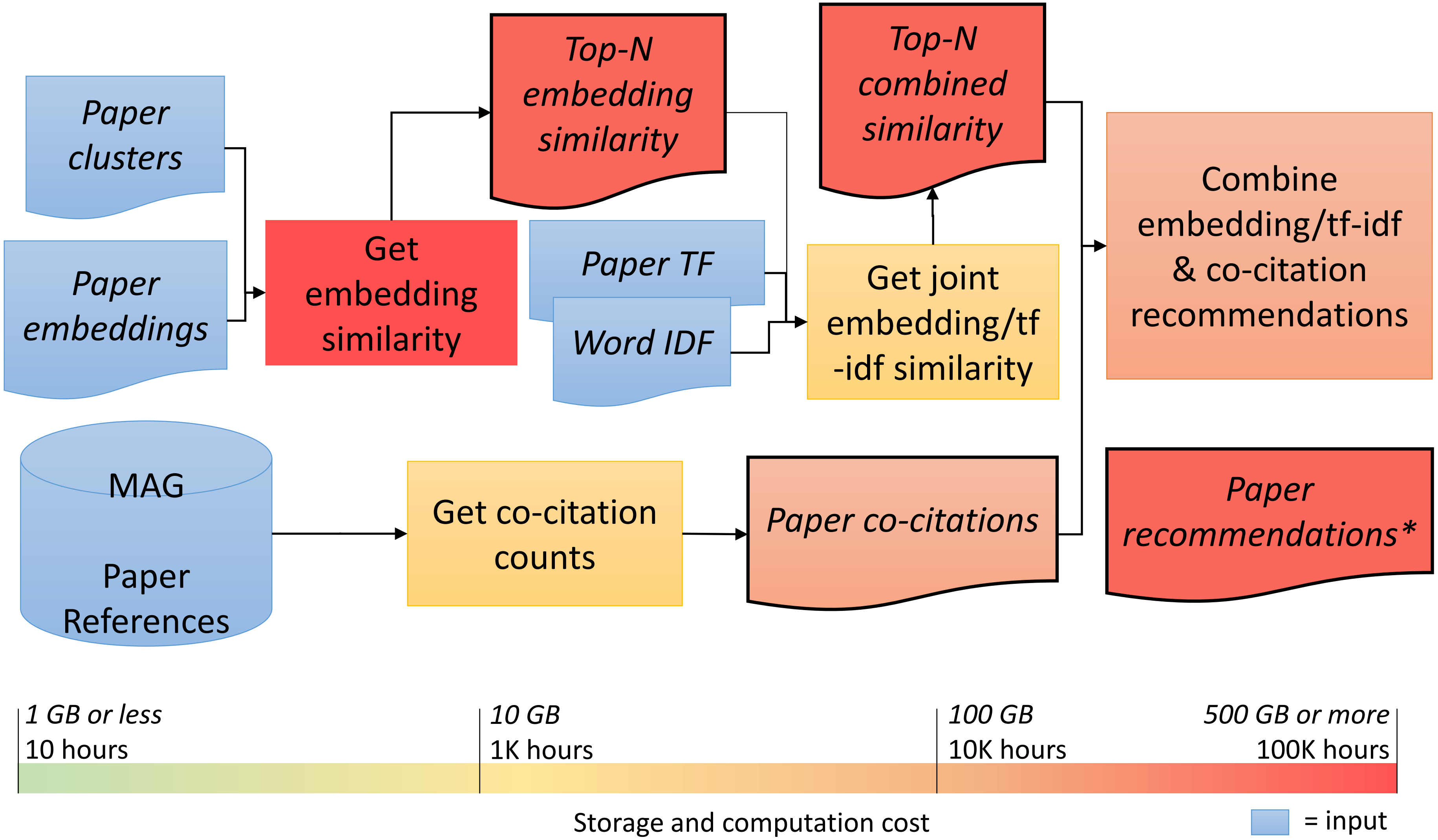
- Need to map co-citation counts to range [0, 1]
- Tunable logistic function

$$\sigma(cc_{i,j}) = \frac{1}{1 + e^{\theta(\tau - cc_{i,j})}}$$

- θ controls slope. Tuned using co-citation mean
- τ controls x-offset. Turned using co-citation variance
- Tuning both parameters can balance novel vs authoritative recommendations

Kanakia, A., Shen, Z., Eide, D., & Wang, K. (2019). A Scalable Hybrid Research Paper Recommender System for Microsoft Academic. In The World Wide Web Conference on (pp. 2893–2899).

Paper Recommendation Pipeline

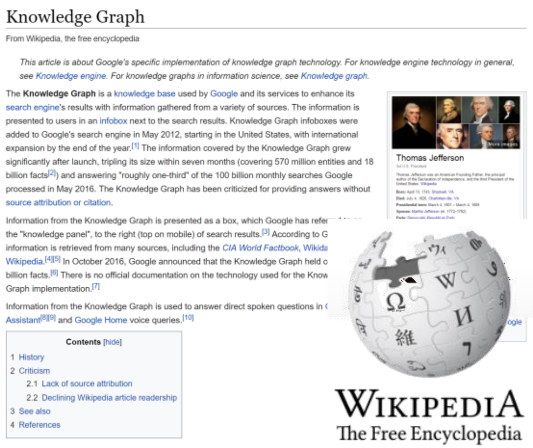


Content Labeling & Language Similarity

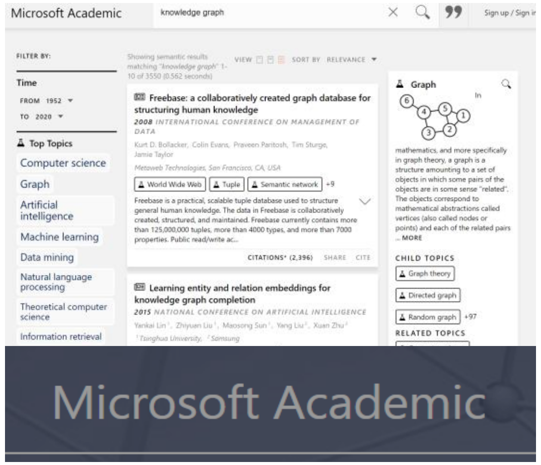
Learning word and topic embeddings from the knowledge graph

- Leverage our topic (field of study) ontology and paper topic labels to learn topic content embeddings
  - External data: Wikipedia content for topics
  - Knowledge graph data: Paper information (titles + abstracts) from papers in topic

- Topic embeddings encode more technical information using paper data
- Topic embeddings can be used to label arbitrary text



+



Lab 5: Knowledge Graph Applications

1. Content Labeling & Language Similarity
2. Network and Content based Paper Recommendation