CMPT 880 Advanced Topic Presentation
February 26,2019

Network Embedding & Heterogeneous Representation Learning

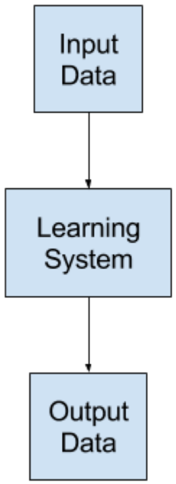
Presented by
Ankita Sakhuja, Kiarash Zahirnia

Outline

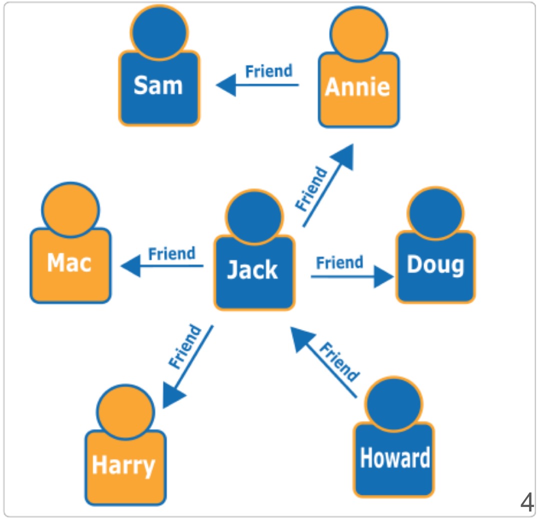
- Graph Embedding, Importance and Application
- Homogeneous Networks and Heterogeneous Networks
- Graph embedding on Heterogeneous Networks
- Recent publications: metapath2vec , metapath2vec++

Conventional Learning

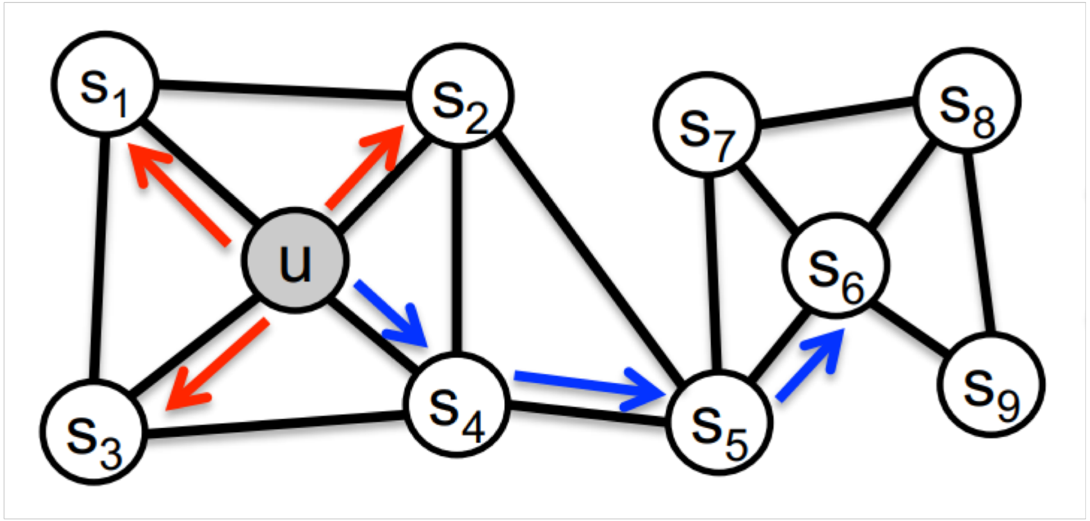
In a traditional machine learning classification setting, we aim to learn a hypothesis H that maps elements of X to the labels set Y.



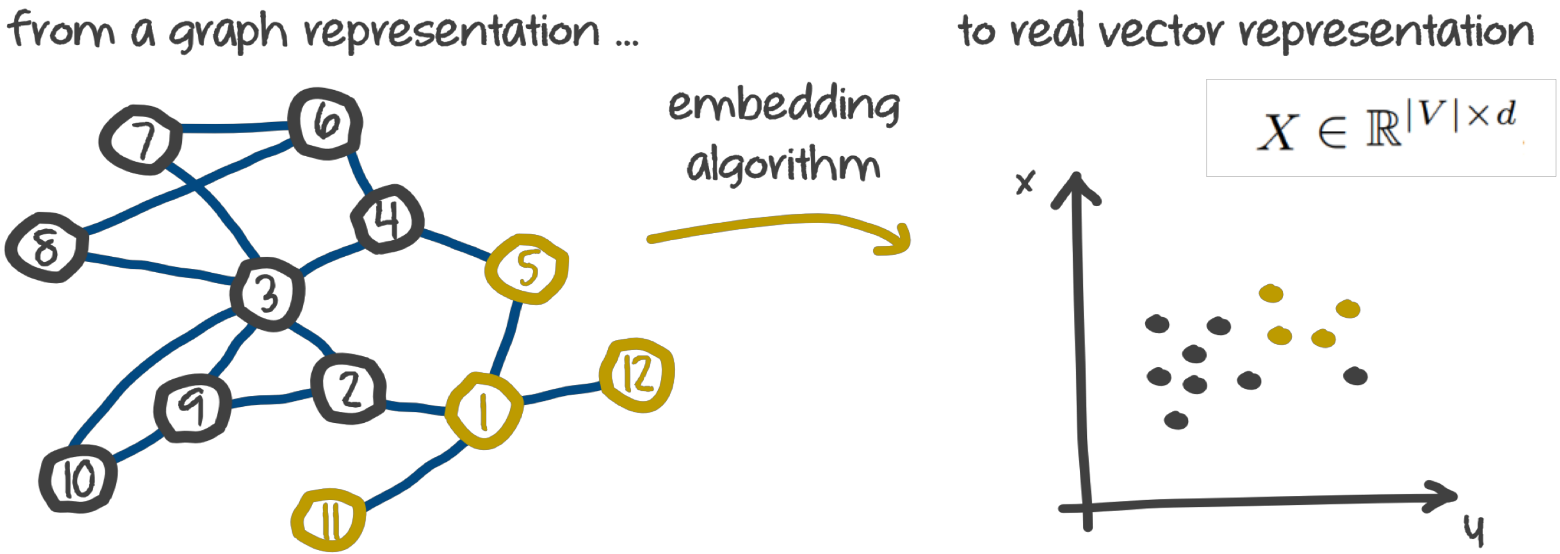
Information about the Dependence of instances!!!



Seeks to Preserve Local Neighborhoods of Nodes.



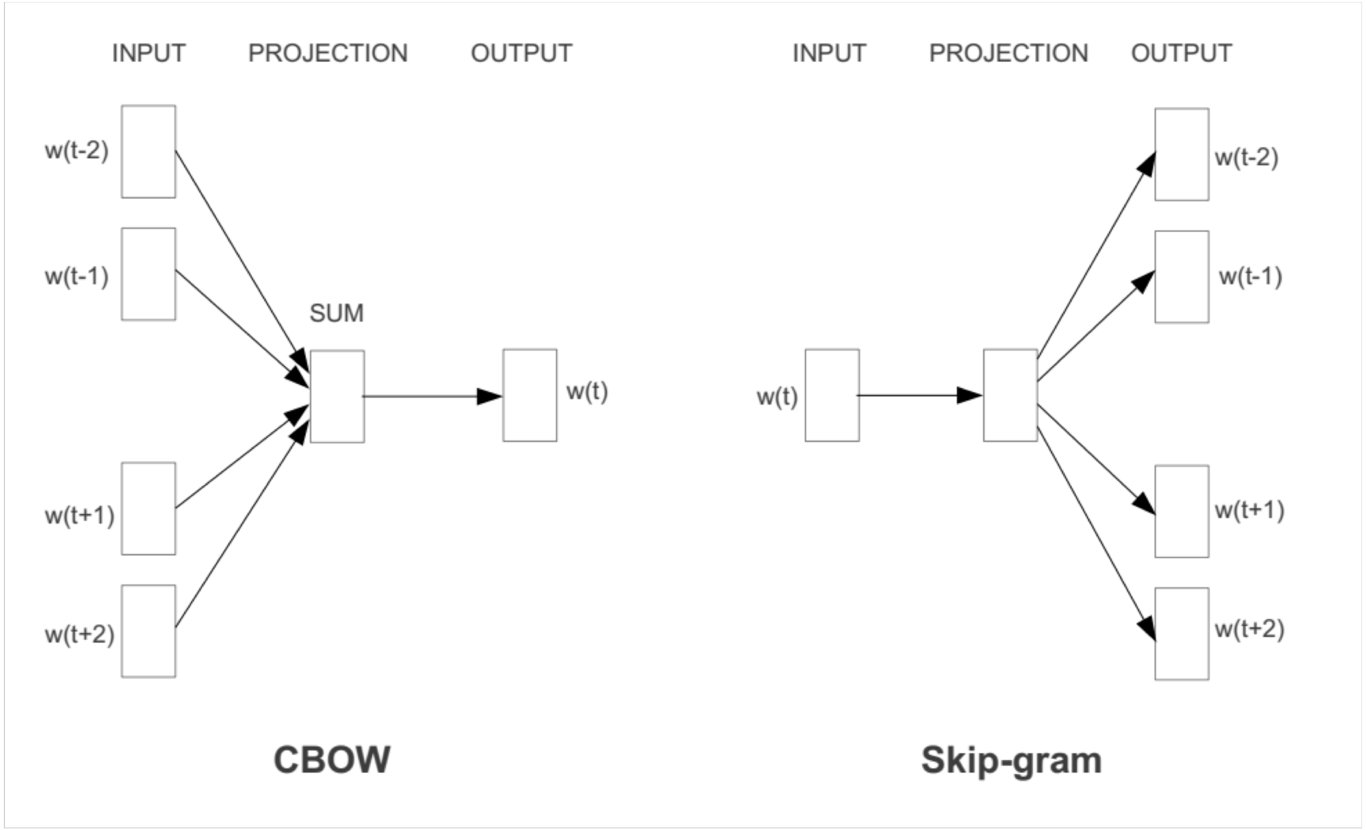
Graph Embedding



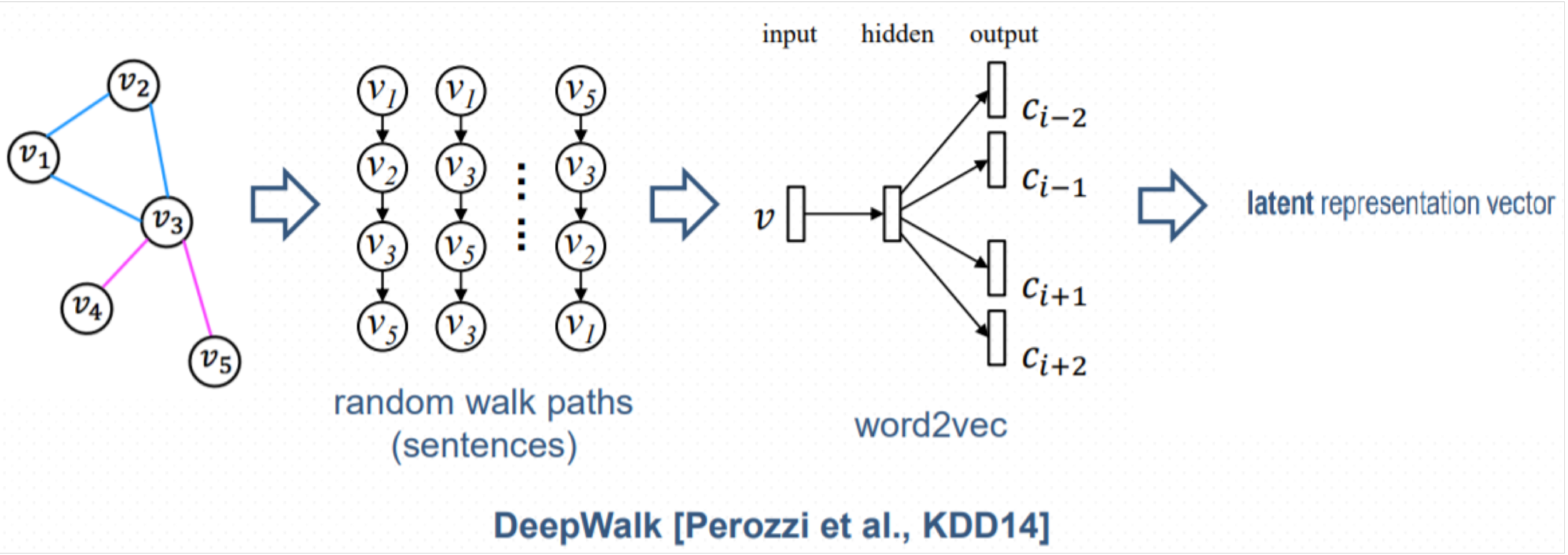
Many problems can be solved by Information Networks -

- Link Prediction
- Community Detection
- Recommendation
- Structure Analysis
- Visualization and so on..

Word Embedding in NLP



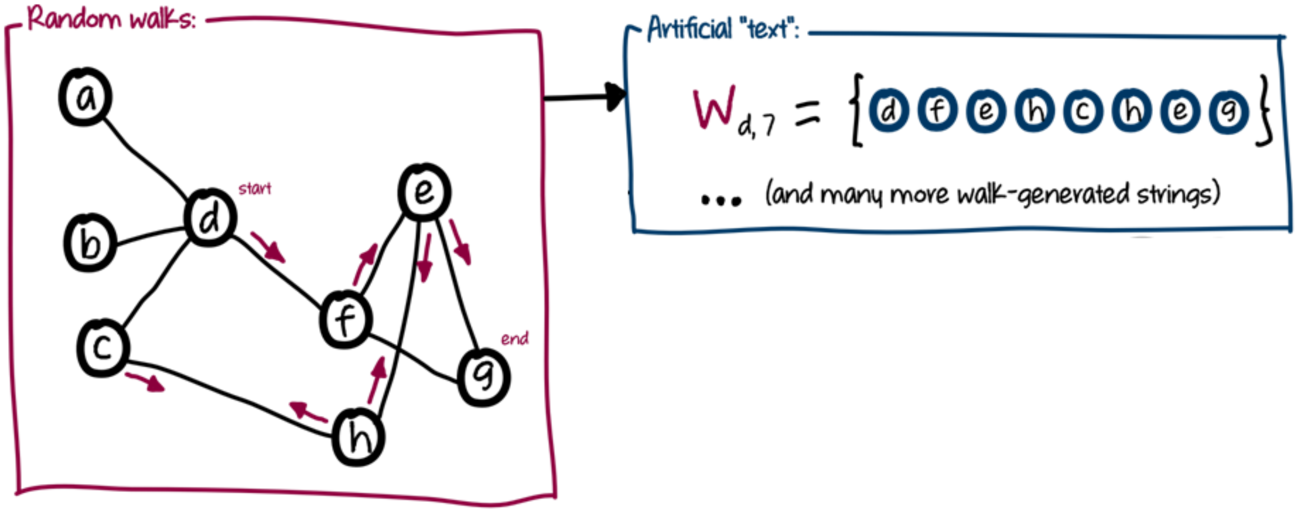
Most Important works on Feature Learning on Homogeneous Networks



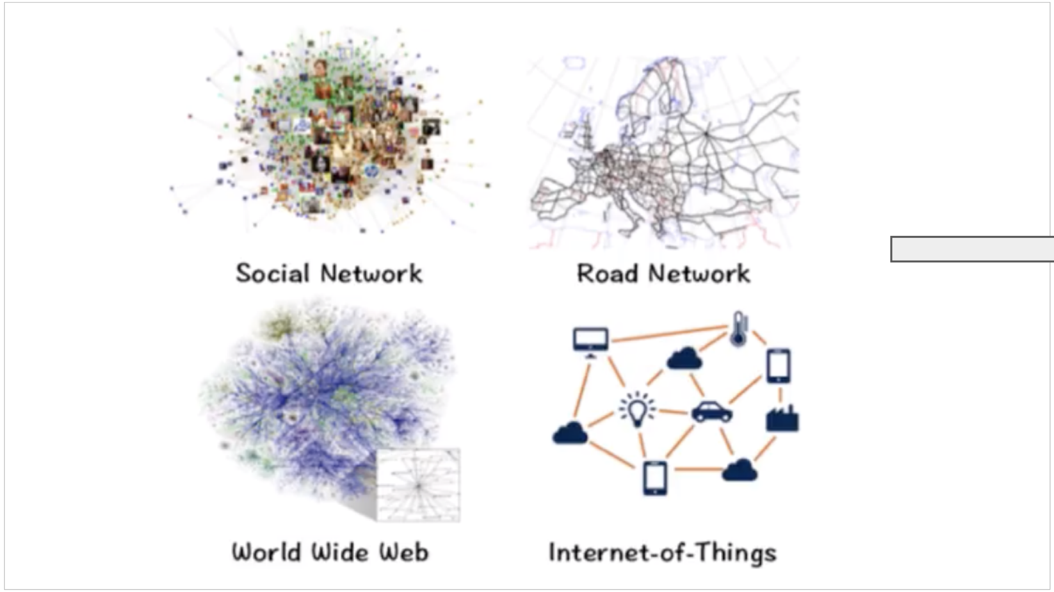
Neighborhood

- DFS
- BFS

- Random Walk



But, how does the real world Network looks like ?



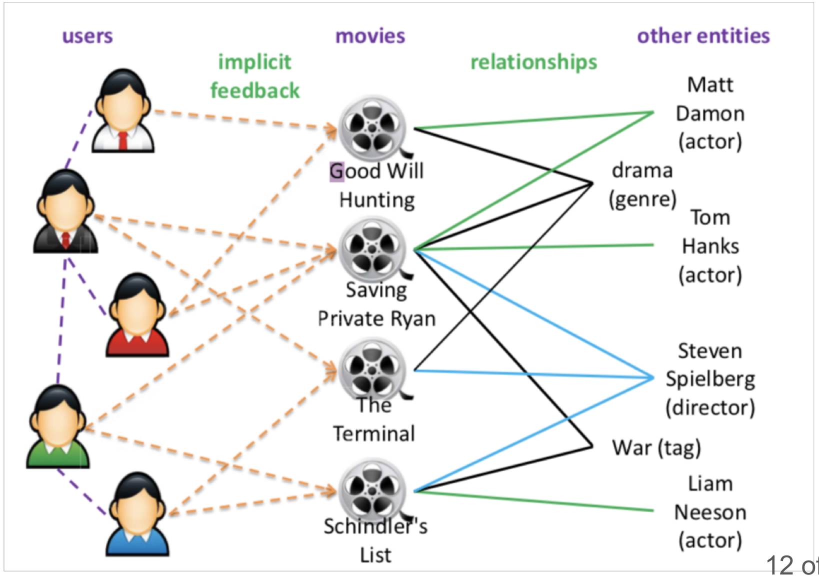
Heterogeneous Network Modelling

Heterogeneous Networks

Heterogeneous Information Networks (HINs) -

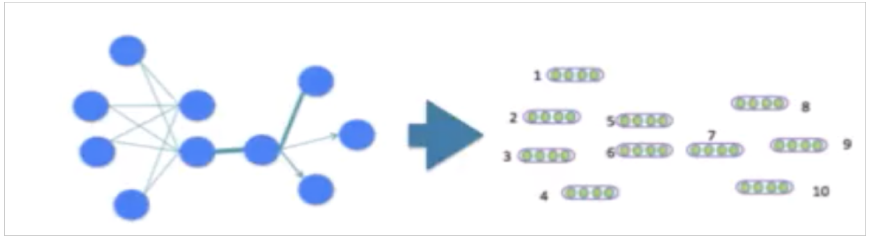
- Multiple Type of Nodes
- Multiple Type of Links

Such as Publication networks, Knowledge base graphs , E-commerce



Heterogeneous Networks Modelling

Every vertex is represented as a low-dimensional vector.



Two approaches discussed in KDD’17 (*metapath2vec: Scalable representation learning for heterogeneous networks." Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2017*)

- metapath2vec
- metapath2vec++

Heterogeneous Network Embedding : Problem

Input: **a heterogeneous information network** $G = (V, E, T)$

Output: $\boldsymbol{X} \in R^{V \times d}, d \ll |V|$, *d*-dim vector $\boldsymbol{X}_v$ for each node *v*.

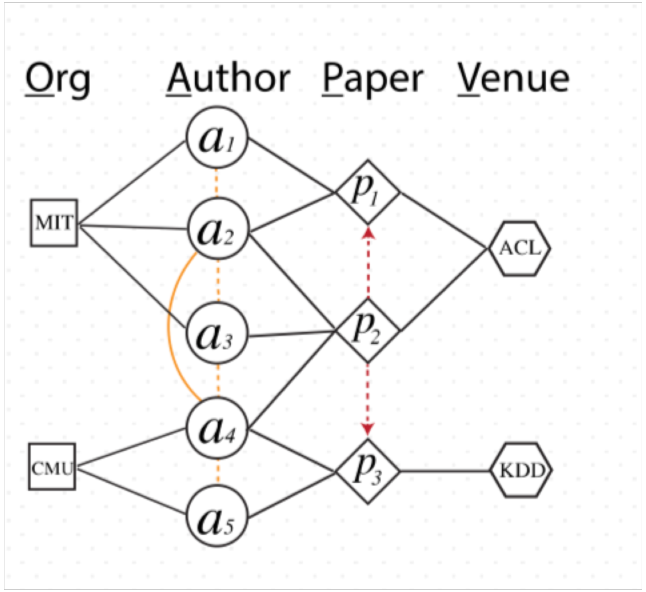
Step 1 : Meta-path-based random walks

Step 2 : Skip-gram

Perform random walks on nodes in a graph to generate node sequences according to meta-path scheme

Run skip-gram to learn the embedding of each node based on the node sequences generated in step 1

Meta-path: Concept is proposed to capture numerous semantic relationships across multiple types of nodes



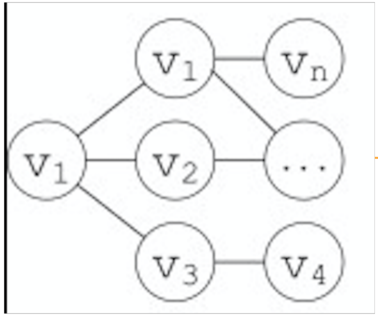
1. Meta-PathGiven a meta-path scheme (Example) ***OAPVPAO***

2. In a traditional random walk procedure, in the toy example, the next step of a walker on node a4 transitioned from node CMU can be all types of nodes surrounding it—**a2, a3, a5, p2, p3, and CMU.**

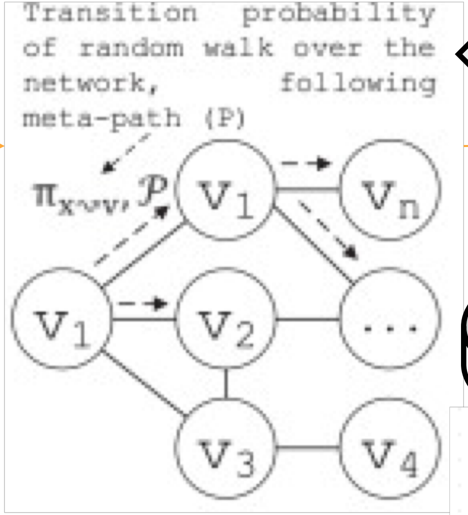
3. Under the meta-path scheme 'OAPVPAO', for example, the walker is biased towards **paper nodes (P)** given its previous step on an organization node CMU (O), following the semantics of this meta-path.

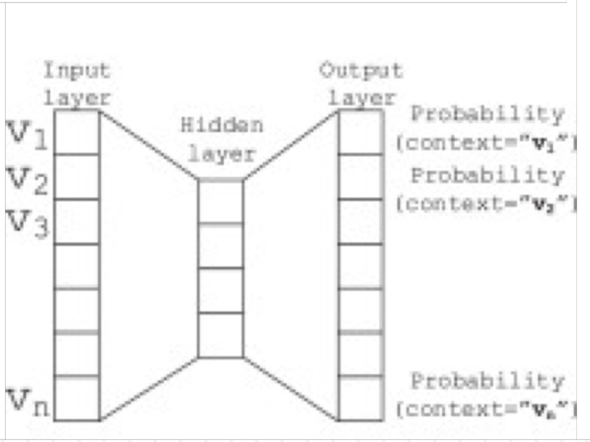
Meta path Based Walk

$$p(v^{i+1}|v_t^i, \mathcal{P}) = \begin{cases} \frac{1}{|N_{t+1}(v_t^i)|} & (v^{i+1}, v_t^i) \in E, \phi(v^{i+1}) = t{+}1 \\ 0 & (v^{i+1}, v_t^i) \in E, \phi(v^{i+1}) \neq t{+}1 \\ 0 & (v^{i+1}, v_t^i) \notin E \end{cases}$$



1. Heterogeneous Network



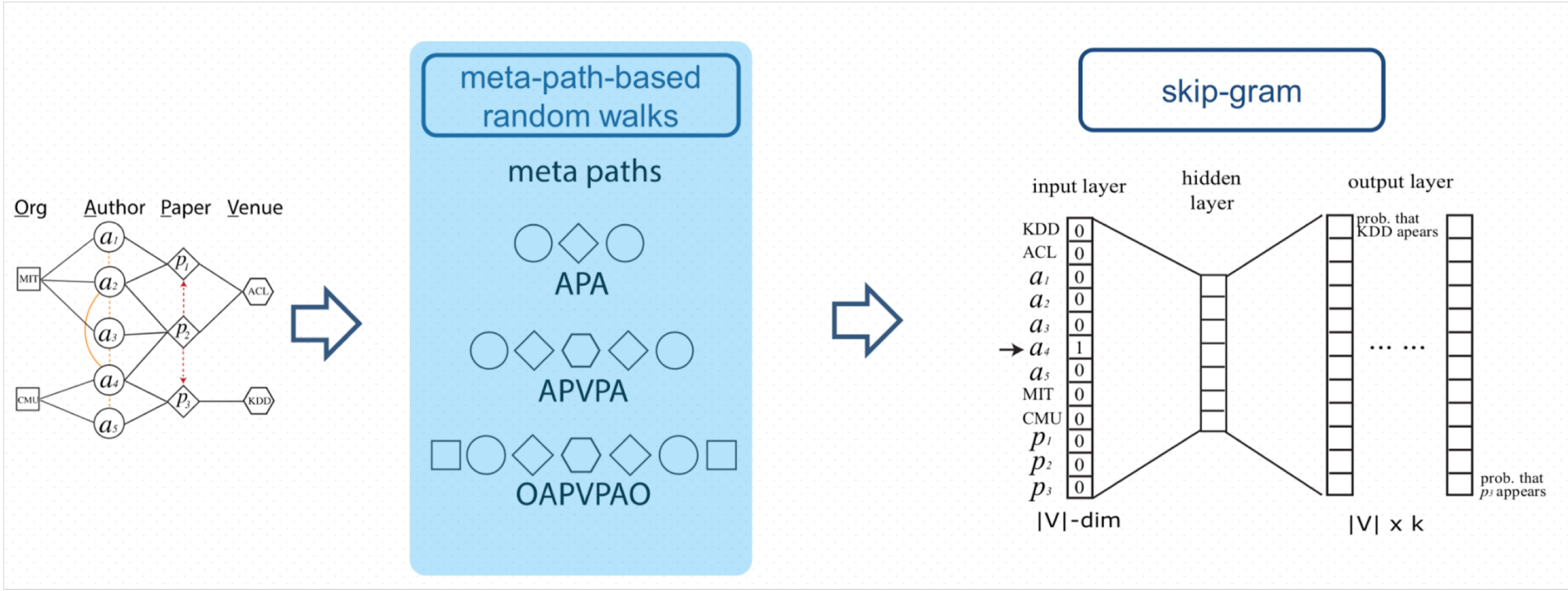


$$p(c_t|v;\theta) = \frac{e^{X_{c_t}} \cdot e^{X_v}}{\sum_{u \in V} e^{X_u} \cdot e^{X_v}}$$

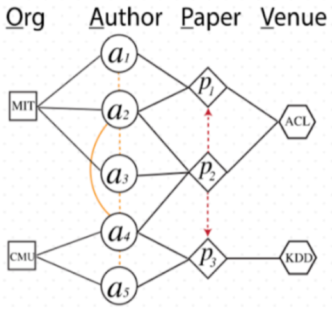
2. Meta-path-based walk

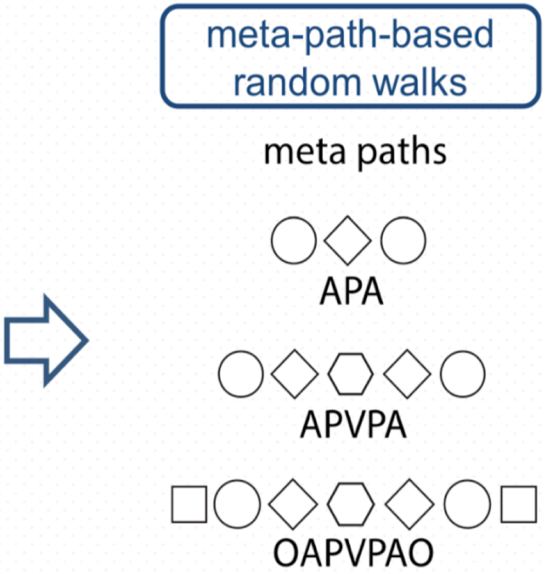
3. Skip-gram architecture

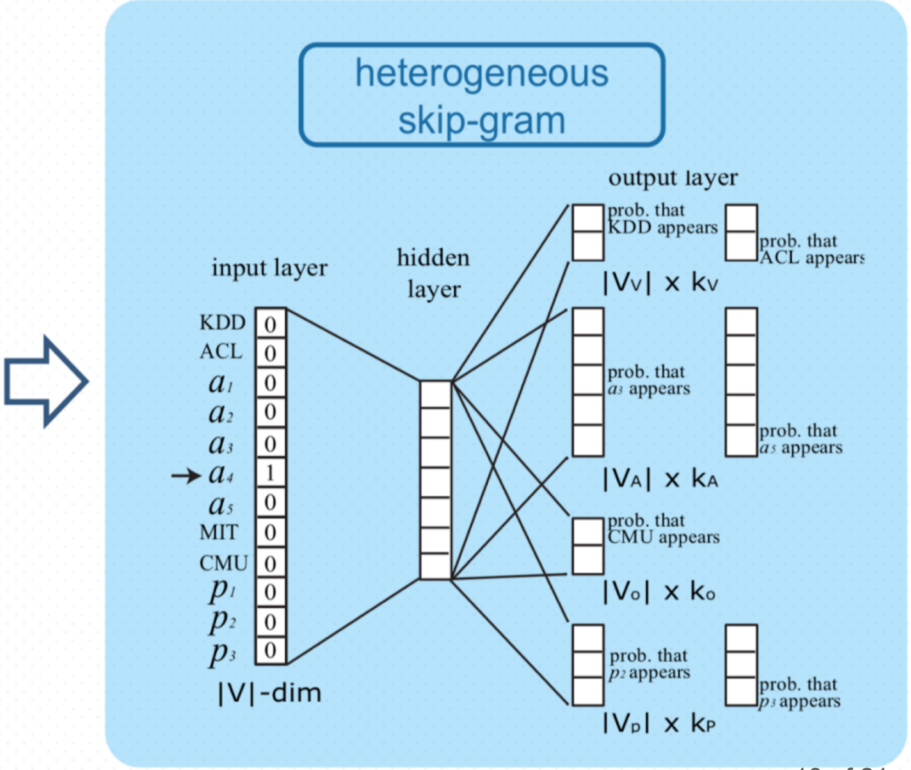
Metapath2vec



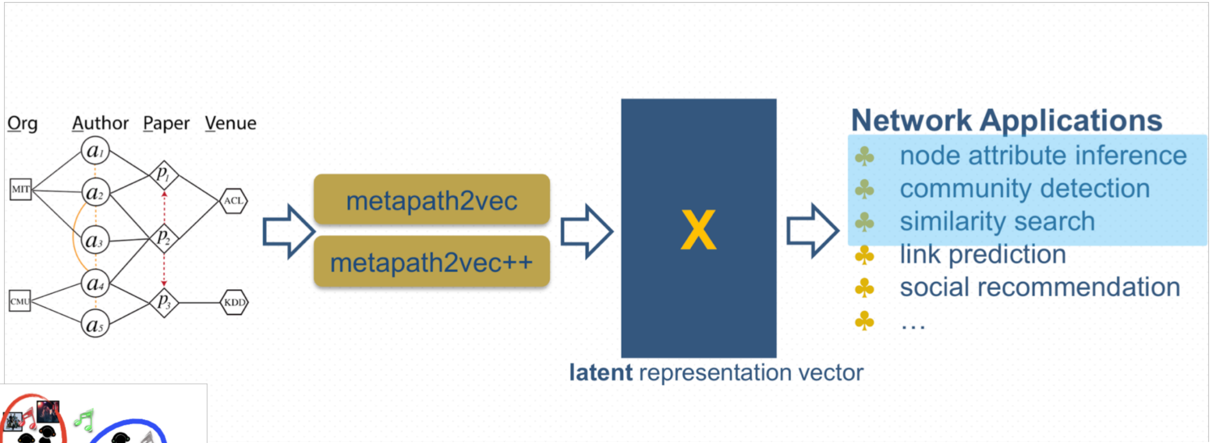
Metapath2vec++

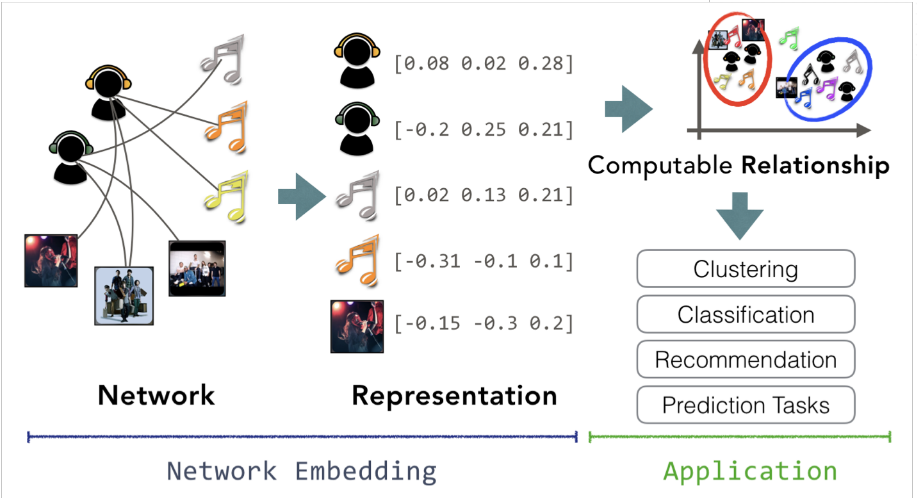






Solution to our "Problem"





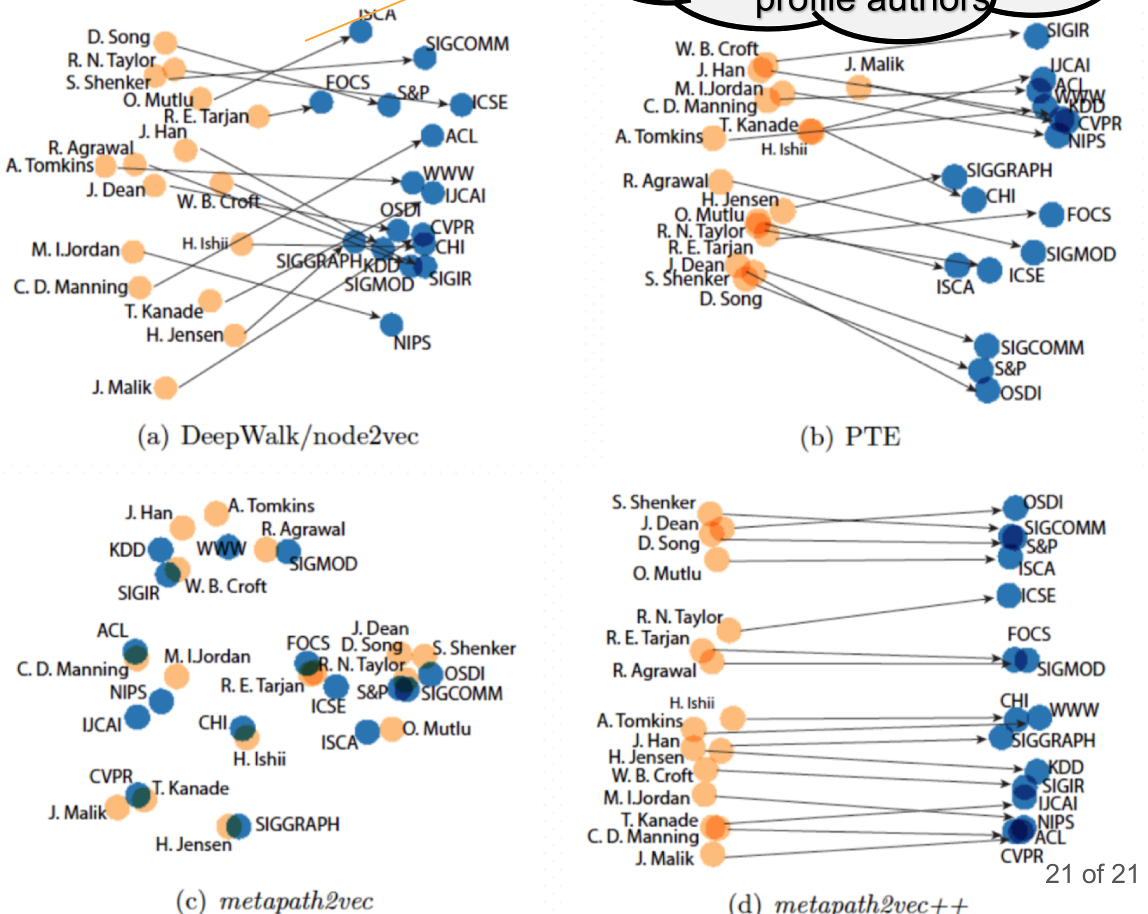
Visualization

Aminer Academic Network
1.7 million authors
3 million papers
3800+ Venues

16 Top CS conferences and corresponding top-profile authors

- Metapath2vec++
  - ➢ Automatically organizes two nodes
  - ➢ Implicitly learn relationships between them indicated by distances and arrows connecting each pair.

- Metapath2vec
  - ➢ Capable of grouping pair of venue and author closely



(a) DeepWalk/node2vec
(b) PTE
(c) *metapath2vec*
(d) *metapath2vec++*

Mathematical Formulation

- ➔ Meta-path-based Random Walks Formulations
  - Transition Probability at Step i

$$p(v^{i+1}|v_t^i, \mathcal{P}) = \begin{cases} \frac{1}{|N_{t+1}(v_t^i)|} & (v^{i+1}, v_t^i) \in E, \phi(v^{i+1}) = t{+}1 \\ 0 & (v^{i+1}, v_t^i) \in E, \phi(v^{i+1}) \neq t{+}1 \\ 0 & (v^{i+1}, v_t^i) \notin E \end{cases}$$

- ➔ Skip-Gram Formulations
  - Softmax in metapath2vec :-

$$p(c_t|v;\theta) = \frac{e^{X_{c_t}} \cdot e^{X_v}}{\sum_{u \in V} e^{X_u} \cdot e^{X_v}}$$

  - Softmax in metapath2vec++ :-

$$p(c_t|v;\theta) = \frac{e^{X_{c_t}} \cdot e^{X_v}}{\sum_{u_t \in V_t} e^{X_{u_t}} \cdot e^{X_v}}$$

References

[1]. Yu, Xiao, et al. "Personalized entity recommendation: A heterogeneous information network approach." *Proceedings of the 7th ACM international conference on Web search and data mining*. ACM, 2014.

[2]. Rong, Xin. "word2vec parameter learning explained." arXiv preprint arXiv:1411.2738 (2014).

[3]. Perozzi, Bryan, Rami Al-Rfou, and Steven Skiena. "Deepwalk: Online learning of social representations." Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2014.

[4]. Grover, Aditya, and Jure Leskovec. "node2vec: Scalable feature learning for networks." Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2016.

[5]. Dong, Yuxiao, Nitesh V. Chawla, and Ananthram Swami. "metapath2vec: Scalable representation learning for heterogeneous networks." Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2017.

[6]. Nickel, Maximilian, et al. "A review of relational machine learning for knowledge graphs." Proceedings of the IEEE 104.1 (2016): 11-33.

Thankyou !