



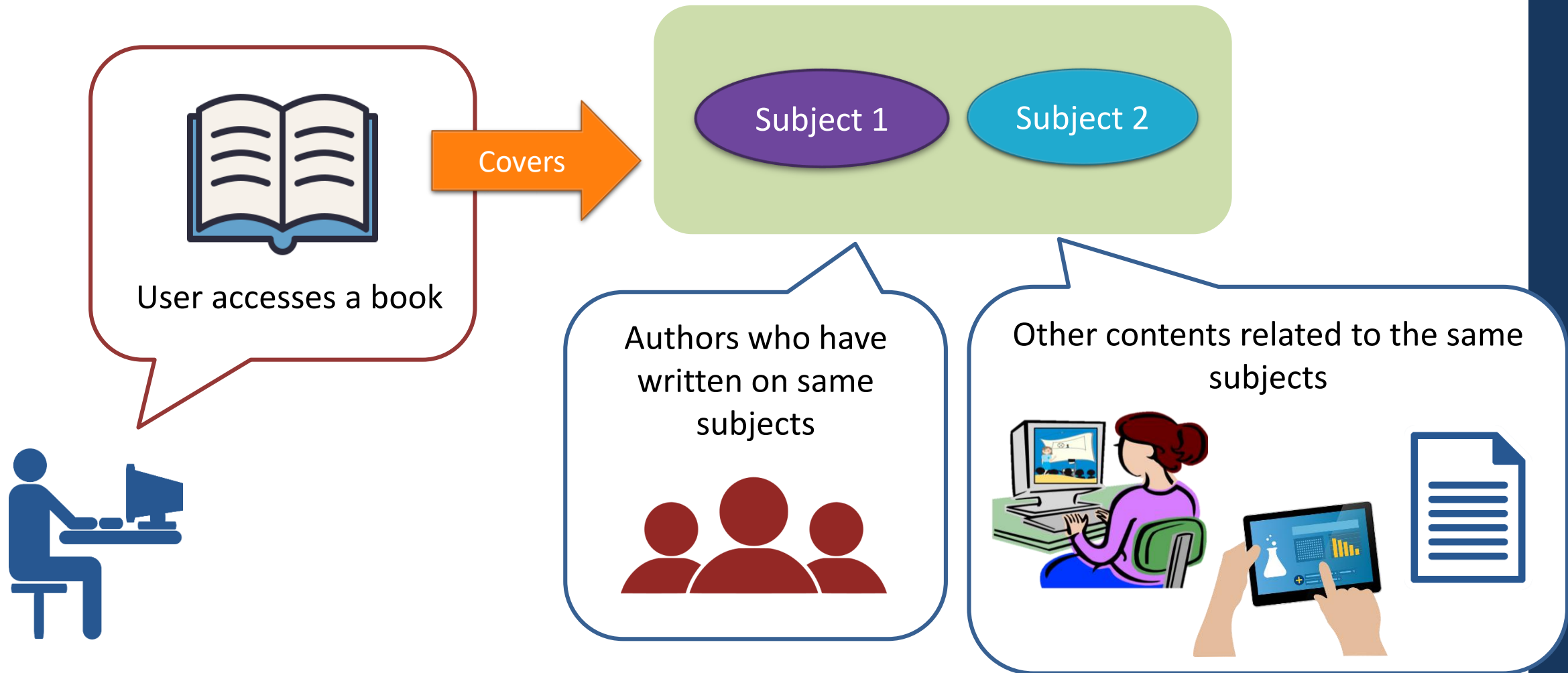
National
Digital Library
of India



Learning to Retrieve Related Resources in a Bibliographic Information Network

Poonam Anthony and Dr. Plaban Kumar Bhowmick

Centre for Educational Technology, IIT Kharagpur



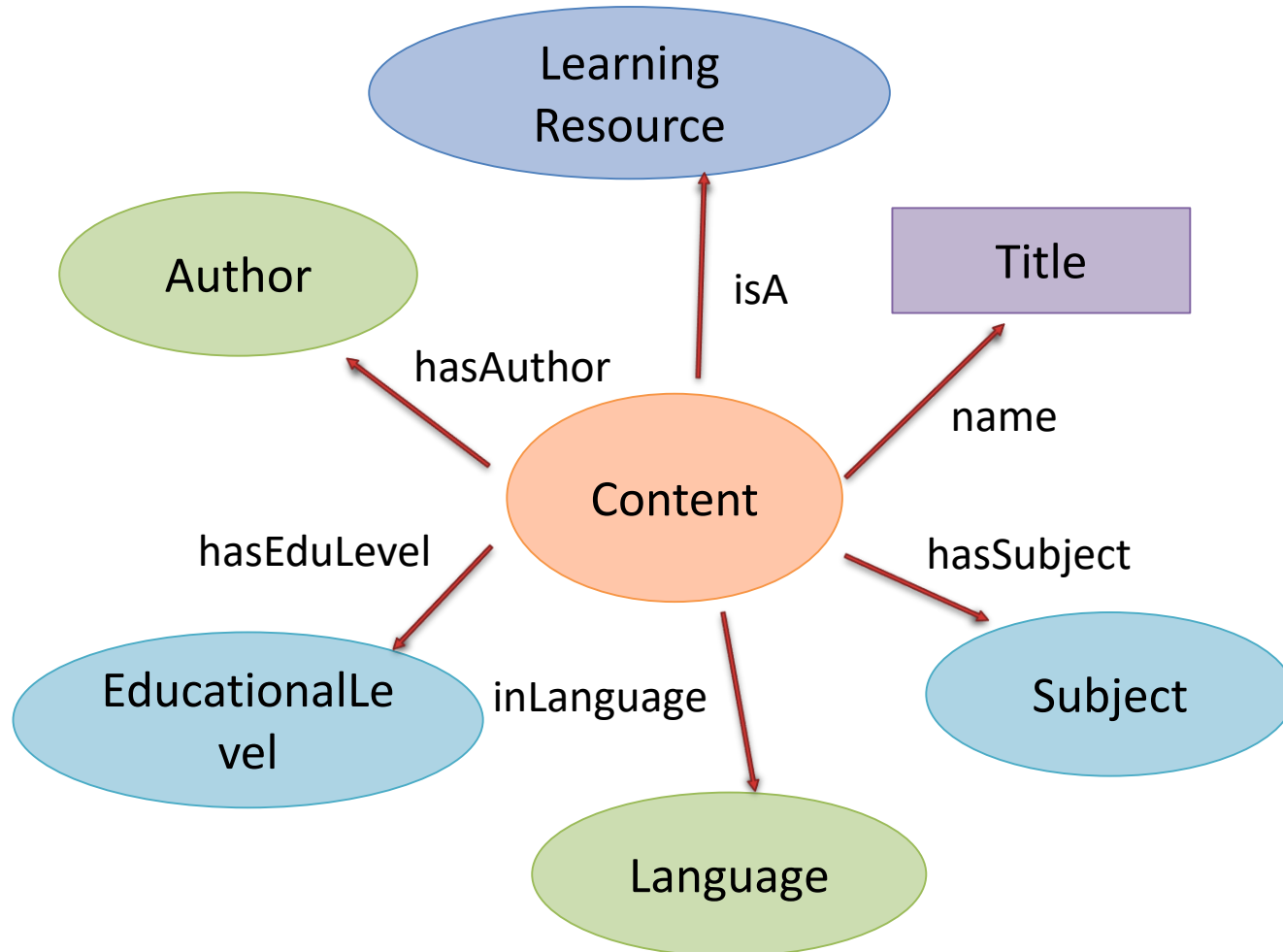
Related entities in a digital library

Image Source: Google

Motivation

- Recommendations are desirable for enhancing the learning experience of users
- Different types of resources can be included in recommendations
- Related resources can be useful for generating recommendations

Graph representation of bibliographic database



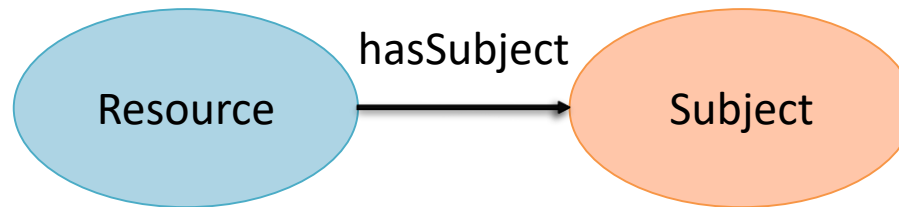
Different types of Nodes and Links

Can be modelled as a Heterogeneous Information Network (HIN)

Related resources
↓
Related entities in a HIN

Relatedness between entities in HIN

- Connected via a single property edge



- Connected via a sequence of edges or property path



Formally known as **Meta-path**

Objectives

- Automatic enumeration of meta-paths
 - Inclusion of HeteSim similarity measure
- Finding related resources on the basis of these meta-paths

Meta-path Generation

Existing Model – Forward Stage-wise Path Generation (FSPG)

- Based on LARS – where meta-paths form the features
- Uses forward search to find a meta-path at each regression stage
- The search algorithm is based on Path Constrained Random Walk

Proposed Model – Bi-directional Meta-path Search (BMS)

- Uses the same LARS model
- Incorporates HeteSim as the underlying similarity measure
- Uses a two-way search strategy to accommodate HeteSim

Bi-directional Meta-path Search

Input

- A set of example pairs / training pairs
- Residual vector $r = \{1, \dots, -1\}$
- The HIN G

Output

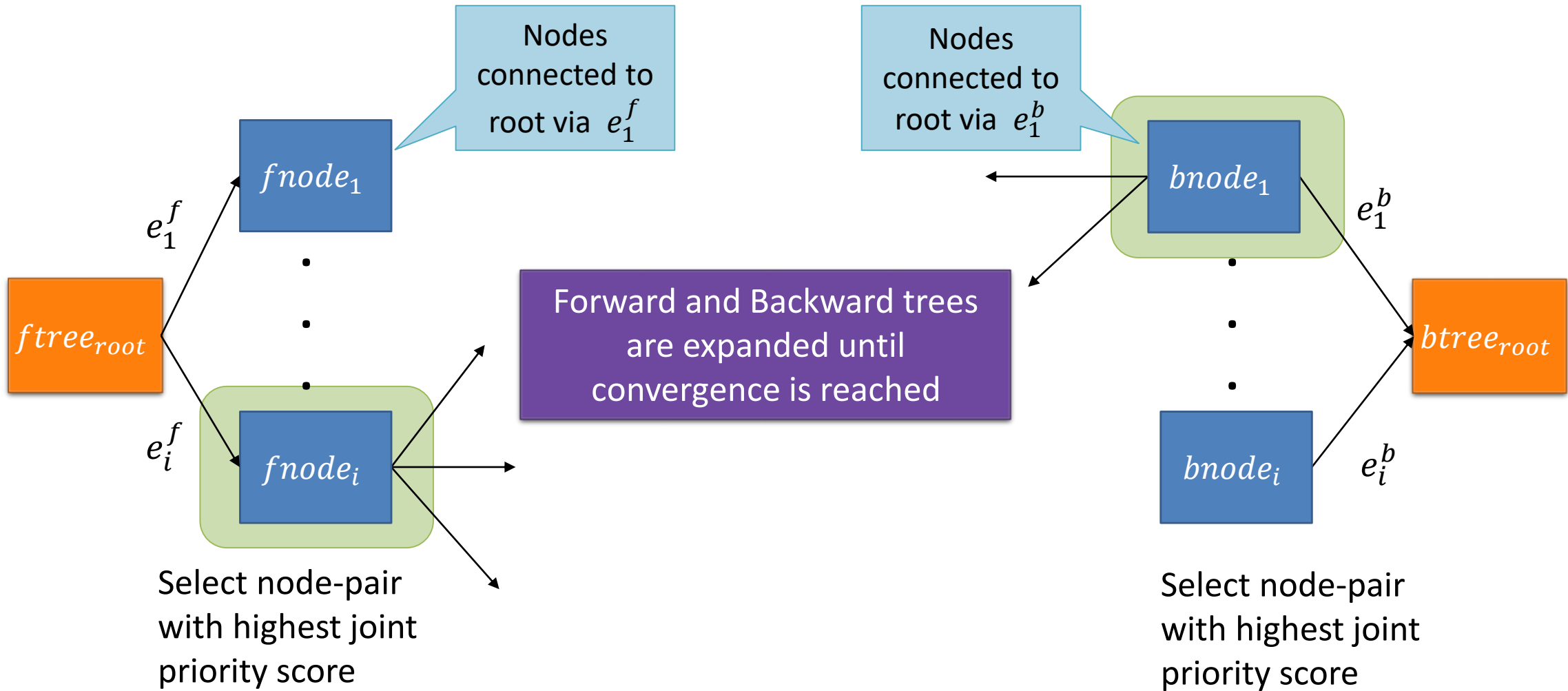
- Set of meta-paths
- Set of weights

Initialization of forward and backward search tree nodes

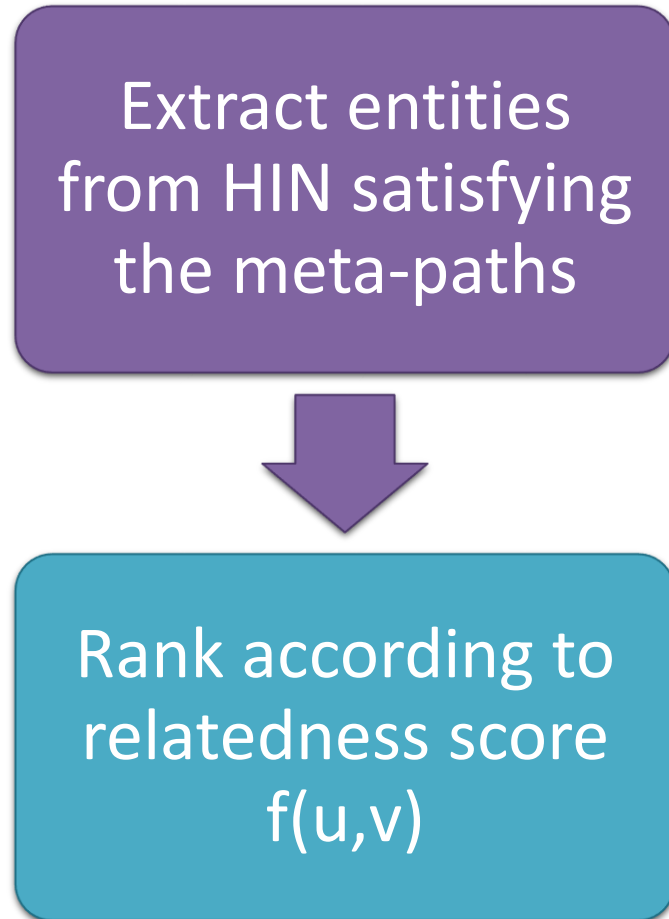
Expansion of forward and backward trees

Termination – convergence of forward and backward trees

Expansion of Forward and Backward Trees



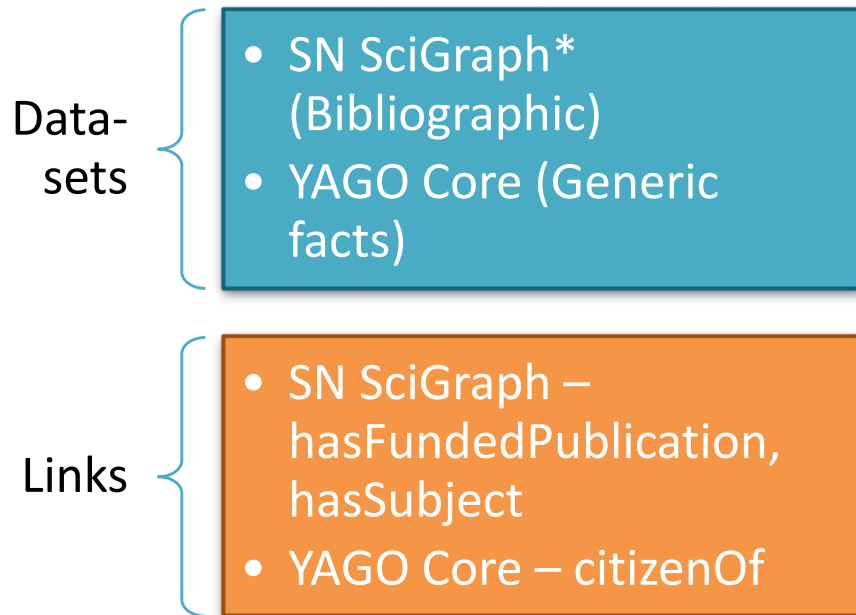
Finding Related Resources



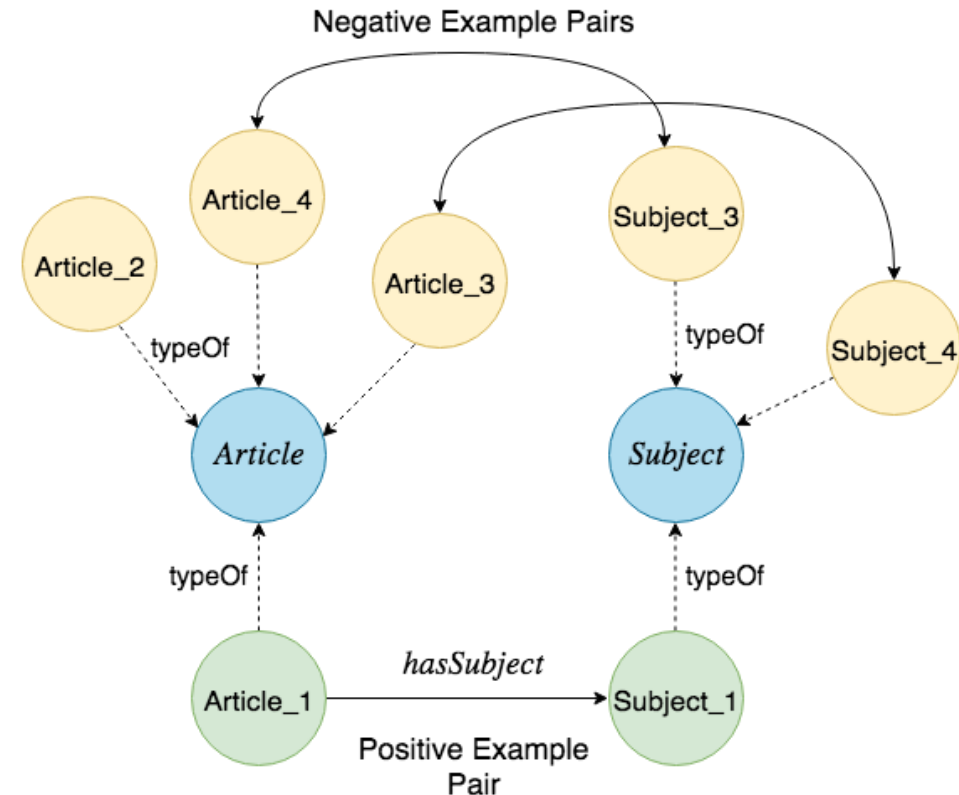
Relatedness score is computed as:

$$f(u, v) = \sum_{\Pi_i} HeteSim(u, v | \Pi_i) \times weight_i$$

Experiment: Comparison of BMS with Baseline model (FSPG)



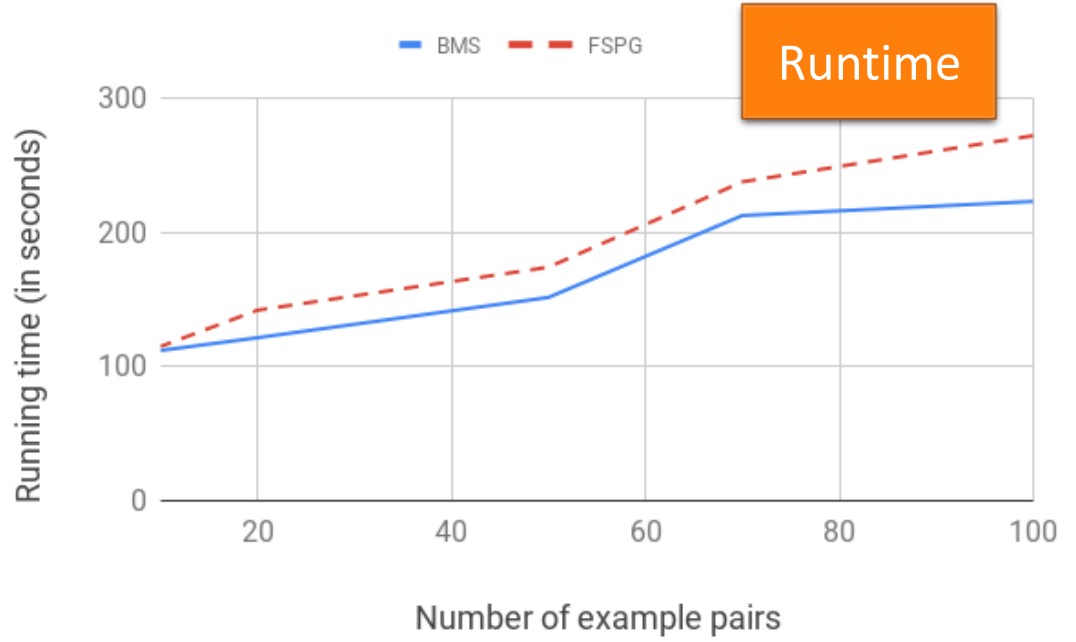
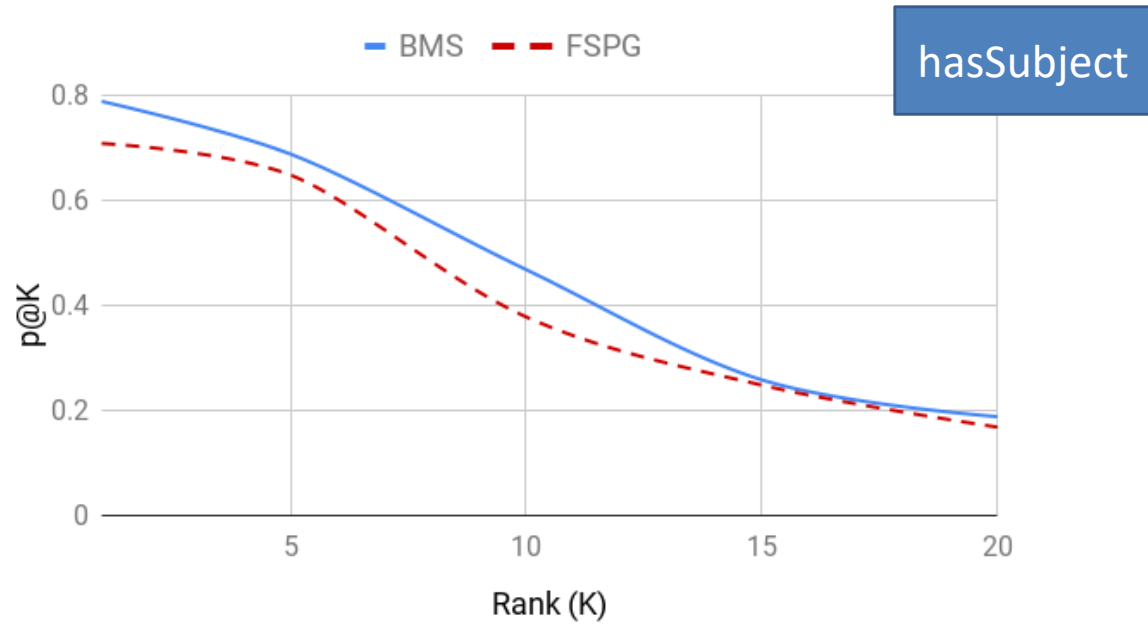
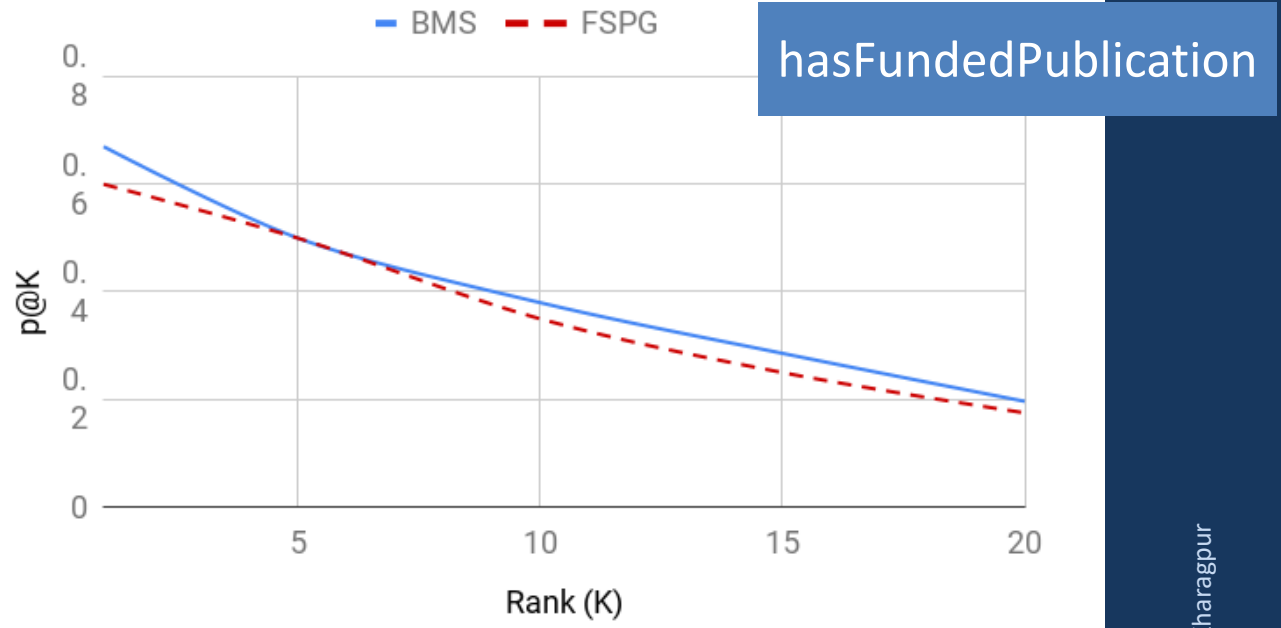
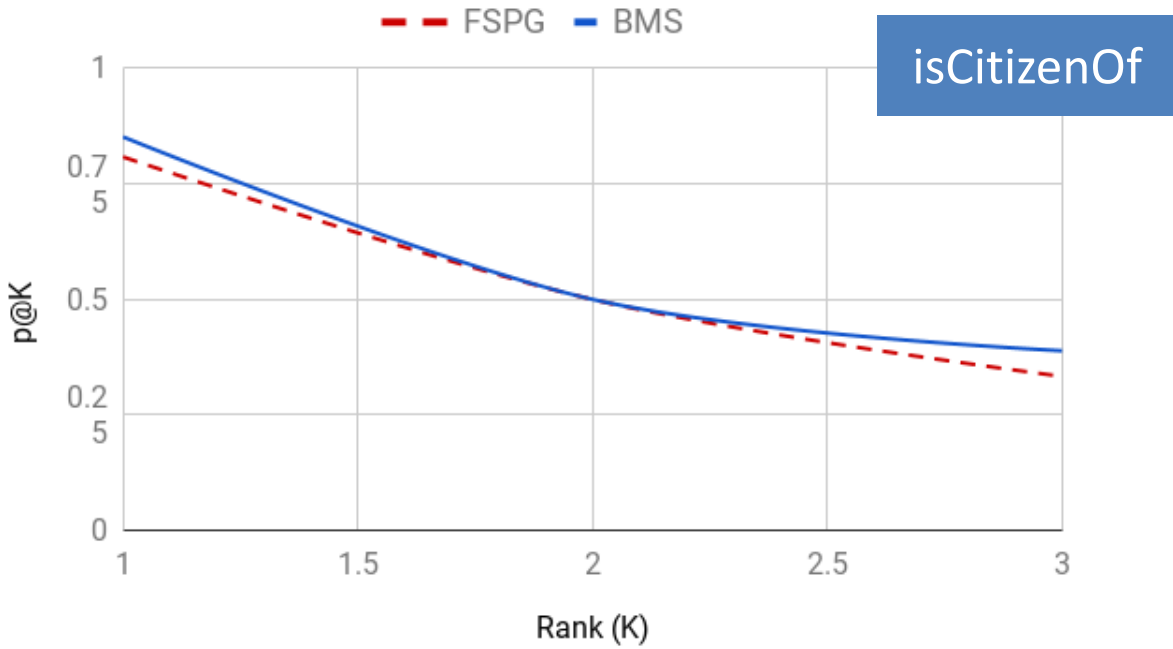
Training data	100 example pairs
Test data	50 example pairs



*<https://www.springernature.com/gp/researchers/scigraph>

Meta-paths discovered

Link type	Meta-paths	Weight (FSPG)	Weight (BMS)
isCitizenOf	<bornIn,locatedIn>	0.53424	0.55023
	<livesIn>	0.31825	0.49674
	<graduatedFrom,locatedIn>	0.18411	0.08414
	<diedIn,locatedIn>	0.01580	0.13666
hasFundedPublication	<hasRecipientOrg,^hasContributingOrg>	1.26714	2.30647
	<hasRecipientOrg,isOrganizationOf,isAffiliationOf,isContributionOf>	0.54694	0.86630
hasSubject	<hasCategory>	0.36198	0.12128
	<hasCategory,broader,narrower>	0.02334	0.09416



Conclusion and future scope

- BMS shows an improvement over FSPG in performance
- Precision is higher in higher ranks – suited for recommendation
- Augmentation of link only meta-paths with node classes
- Considering the context of nodes to reduce false positives

References

- [1] Y. Sun, J. Han, X. Yan, P. S. Yu, and T. Wu, "Pathsim: Meta path-based top-k similarity search in heterogeneous information networks," Proceedings of the VLDB Endowment, vol. 4, no. 11, pp. 992-1003, 2011.
- [2] C. Shi, X. Kong, Y. Huang, S. Y. Philip, and B. Wu, "Hetesim: A general framework for relevance measure in heterogeneous networks," IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 10, pp. 2479-2492, 2014.
- [3] Fabian M Suchanek, Gjergji Kasneci, and Gerhard Weikum. 2007. Yago: a core of semantic knowledge. In Proceedings of the 16th international conference on WorldWide Web. ACM, 697–706
- [4] C. Meng, R. Cheng, S. Maniu, P. Senellart, and W. Zhang, "Discovering meta-paths in large heterogeneous information networks," in Proceedings of the 24th International Conference on World Wide Web, pp. 754-764, International World Wide Web Conferences Steering Committee, 2015.

Acknowledgement

- JCDL Organizers
- Our Sponsors –
 - **National Digital Library of India** and **IBM Research**

Thank You