Attributed and Predictive Entity Embedding for Fine-Grained Entity Typing in Knowledge Bases

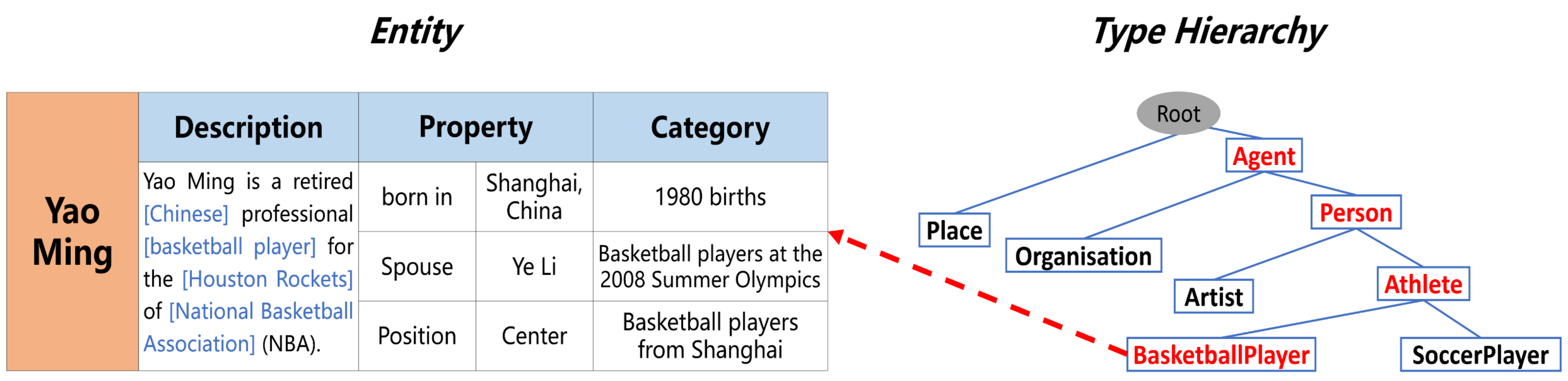
Hailong Jin, Lei Hou and Juanzi Li
{jinhl15@mails.,houlei@mail.,lijuanzi@}tsinghua.edu.cn

Introduction

- In the real world, some large $\mathcal{KB}$s are lack of type information. Fine-grained entity typing aims at identifying semantic types of an entity in $\mathcal{KB}$. Existing methods suffer from two main problems:
  - Ignoring rich structural and partial-labeled information in $\mathcal{KB}$.
  - Requiring large scale corpus in which entity mentions are annotated.
- We propose **APE** model, which can fully utilize various kinds of information comprehensively. Our work benefits many real applications, such as relation extraction, entity linking and QA system.

Approach

An Example in $\mathcal{KB}$

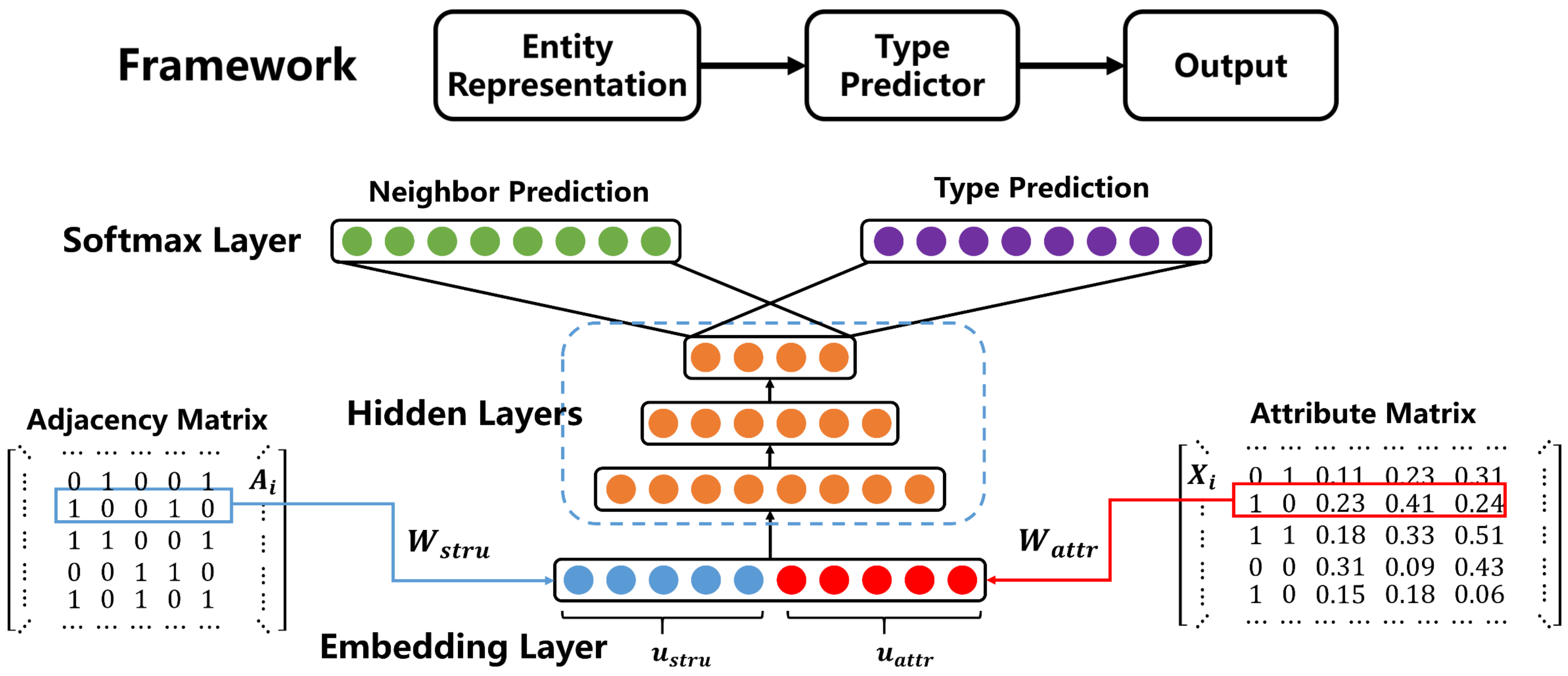


Various Kinds of Information

Three kinds of entity knowledge:

- ***Entity Link Relation***: convert entities into a network structure, which reflects the overall structure of $\mathcal{KB}$.
- ***Entity Attribute***: textual description, properties and categories.
- ***Entity Type***: part of entities in $\mathcal{KB}$ have been assigned with types, which are useful signals to infer missing type.

APE Model



Entity Representation

Attributed Embedding

Information Fusion: We present a tower structure neural network model to integrate link structure and entity attributes.

Predictive Embedding

Type Prediction: For a labeled entity, we jointly predict its type and its neighbor entities in the entity network during training.

Type Hierarchy Modeling

Type Order: We define two kinds of type order to model the structure of taxonomy via *Learning to Rank* method.

A general state-of-the-art architecture consists of two components:

- **Entity Representation:** It learns distributed representation $\vec{v}(e)$ for each entity in $\mathcal{KB}$ (Attributed and Predictive embedding).
- **Type Predictor**: It learns a predict function P. $P(t|e)$ denotes the probability that entity e belongs to type t.

Experiments

Datasets

We construct a dataset from DBpedia. For each entity, we use the single type-path assigned in DBpedia as the ground truth. Statistics of the dataset as fellow:

Type	Entity	Link	Attribute		
			word	property	category
214	300,000	5,243,230	10,000	450	550

Metrics

We use Strict-F1, Mi-F1 and Ma-F1 to evaluate the performance.

Comparison Results

We compare APE with 8 state-of-the-art entity typing methods.

Type	Algorithm	Acc	Ma-F1	Mi-F1
Entity Typing	**FIGMENT**	0.463	0.619	0.628
	CUTE	0.515	0.673	0.677
	MuLR	0.501	0.654	0.662
	Corpus	0.488	0.662	0.659
	Global	0.457	0.608	0.615
Network Embedding	**Planetoid**	0.434	0.585	0.590
	ASNE	0.417	0.568	0.573
	PTE	0.352	0.512	0.518
Our Variants	**APE**$_{no\_type}$	0.492	0.649	0.657
	APE$_{no\_hierarchy}$	0.531	0.694	0.689
	APE	**0.545**	**0.702**	**0.711**

References

1. Liao et al. Attributed Social Network Embedding. TKDE 2018.
2. Yaghoobzadeh et al. Corpus-Level Fine-Grained Entity Typing. JAIR 2017.
3. Yang et al. Revisiting semi-supervised learning with graph embeddings. ICML 2016.
4. Ren et al. AFET: Automatic Fine-Grained Entity Typing by Hierarchical Partial-Label Embedding. EMNLP 2016.

A Joint Research Collaboration between
NUS National University of Singapore
Tsinghua University
University of Southampton

Our research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its IRC@SG Funding Initiative.