Knowledge Graph Completion with Mixed Geometry Tensor Factorization

Yusupov Viacheslav[1] Maxim Rakhuba[1] Evgeny Frolov[2,1]

1. HSE University, 2. AIRI

v.yusupov.lab@gmail.com

Motivation

We observe that real-world knowledge graphs may not consistently align with the assumption of strict hierarchical internal structure and may only partially follow a power-law distribution.

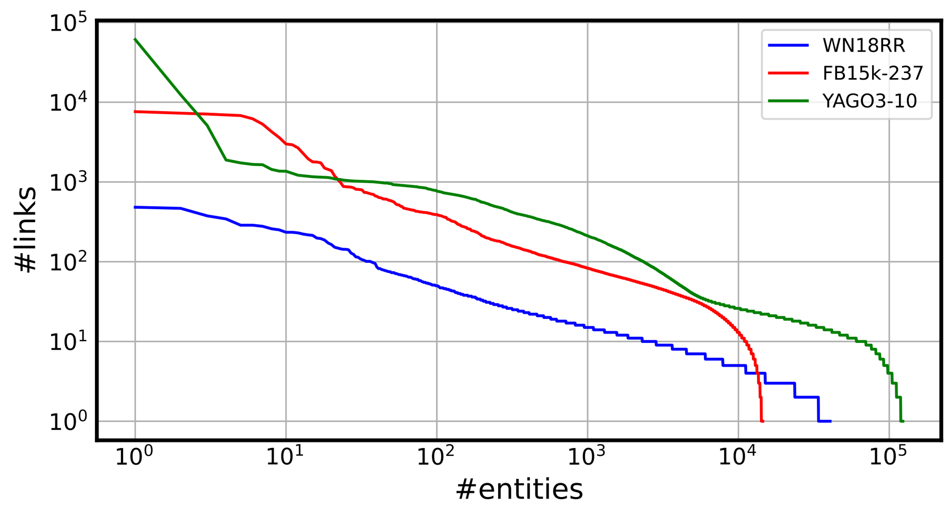

Figure 1: Links distribution on three benchmark knowledge graphs considered in this work. Our MIG-TF approach outperforms both Euclidean and hyperbolic models, see Table 1.

Hyperbolic Geometry

The Lorentz inner product for vectors $x, y \in \mathbb{R}^{n+1}$:

$$\langle x, y\rangle_{\mathcal{L}} = -x_0 y_0 + \sum_{i=1}^{n} x_i y_i$$

$$x_0 = \sqrt{\beta + \sum_{i=1}^{n} x_i^2},$$

$$\|x\|_{\mathcal{L}}^2 = \langle x, x\rangle_{\mathcal{L}}$$

The corresponding n-dimensional Hyperboloid $\mathcal{H}^{n,\beta} \subset \mathbb{R}^{n+1}$ is defined as follows:

$$\mathcal{H}^{n,\beta} = \left\{ x \in \mathbb{R}^{n+1} \,\middle|\, \|x\|_{\mathcal{L}}^2 = -\beta,\ \beta \geq 0 \right\}.$$

The origin vector of the hyperboloid $\mathcal{H}^{n,\beta}$ equals to $0 = (\beta, 0, ..., 0) \in \mathbb{R}^{n+1}$. The inner product of 0 and x is, hence, $\langle 0, x\rangle_{\mathcal{L}} = -\beta x_0$.

Geodesic and Lorentz Distance

The associated geodesic distance is defined as

$$d_l(x, y) = \operatorname{arccosh}(-\langle x, y\rangle_{\mathcal{L}}).$$

We introduce the square Lorentz distance between $x, y \in \mathcal{H}^n$:

$$d_{\mathcal{L}}^2(x, y) = \|x - y\|_{\mathcal{L}}^2 = -2 - 2\langle x, y\rangle_{\mathcal{L}}. \quad (1)$$

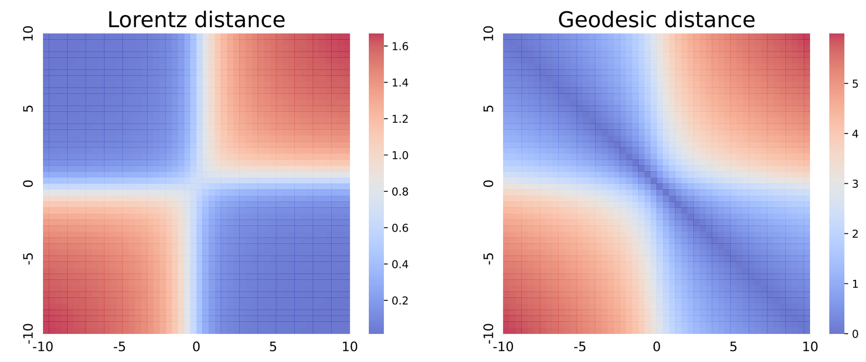

Figure 2: Left score function landscape corresponds to our Lorentz distance and right to Geodesic one.

Contributions

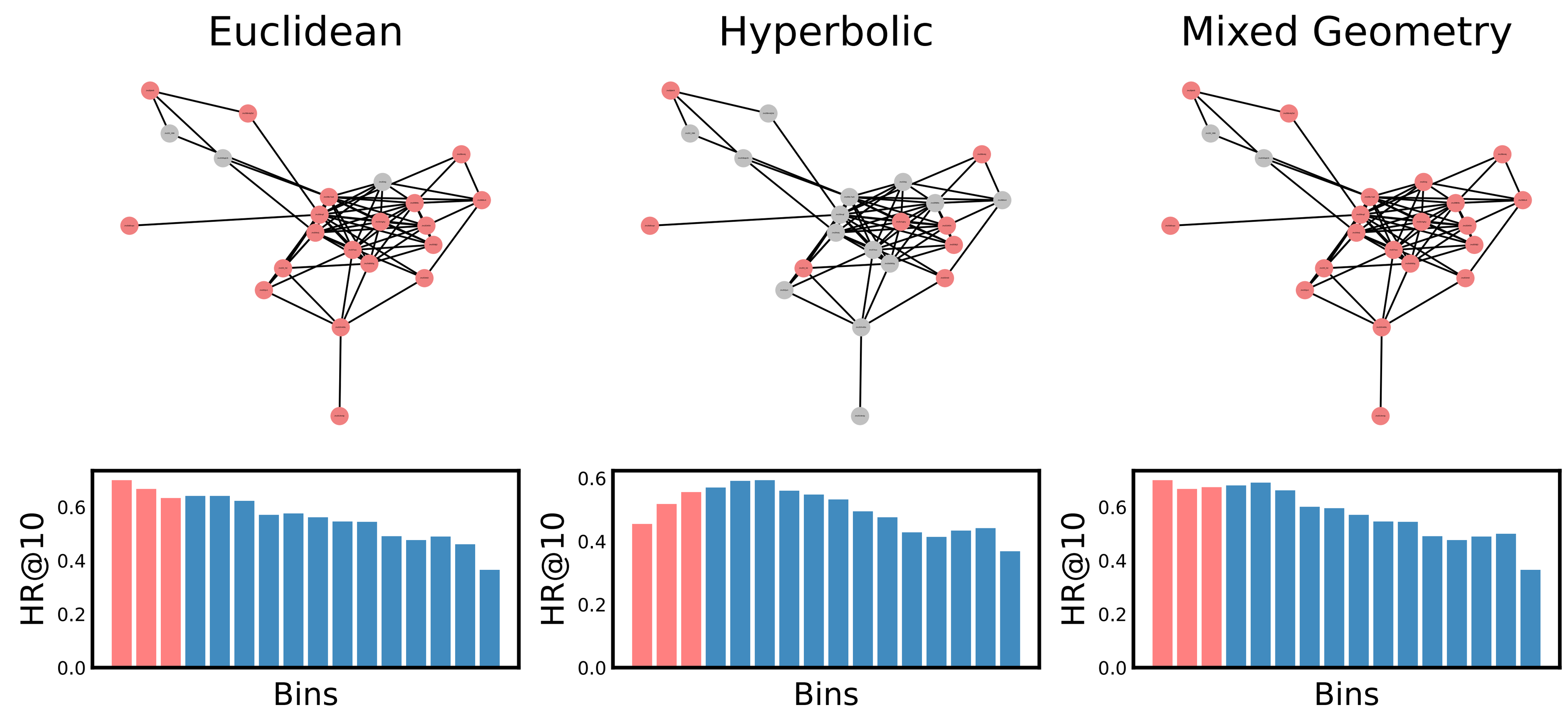

Figure 3: Comparison of performances of the Euclidean, hyperbolic and mixed-geometry models on FB15k-237.

- We introduce a new mixed-geometry tensor factorisation (MIG-TF) model that combines Tucker decomposition defined in the Euclidean space with a new hyperbolic ternary interaction term.
- We highlight intricacies of applying geometric approach to real-world knowledge graphs and demonstrate the associated with it limitations of using single-geometry modelling.
- The proposed combined approach significantly reduces the number of model parameters compared to state-of-the-art methods. It does so without sacrificing expressive power and achieves more accurate results in most of the common benchmarks.

Mixed Geometry Tensor Factorisation

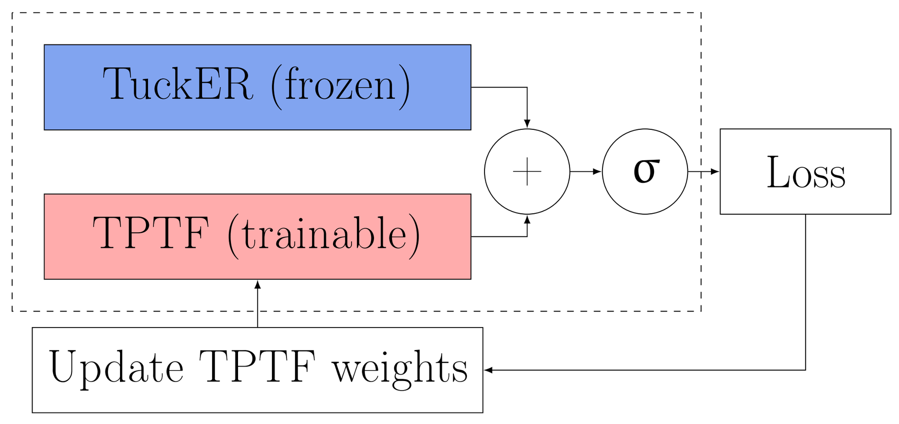

Figure 4: The proposed MIG-TF model architecture.

We introduce a shared-factor mixed geometry model combining Euclidean TuckER and hyperbolic TPTF models. The score function of our mixed-geometry model MIG-TF is the sum of score functions of the Lorentzian and Euclidean models:

$$(S_{\text{MIG-TF}})_i = (S_E)_i + (S_H)_i, \quad (2)$$

In MIG-TF model, we utilize pretrained TuckER model and optimize the hyperbolic term parameters (4) of the score function (2) to minimize the BCE loss:

$$\mathcal{L}_{\text{MIG-TF}} = \frac{1}{n_e} \sum_{i=1}^{n_e} l_{BCE}(a_i, \sigma((S_{\text{MIG-TF}})_i)),$$

Tetrahedron Pooling Tensor Factorization

We propose to modify triangle inequality to capture ternary interactions in hyperbolic space. In particular, we utilize the so-called tetrahedron inequality: for the points u, v, t, o in the Euclidean space, it holds

$$d(u, v) + d(o, t) \leq d(u, t) + d(v, t) + d(o, u) + d(o, v). \quad (3)$$

Consequently, we can naturally introduce the following "smoothed" and differentiable everywhere score function:

$$S_H(u, v, t) = \frac{\frac{1}{2}\left(\begin{matrix} d_{\mathcal{L}}^2(u, v) + d_{\mathcal{L}}^2(0, t) - d_{\mathcal{L}}^2(u, t) - \\ -d_{\mathcal{L}}^2(t, v) - d_{\mathcal{L}}^2(0, u) - d_{\mathcal{L}}^2(0, v) \end{matrix} \right)}{\langle 0, v\rangle_{\mathcal{L}}\langle 0, t\rangle_{\mathcal{L}} + \langle 0, u\rangle_{\mathcal{L}}\langle 0, v\rangle_{\mathcal{L}} + \langle 0, t\rangle_{\mathcal{L}}\langle 0, v\rangle_{\mathcal{L}}}, \quad (4)$$

We optimize the Euclidean embeddings and map them to the hyperboloid: $v \in \mathbb{R}^n$ maps to $[\sqrt{\beta + \|v\|_2^2}, v] \in \mathbb{R}^{n+1}$.

We propose to model ternary interactions in a knowledge graph via combination of pairwise interactions in hyperbolic space.

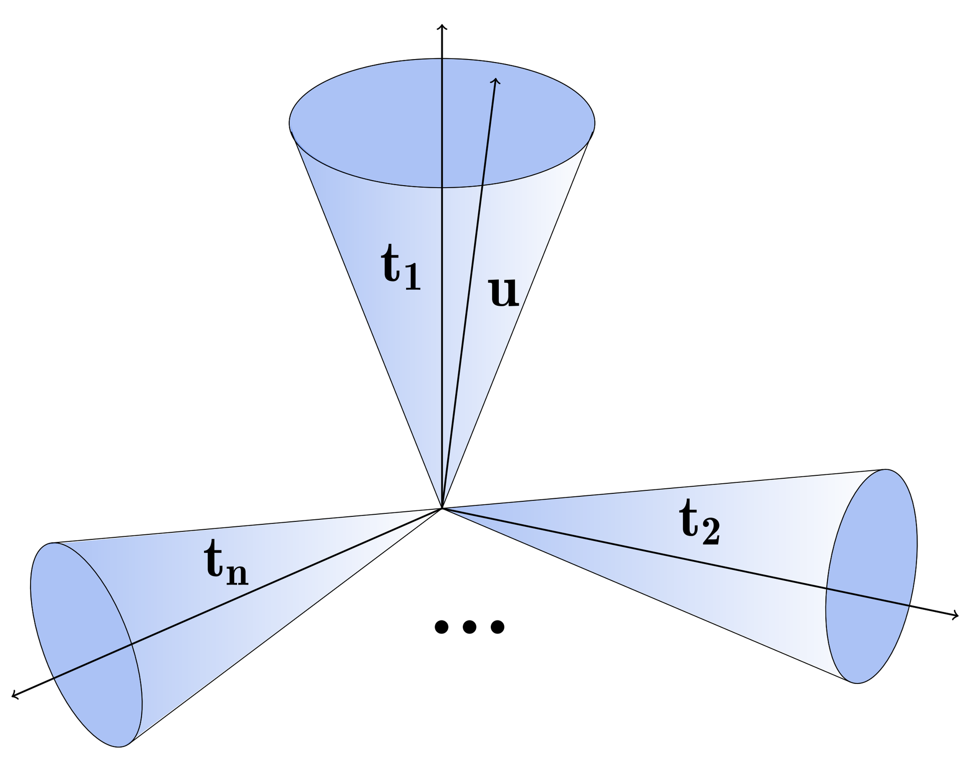

Figure 5: Each embedding of a relation $(t_1, ..., t_n)$ defines a cone that encompasses the embeddings of the entities associated with that relation.

Results

Models	FB15k-237	WN18RR	YAGO3-10
$\text{TuckER}_{S_E+0\cdot S_H}$	$4\cdot 10^6$	$8\cdot 10^6$	$25\cdot 10^6$
RotH	$40\cdot 10^6$	$80\cdot 10^6$	$120\cdot 10^6$
Our models			
$\text{TPTF}_{0\cdot S_E+S_H}$	$2\cdot 10^6$	$4\cdot 10^6$	$12\cdot 10^6$
$\text{MIG-TF}_{S_E+S_H}$	$5\cdot 10^6$	$10\cdot 10^6$	$31\cdot 10^6$

Table 1: Approximate number of models' parameters.

	FB15k-237		WN18RR		YAGO3-10	
Models	HR@1	MRR	HR@1	MRR	HR@1	MRR
TuckER	0.266	0.358	0.443	0.470	0.466	0.544
RotH	0.246	0.344	0.449	0.496	0.495	0.573
Our models						
TPTF	0.186	0.238	0.252	0.314	0.383	0.481
MIG-TF	<u>0.277</u>	<u>0.367</u>	0.450	0.496	0.501	0.579
MIG-TF_{QR}	0.276	0.365	<u>0.452</u>	<u>0.499</u>	<u>0.502</u>	<u>0.580</u>

Table 2: Metrics on knowledge graphs WN18RR, FB12k-237 and YAGO3-10. <u>Underlined</u> means the best metric.