Knowledge Graph Embedding via Adaptive Sparse Dynamic Mapping Matrix

汇报人：黄婷

2018.04.27

目录

CONTENTS

01

理　论

2013-TransE

$$f_r(h,t) = \|\mathbf{h} + \mathbf{r} - \mathbf{t}\|_\ell$$

2015-TransR

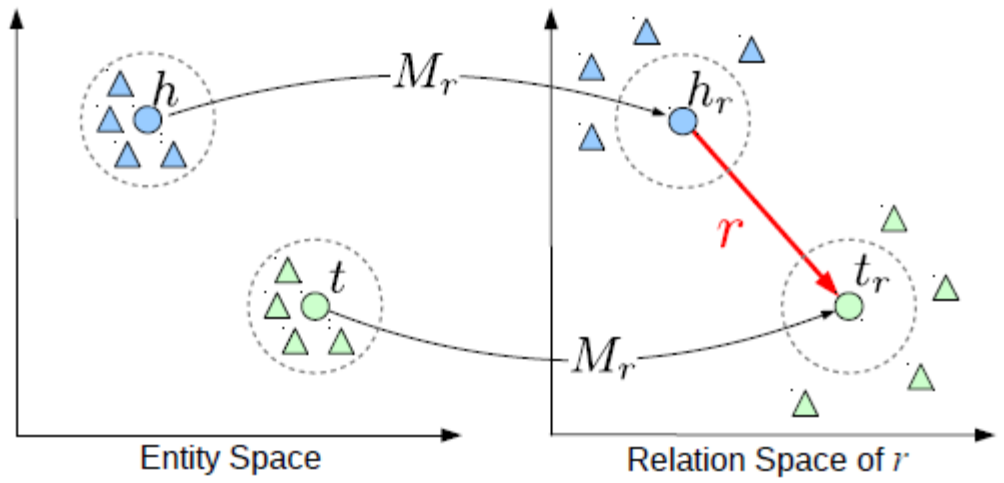


Figure 1: Simple illustration of TransR.

$$\mathbf{h}_r = \mathbf{h}\mathbf{M}_r, \quad \mathbf{t}_r = \mathbf{t}\mathbf{M}_r.$$

$$f_r(h,t) = \|\mathbf{h}_r + \mathbf{r} - \mathbf{t}_r\|_2^2.$$

- Heterogeneous
  - some relations link many entity pairs (called complex relations) and others do not (called simple relations)
- Unbalanced
  - some relations link many head (tail) entities and fewer tail (head) entities

2016-TranSparse

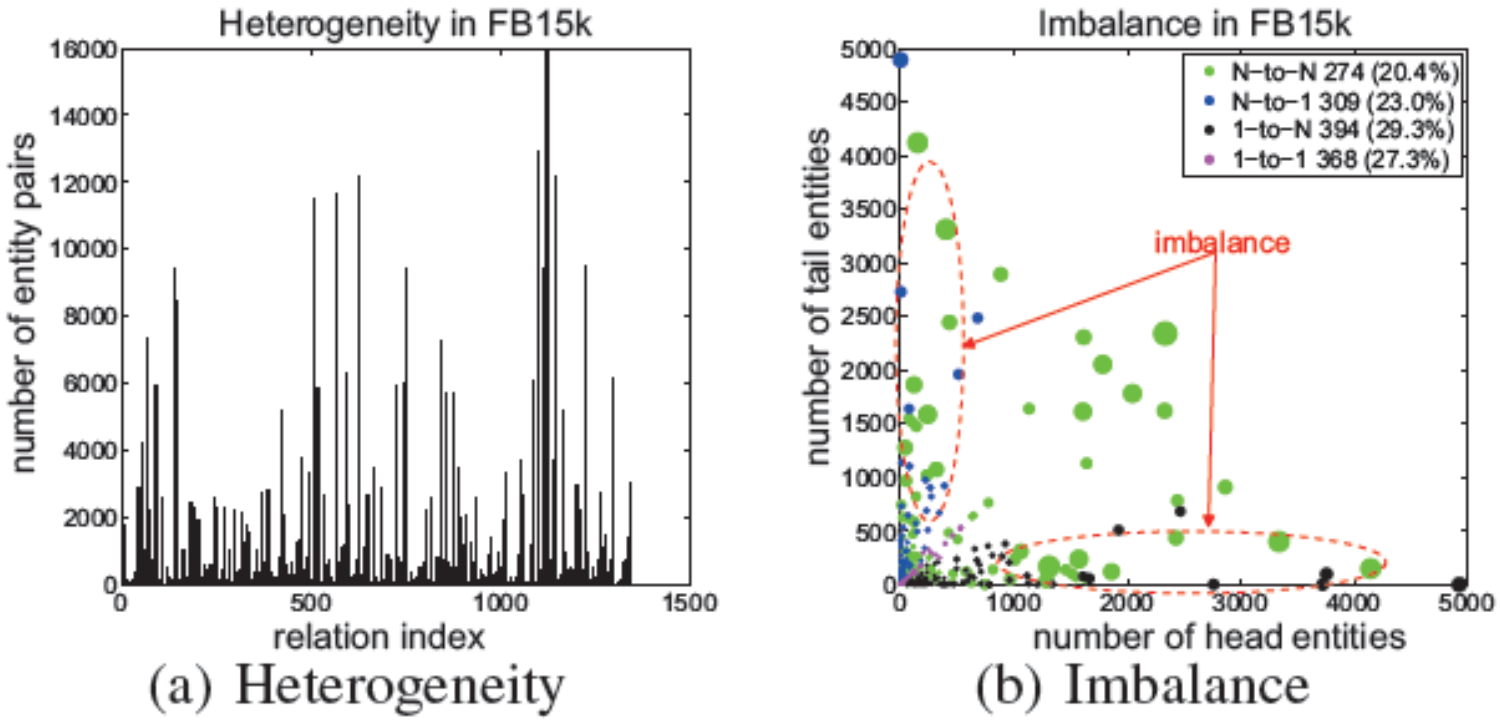
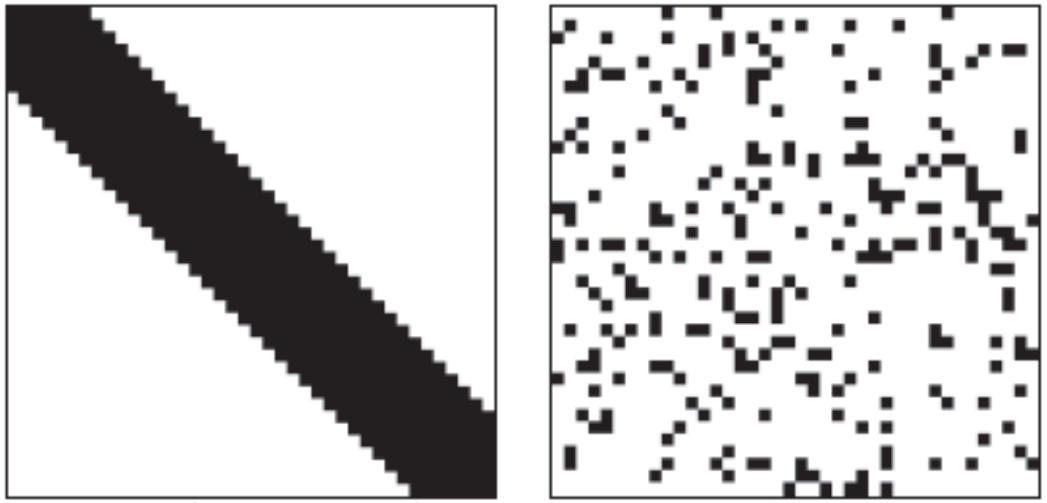


Figure 1: Relation statistics of FB15k which contains 1,345 relations. (a) shows the heterogeneity of FB15k. In (b), each circle represents a relation. Area of the circle is proportional to the number of entity pairs linked by relations in train data.

- Sparse Matrix
  - Sparse matrices refer to the matrices in which most entries are **zeros**. The fraction of zero elements over the total number of elements in a matrix is called **sparse degree** (denoted by θ).We use $M(\theta)$ to denote a matrix M with sparse degree θ.

- Sparse Matrix



(a) Structured (b) Unstructured

Figure 2: Sparse Patterns: (a) Structured. (b) Unstructured. Black squares represent the nonzero elements. The two sparse matrices have the same number of nonzero entries.

- Heterogeneous
  - We propose a model TranSparse(share), in which the sparse degrees of transfer matrices are determined by the number of entity pairs linked by relations and the two sides of relations share the same transfer matrices.
- Unbalanced
  - TranSparse(separate)

- TranSparse(share)
  - Sparse degrees

$$\theta_r = 1 - (1 - \theta_{min})N_r/N_{r^*}.$$

  - Projected vectors

$$\mathbf{h}_p = \mathbf{M}_r(\theta_r)\mathbf{h}, \quad \mathbf{t}_p = \mathbf{M}_r(\theta_r)\mathbf{t}. \tag{2}$$

where $\mathbf{M}_r(\theta_r) \in \mathbb{R}^{m\times n}$, $\mathbf{h}, \mathbf{t} \in \mathbb{R}^n$ and $\mathbf{h}_p, \mathbf{t}_p \in \mathbb{R}^m$.

- TranSparse(separate)
  - Sparse degrees

$$\theta_r^l = 1 - (1 - \theta_{min}) N_r^l / N_{r^*}^{l^*} \quad (l = h, t).$$

  - Projected vectors

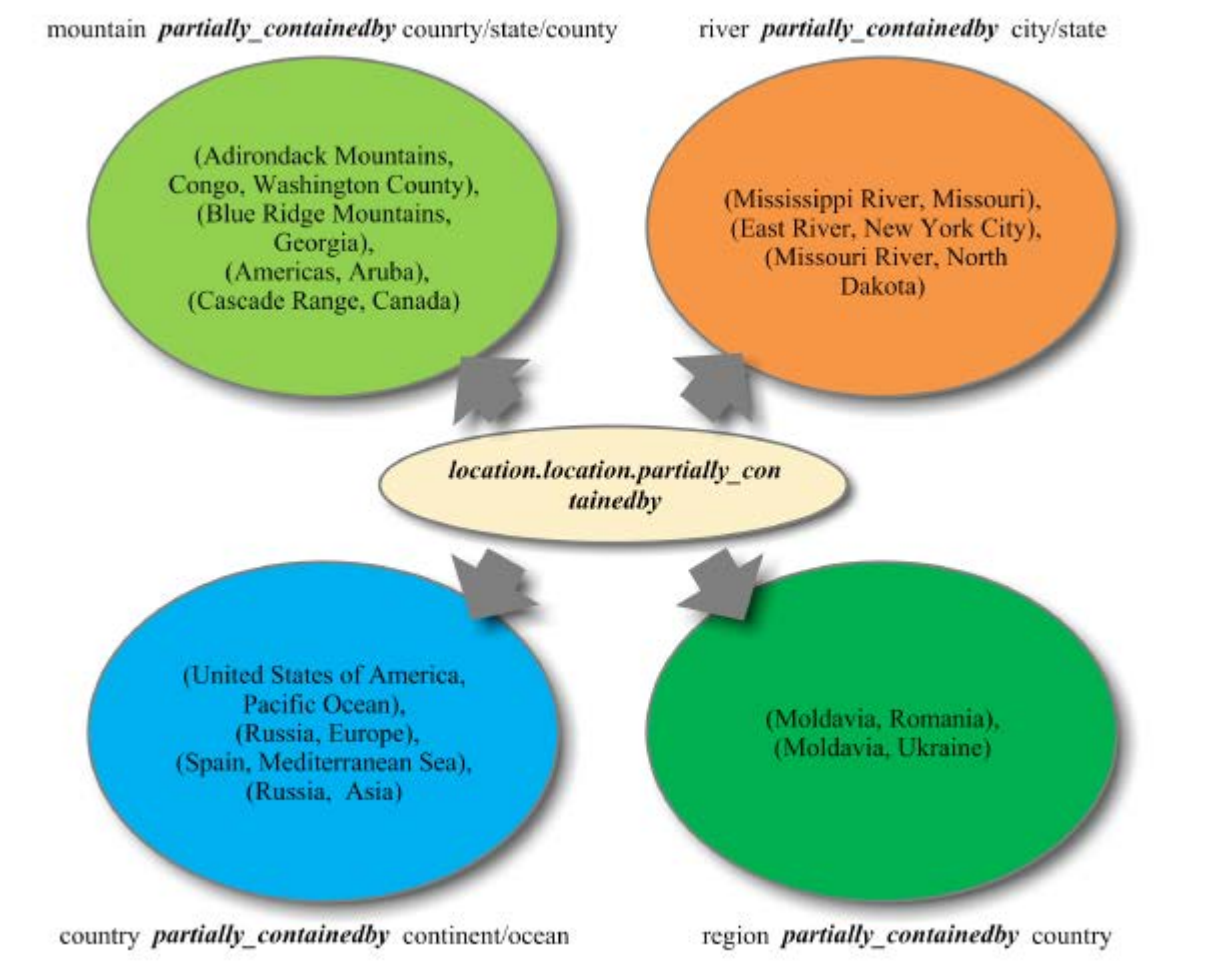
$$\mathbf{h}_p = \mathbf{M}_r^h(\theta_r^h)\mathbf{h}, \quad \mathbf{t}_p = \mathbf{M}_r^t(\theta_r^t)\mathbf{t}.$$

  - Score function for both

$$f_r(\mathbf{h}, \mathbf{t}) = \|\mathbf{h}_p + \mathbf{r} - \mathbf{t}_p\|_{\ell_{1/2}}^2.$$

2015-TranD:Dynamic Mapping Matrix

- Multiple Types of Entities and Relations



2015-TranD:Dynamic Mapping Matrix

- Entity : h, h_p, t, t_p
- Relation : r, r_p

$$\mathbf{M}_{rh} = \mathbf{r}_p \mathbf{h}_p^\top + \mathbf{I}^{m \times n} \quad (11)$$

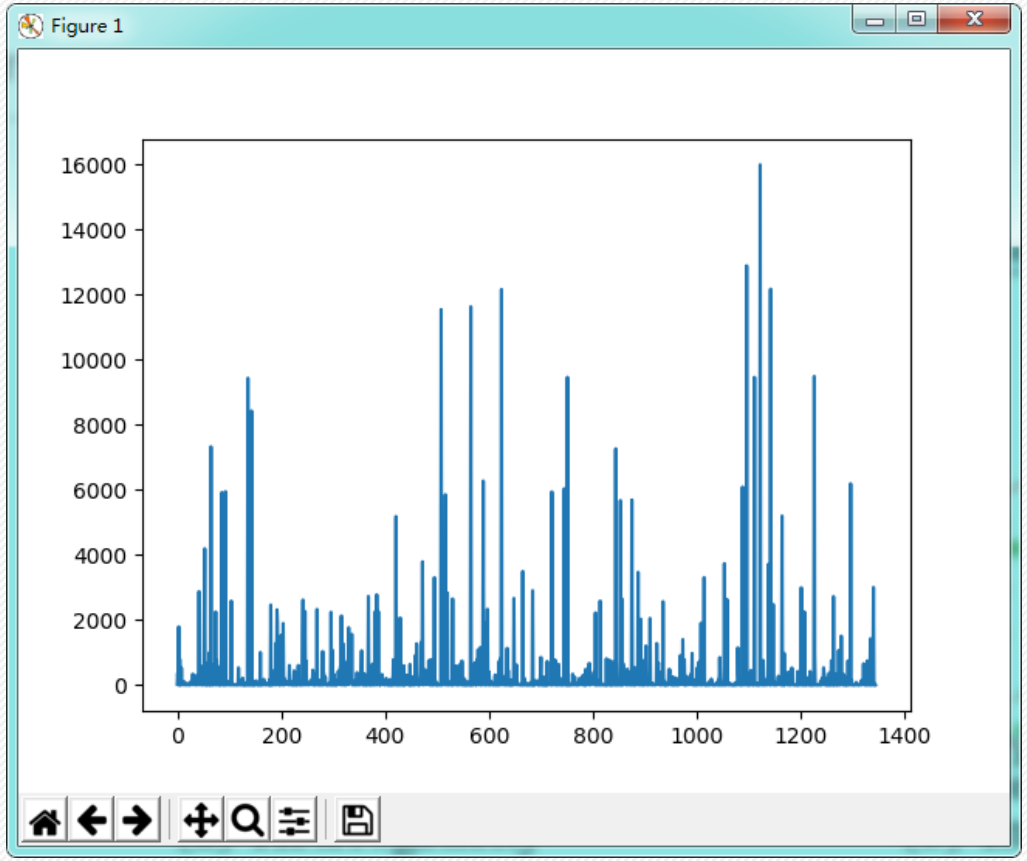
$$\mathbf{M}_{rt} = \mathbf{r}_p \mathbf{t}_p^\top + \mathbf{I}^{m \times n} \quad (12)$$

$$\mathbf{h}_\perp = \mathbf{M}_{rh}\mathbf{h}, \quad \mathbf{t}_\perp = \mathbf{M}_{rt}\mathbf{t} \quad (13)$$

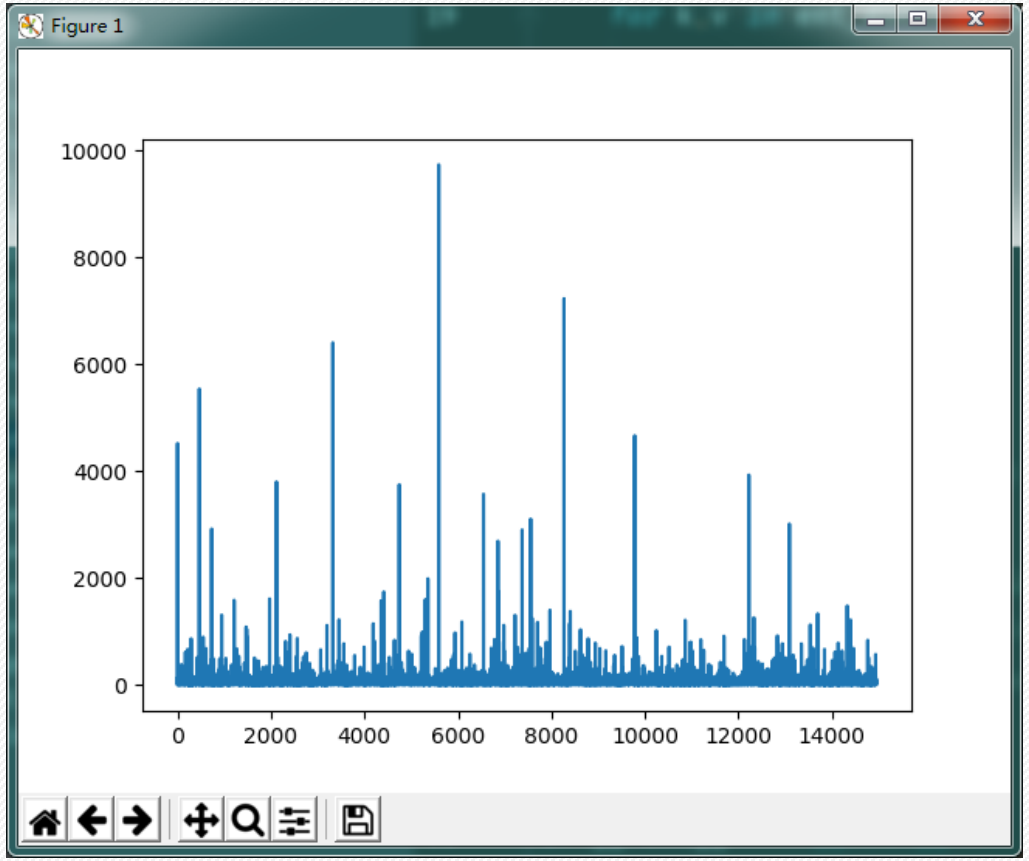
$$f_r(\mathbf{h}, \mathbf{t}) = -\|\mathbf{h}_\perp + \mathbf{r} - \mathbf{t}_\perp\|_2^2 \quad (14)$$

Adaptive Sparse Dynamic Mapping Matrix

- Heterogeneity in relations and entities on FB15k



relations



entities

Adaptive Sparse Dynamic Mapping Matrix

- sparse projection vectors for entities and relations

$$\theta_l = 1 - (1 - \theta_{\min})\, N_l / N_{l^*} \quad (l = h, t) \tag{1}$$

$$M_{rh} = r_p h_p^T + I^{m \times n} \longrightarrow M_{rh} = r_p(\theta_r) h_p^T(\theta_h) + I^{m \times n} \tag{2}$$

$$\mathbf{h}_\perp = \mathbf{M}_{rh}\mathbf{h}, \quad \mathbf{t}_\perp = \mathbf{M}_{rt}\mathbf{t} \tag{3}$$

$$f_r(\mathbf{h}, \mathbf{t}) = -\|\mathbf{h}_\perp + \mathbf{r} - \mathbf{t}_\perp\|_2^2 \tag{4}$$

02

实　验

Link Prediction

- mean rank and hit@10

Data sets	WN18				FB15K			
Metric	Mean Rank		Hits@10		Mean Rank		Hits@10	
	Raw	Filt	Raw	Filt	Raw	Filt	Raw	Filt
Unstructured (Bordes et al. 2012)	315	304	35.3	38.2	1,074	979	4.5	6.3
RESCAL (Nickle, Tresp, and Kriegel 2011)	1,180	1,163	37.2	52.8	828	683	28.4	44.1
SE (Bordes et al. 2011)	1,011	985	68.5	80.5	273	162	28.8	39.8
SME (linear) (Bordes et al.2012)	545	533	65.1	74.1	274	154	30.7	40.8
SME (Bilinear) (Bordes et al. 2012)	526	509	54.7	61.3	284	158	31.3	41.3
LFM (Jenatton et al. 2012)	469	456	71.4	81.6	283	164	26.0	33.1
TransE (Bordes et al. 2013)	263	251	75.4	89.2	243	125	34.9	47.1
TransH (unif) (Wang et al. 2014)	318	303	75.4	86.7	211	84	42.5	58.5
TransH (bern) (Wang et al. 2014)	401	388	73.0	82.3	212	87	45.7	64.4
TransR (unif) (Lin et al. 2015)	232	219	78.3	91.7	226	78	43.8	65.5
TransR (bern) (Lin et al. 2015)	238	225	79.8	92.0	198	77	48.2	68.7
CTransR (unif) (Lin et al. 2015)	243	230	78.9	92.3	233	82	44.0	66.3
CTransR (bern) (Lin et al. 2015)	231	218	79.4	92.3	199	75	48.4	70.2
TransD (unif)(base)	242	229	79.2	92.5	211	67	49.4	74.2
TransD (bern)(base)	224	212	79.6	92.2	194	91	53.4	77.3
TranSparse (unif)	233	221	79.6	93.4	216	**66**	50.3	78.4
TranSparse (bern)	223	211	80.1	93.2	**190**	82	**53.7**	79.9
Sparse TransD (unif)	233	220	80.0	93.9	247	89	48.7	78.1
Sparse TransD (bern)	**220**	**208**	**81.6**	**95.3**	203	73	51.6	**80.2**

Link Prediction

- mean rank and hit@10

Tasks	Prediction Head (Hits@10)				Prediction Tail (Hits@10)			
Relation Category	1-to-1	1-to-N	N-to-1	N-to-N	1-to-1	1-to-N	N-to-1	N-to-N
Unstructured (Bordes et al. 2012)	34.5	2.5	6.1	6.6	34.3	4.2	1.9	6.6
SE (Bordes et al. 2011)	35.6	62.6	17.2	37.5	34.9	14.6	68.3	41.3
SME (linear) (Bordes et al.2012)	35.1	53.7	19.0	40.3	32.7	14.9	61.6	43.3
SME (Bilinear) (Bordes et al. 2012)	30.9	69.6	19.9	38.6	28.2	13.1	76.0	41.8
TransE (Bordes et al. 2013)	43.7	65.7	18.2	47.2	43.7	19.7	66.7	50.0
TransH (unif) (Wang et al. 2014)	66.7	81.7	30.2	57.4	63.7	30.1	83.2	60.8
TransH (bern) (Wang et al. 2014)	66.8	87.6	28.7	64.5	65.5	39.8	83.3	67.2
TransR (unif) (Lin et al. 2015)	76.9	77.9	38.1	66.9	76.2	38.4	76.2	69.1
TransR (bern) (Lin et al. 2015)	78.8	89.2	34.1	69.2	79.2	37.4	90.4	72.1
CTransR (unif) (Lin et al. 2015)	78.6	77.8	36.4	68.0	77.4	37.8	78.0	70.3
CTransR (bern) (Lin et al. 2015)	81.5	89.0	34.7	71.2	80.8	38.6	90.1	73.8
TransD (unif)(base)	80.7	85.8	47.1	75.6	80.0	54.5	80.7	77.9
TransD (bern)(base)	86.1	95.5	39.8	78.5	85.4	50.6	94.4	81.2
TranSparse (unif)	83.2	85.2	**51.8**	80.3	82.6	**60.0**	85.5	82.5
TranSparse (bern)	**87.1**	**95.8**	44.4	81.2	**87.5**	57.0	**94.5**	83.7
Sparse TransD (unif)	79.0	90.8	42.1	80.2	77.4	56.7	85.4	82.7
Sparse TransD (bern)	81.9	95.1	45.4	**81.4**	81.2	55.4	93.5	**84.4**

感谢聆听

汇报人：黄婷