Attribute-driven Capsule Network for Entity Relation Prediction

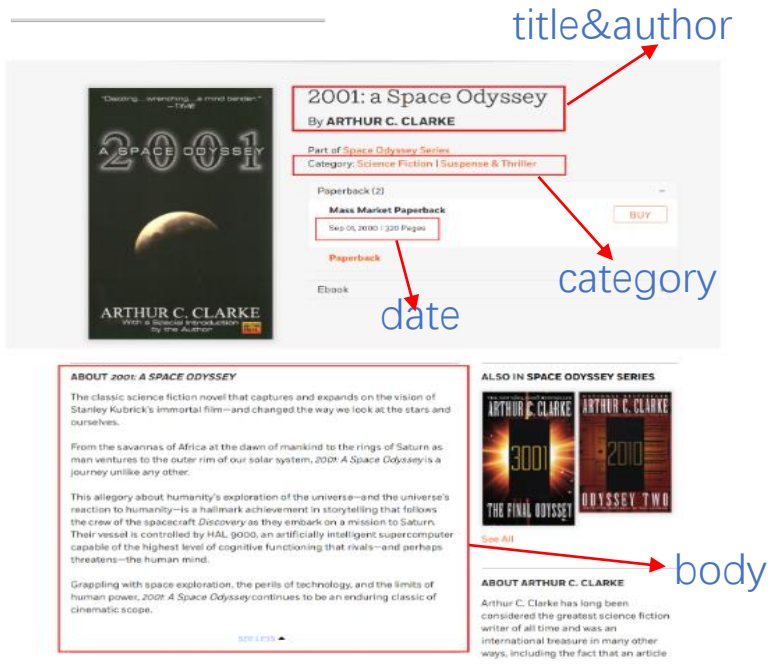
Jiayin Chen[1], Xiaolong Gong[2], Xi Chen[1], Zhiyi Ma[1]

[1] Advanced Institute of Information Technology(AIIT), Peking University, Hangzhou, China
[2] Department of Computer Science, Shanghai Jiao Tong University, Shanghai, China

2020/04/26

目录 / CONTENTS

Motivation



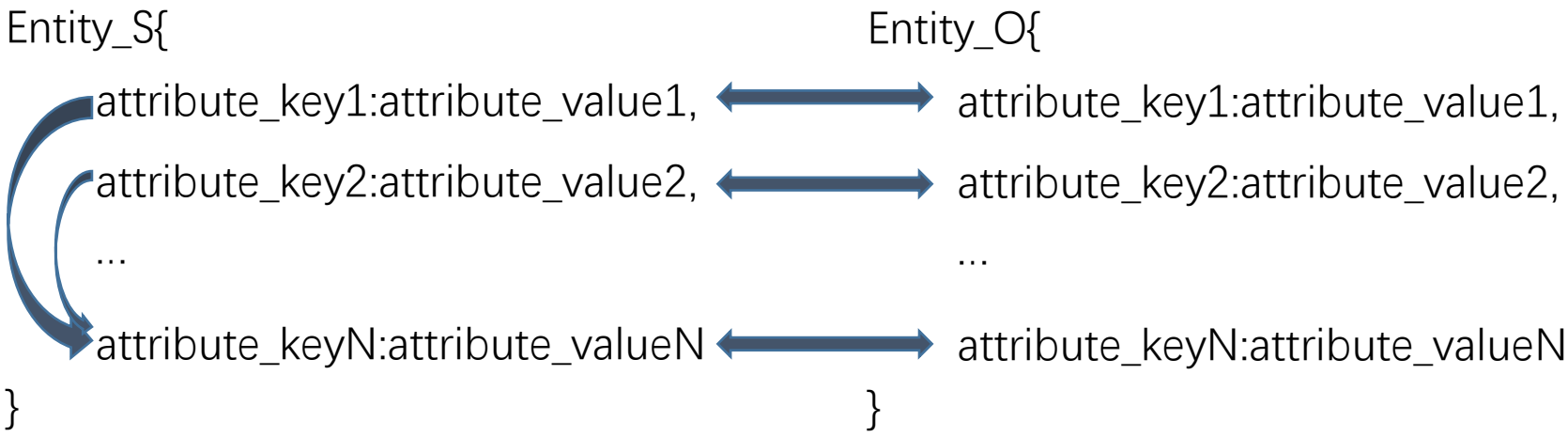
book datasets

company datasets

Task: predict relationship between a pair of entities with common multi-attribute.

Motivation

Problem1: How to learn semantic correlations between attributes?



Problem2: How to learn relation information between common attributes of a pair entities?

Motivation

A real-world scenario CompanyRelationCollection(CRC) dataset

< Entity_S	Customer	Entity_O >


```
"Entity_S": {
 "company_name": "红星美凯龙家居集团股份有限公司",
 "company_address": "上海市浦东新区临御路***号*楼***室",
 "company_type": "股份有限公司/外商投资企业",
 "industry_category": "商业贸易 — 零售",
 "business_scope": "为所投资企业提供管理服务，企业管理咨询..."
 "abstract": "红星美凯龙家居集团股份有限公司（简称：红星美凯龙）..."
},
```

```
"Entity_O": {
 "company_name ": "郑州华商汇房地产开发有限公司",
 "company_address": "新郑市郭店镇***号",
 "company_type": "有限责任公司",
 "industry_category": "房地产业",
 "business_scope": "房地产开发、经营。（凭有效资质证经营）",
 "abstract": "郑州华商汇房地产开发有限公司成立于..."
},
```

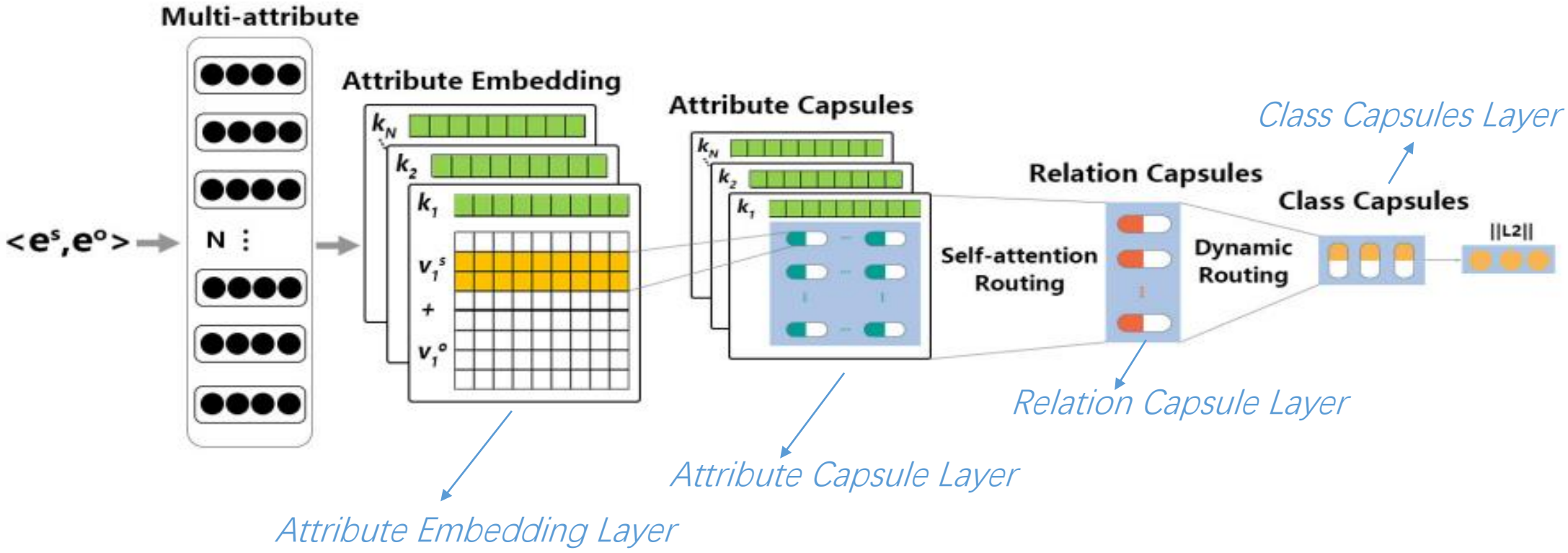

The relations of CRC include customer(C), provider(P), rival(R)

Approach

What we do :

- Apply capsule networks to entity relation prediction.
- Propose self-attention routing method for attribute information representation.
- Construct a new real-world multi-attribute entity relation dataset.

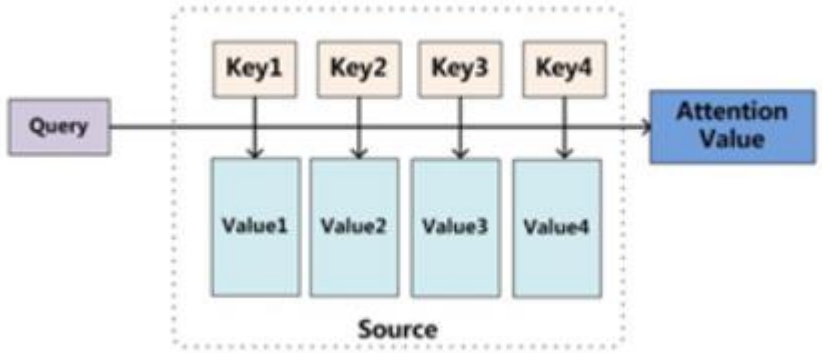
Approach



The whole framework of our model that consists of four layers.

Approach

self- attention



$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V$$

self- attention routing(our proposed)

$$a_i = softmax\left(\frac{E^{k_j}W^q(F_sW^k)^T}{\sqrt{d_w}}\right)$$

$$Q = E^{k_j}W^q \qquad K = F_sW^k$$

$$A = [a_1, a_2, ..., a_B].$$

$$S = C \odot A$$

C represents attribute capsules

A represents weights of attribute capsules

Results

datasets:

- CompanyRelationCollection (CRC)
  https://github.com/cjymz886/ACNet

- BlurbGenreCollection(BGC)
  https://github.com/uhh-lt/BlurbGenreCollection-HMC

Table 1. Quantitative characteristics of both datasets

	CRC	BGC
Number of entities	58,013	91,892
Number of attributes per entity	6	3
Total number of relationships	3(C,P,R)	3(S,P,D)
Number of relational pairs	61,441	918,920
Train set	43,009	735,136
Validation set	9216	91,892
Test set	9216	91,892

Baselines:

- CNN， PCNN,
- BLSTM,ATT-BLSTM
- BERT
- Basic-Caps (our model without self-attention routing)
- ACNet (our model with self-attention routing)

Evaluation metrics:

- Precision
- Recall
- F1

Results

Table 3. The results of Comparison of different methods. Best scores are in bold.

Method	CRC			BGC		
	Precision	Recall	F1	Precision	Recall	F1
CNN[10]	0.7706	0.7012	0.7343	0.8420	0.8265	0.8342
PCNN[11]	0.7825	0.7103	0.7447	0.8578	0.8299	0.8436
BLSTM[12]	0.7682	0.7066	0.7361	0.85378	0.8122	0.8324
ATT-BLSTM[13]	0.7694	0.7043	0.7359	0.8621	0.8017	0.8308
BERT[27]	**0.8067**	0.6936	0.7459	**0.8628**	0.8345	0.8484
Basic-Caps	0.7528	0.7331	0.7428	0.8534	0.8304	0.8417
ACNet	0.7662	**0.7405**	**0.7531**	0.8612	**0.8405**	**0.8507**

- Our attribute-driven capsule network achieves the highest F1 scores.
- The self-attention routing approach is effective.
- Powerful pre-training model can achieve richful representation information.

Results

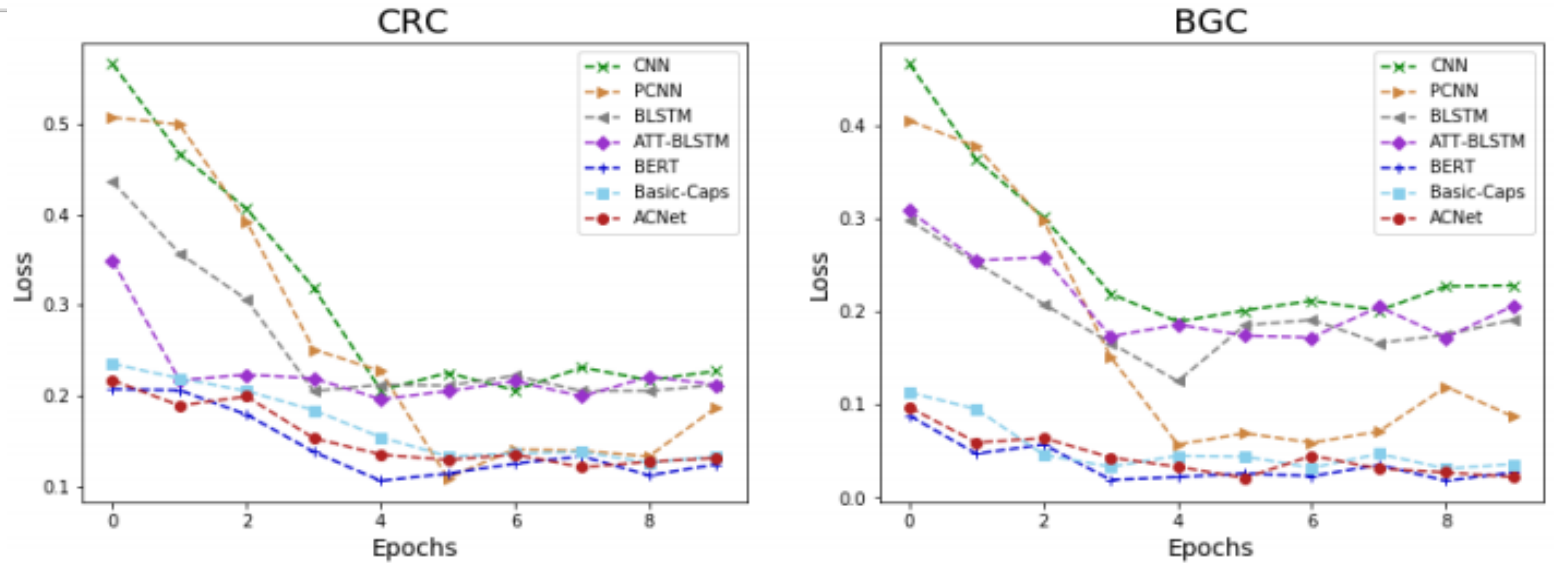


Fig. 3. The result of training loss from all models on two datasets.

- Convergence of capsule networks is much faster than other models.
- Capsule networks get a stable training process and achieve less fluctuation.

THANKS FOR YOUR ATTENTION

欢迎关注