

CANE: Context-Aware Network Embedding for Relation Modeling

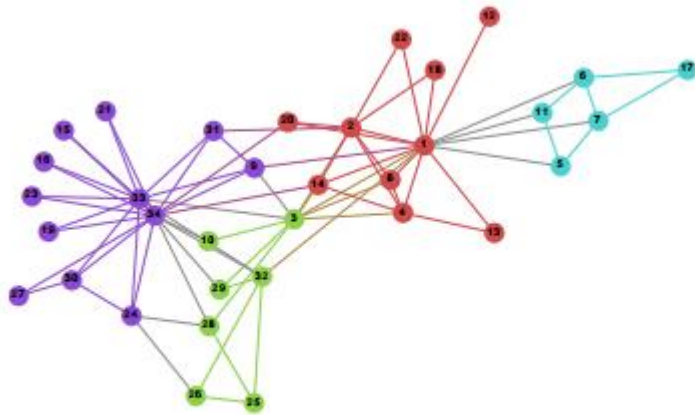
Cunchao Tu

NLP Group, Dept. of CS&T, Tsinghua University

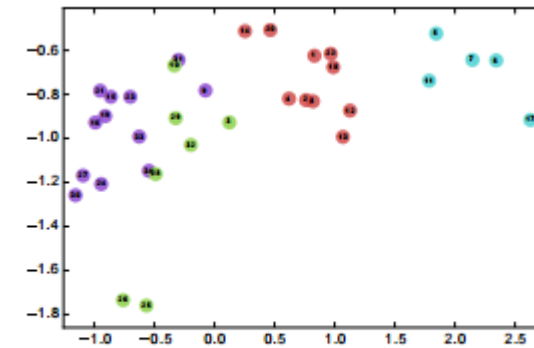


Network Embedding

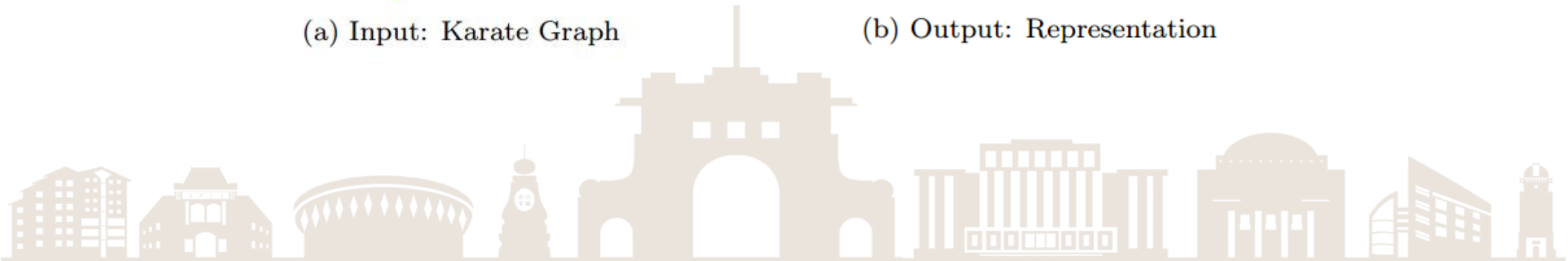
- Learn a low-dimensional embedding for each vertex in networks



(a) Input: Karate Graph

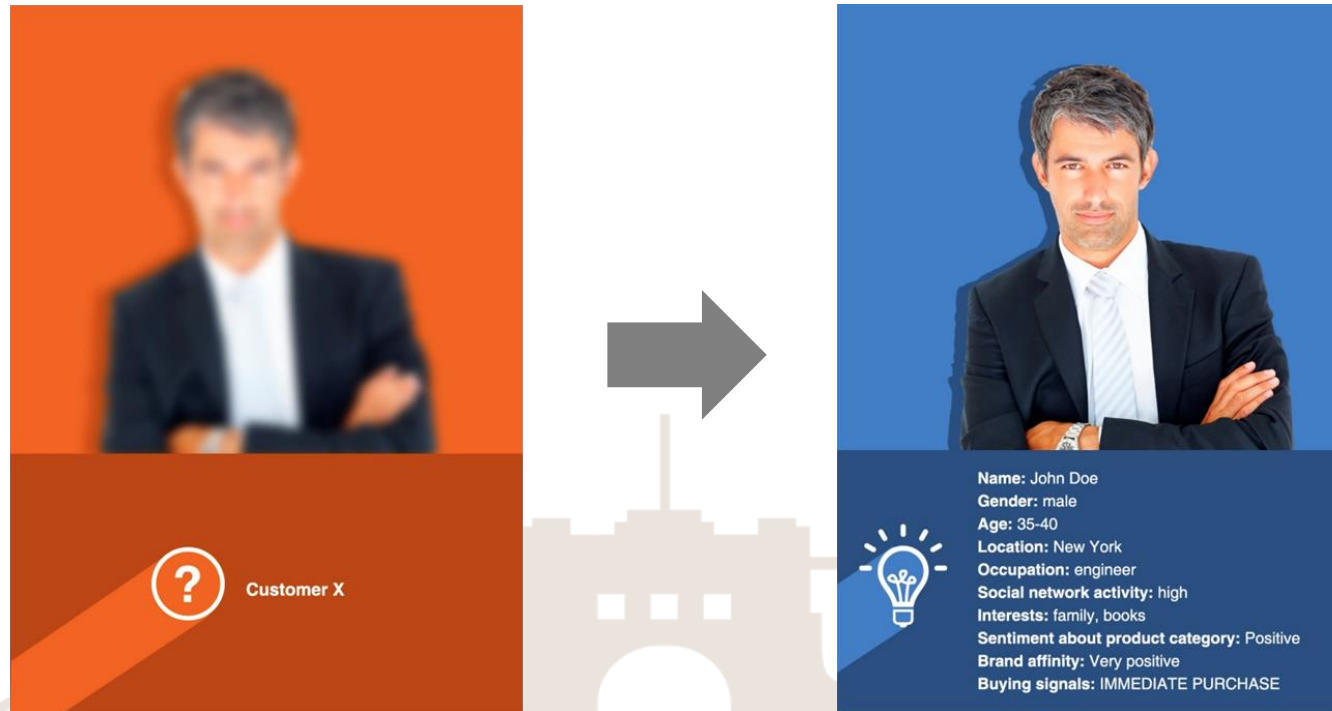


(b) Output: Representation



Network Embedding

- Embeddings as features
 - Vertex classification -> **User profiling/ Anomaly detection**



Network Embedding

- Embeddings as features
 - Vertex classification -> User profiling/ Anomaly detection
 - Link prediction -> Friend recommendation



Typical NE Models

- Context-free embeddings
 - Learn a **fixed** embedding for each vertex
 - Vertices show **various aspects** when interacting with different neighbors

Existing models ignore aspect transition of vertices!



Example

- An NLP researcher



I am studying NLP problems,
including syntactic parsing,
machine translation and so on.



Example

- Co-author with other researchers



← Co-author →



I am studying **NLP** problems,
including syntactic parsing,
machine translation and so on.

My research focuses on typical
NLP tasks, including word
segmentation, tagging and
syntactic parsing.



Example

- Co-author with other researchers



◀ Co-author ▶



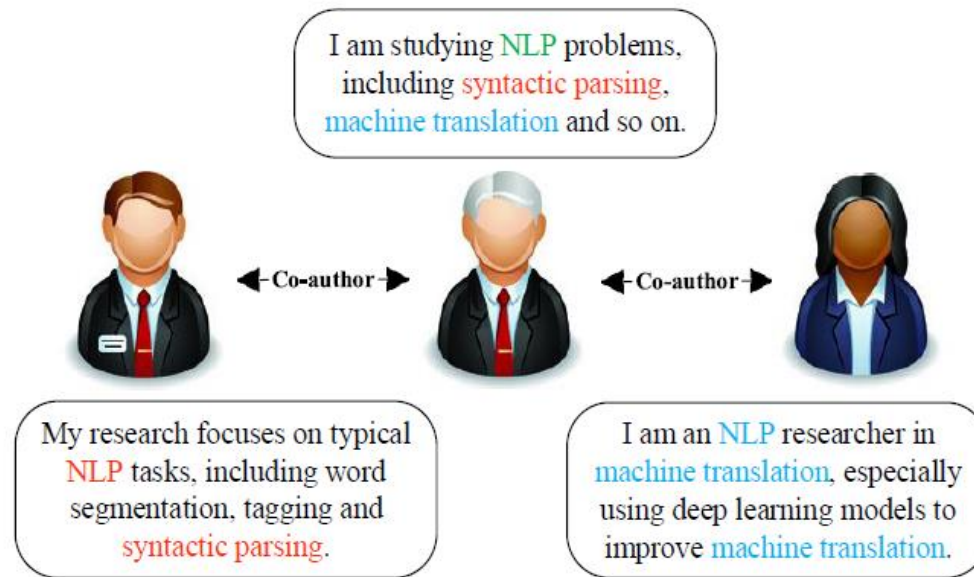
I am studying **NLP** problems, including syntactic parsing, **machine translation** and so on.

I am an **NLP** researcher in **machine translation**, especially using deep learning models to improve **machine translation**.



Context-Aware Network Embedding

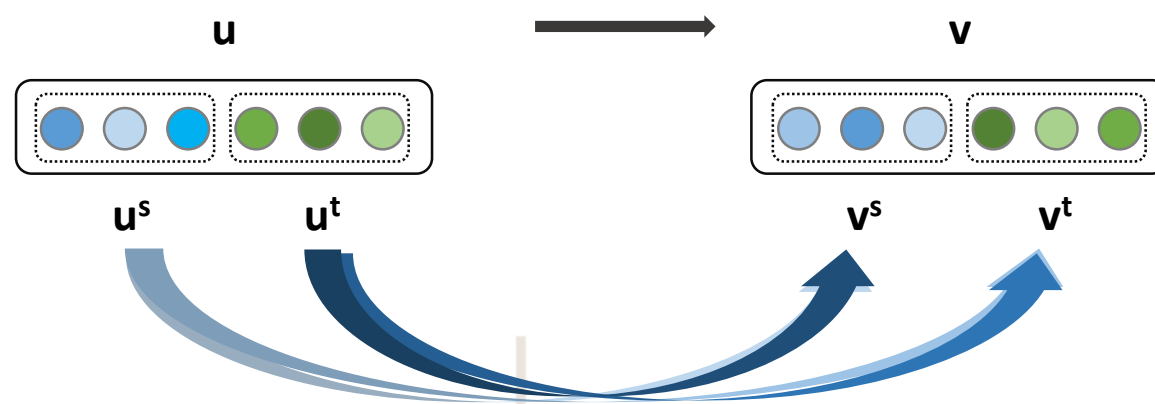
- **Text-based** information networks
- **Dynamic** embeddings according to different **neighbors**
- Model **relationships** between vertices more precisely



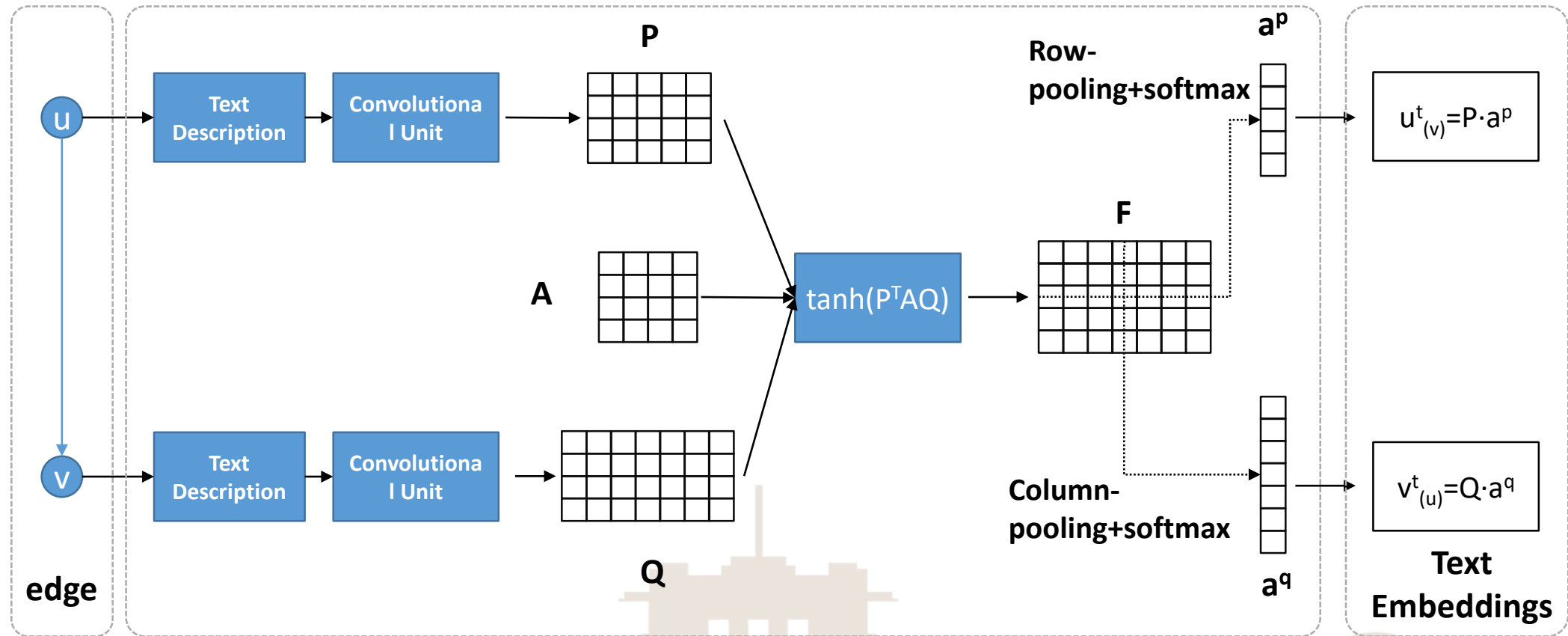
Context-Aware Text Embedding

- Maximize the log-likelihood of each edge $e=(u, v)$

$$L_{ss}(e) + \alpha \cdot L_{st}(e) + \beta \cdot L_{ts}(e) + \gamma \cdot L_{tt}(e)$$



Context-Aware Text Embedding



Experiments

- Datasets
 - **Cora**: paper citation network
 - **HepTh**: (High Energy Physics Theory) citation network from arxiv.
 - **Zhihu**: online Q&A website in China

Datasets	Cora	HepTh	Zhihu
#Vertices	2,277	1,038	10,000
#Edges	5,214	1,990	43,894
#Labels	7	—	—



Experiments

- NE Baselines
 - **Structure-only:** MMB, DeepWalk, LINE, node2vec
 - **Structure and text:** naive combination, TADW, CENE
- Tasks
 - Link prediction (AUC value)
 - Vertex classification (Accuracy)



Link Prediction

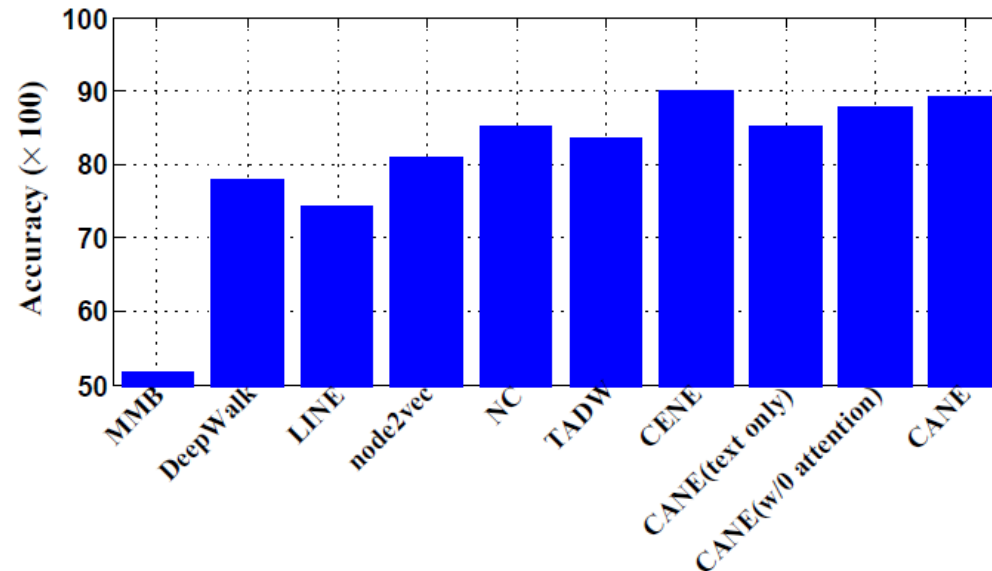
#Removed edges	15%	25%	35%	45%	55%	65%	75%	85%	95%
MMB	54.6	57.9	57.3	61.6	66.2	68.4	73.6	76.0	80.3
DeepWalk	55.2	66.0	70.0	75.7	81.3	83.3	87.6	88.9	88.0
LINE	53.7	60.4	66.5	73.9	78.5	83.8	87.5	87.7	87.6
Node2vec	57.1	63.6	69.9	76.2	84.3	87.3	88.4	89.2	89.2
NC	78.7	82.1	84.7	88.7	88.7	91.8	92.1	92.0	92.7
TADW	87.0	89.5	91.8	90.8	91.1	92.6	93.5	91.9	91.7
CENE	86.2	84.6	89.9	91.2	92.3	91.8	93.2	92.9	93.2
CANE(Text)	83.8	85.2	87.3	88.9	91.1	91.2	91.8	93.1	93.5
CANE(w/o attention)	84.5	89.3	89.2	91.6	91.1	91.8	92.3	92.5	93.6
CANE	90.0	91.2	92.0	93.0	94.2	94.6	95.4	95.7	96.3

Vertex Classification

- Transform context-aware embeddings to context-free embeddings

$$u = \frac{1}{N} \sum_{(u,v)|(v,u) \in E} u_{(v)}$$

- Comparable performance



Mutual Attention

- Edge (A, B) and (A, C)

Machine Learning research making great progress many directions This article summarizes four directions discusses current open problems The four directions improving classification accuracy learning ensembles classifiers methods scaling supervised learning algorithms reinforcement learning learning complex stochastic models

The problem making optimal decisions uncertain conditions central Artificial Intelligence If state world known times world modeled Markov Decision Process MDP MDPs studied extensively many methods known determining optimal courses action policies The realistic case state information partially observable Partially Observable Markov Decision Processes POMDPs received much less attention The best exact algorithms problems inefficient space time We introduce Smooth Partially Observable Value Approximation SPOVA new approximation method quickly yield good approximations improve time This method combined reinforcement learning methods combination effective test cases

Machine Learning research making great progress many directions This article summarizes four directions discusses current open problems The four directions improving classification accuracy learning ensembles classifiers methods scaling supervised learning algorithms reinforcement learning learning complex stochastic models

In context machine learning examples paper deals problem estimating quality attributes without dependencies among Kira Rendell developed algorithm called RELIEF shown efficient estimating attributes Original RELIEF deal discrete continuous attributes limited twoclass problems In paper RELIEF analysed extended deal noisy incomplete multiclass data sets The extensions verified various artificial one well known realworld problem

Conclusion

- Context-aware network embedding
 - Learns **dynamic** text embeddings according to the neighbors.
- Mutual attention
 - Makes the relations between vertices **explicit** and **interpretable**.
- Experiments
 - CANE is naturally applicable to **link prediction** task.
 - The learnt context-aware embeddings can **transform** into high-quality context-free embeddings.



Q&A



Homepage



Code and Datasets



Email: tucunchao@gmail.com