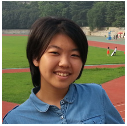
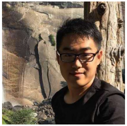
TIMME: Twitter Ideology-detection via Multi-task Multi-relational Embedding

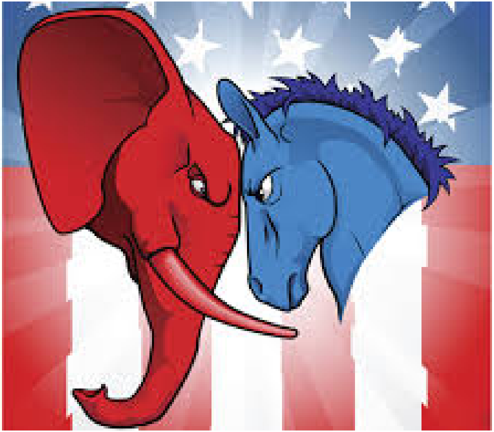
Zhiping (Patricia) Xiao, Weiping Song, Haoyan Xu,
Zhicheng Ren, Yizhou Sun

KDD'20 Applied Data Science Track

Outline

TIMME



The picture comes from `http://www.marekrei.com/blog/political-ideology-detection/`.

Motivation: Data Source

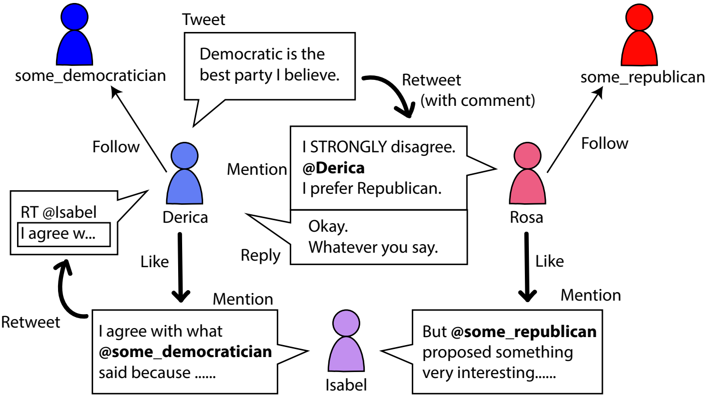
https://www.congress.gov/congressional-record

Motivation: Twitter Data

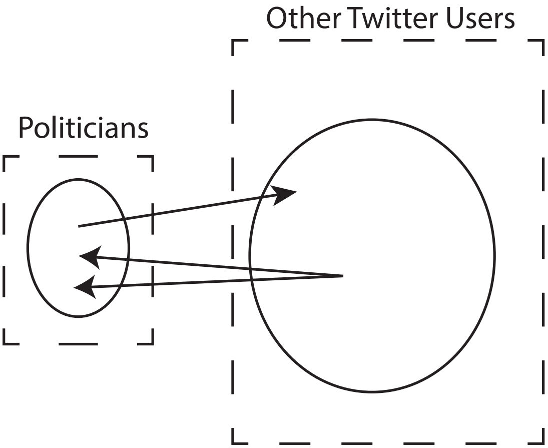
Example: US Presidents on Twitter

API: `https://developer.twitter.com/en`

Motivation: Ideology Detection on Twitter



Problem: Ideology Classification on Twitter


Other Twitter Users
Politicians

Datasets

	PureP	P50	P20∼50	P+all
# User	583	5,435	12,103	20,811
# Link	122,347	1,593,721	1,976,985	6,496,107
# Labeled User	581	759	961	1,206
# Featured User	579	5,149	11,725	19,418
# Follow-Link	59,073	529,448	158,746	915,438
# Reply-Link	1,451	96,757	121,133	530,598
# Retweet-Link	19,760	311,359	595,030	1,684,023
# Like-Link	14,381	302,571	562,496	1,794,111
# Mention-Link	27,682	353,586	539,580	1,571,937

Political-Centered Social Network Dataset

Literature Review:

- GCN (Graph Convolutional Network)
  `http://web.cs.ucla.edu/~patricia.xiao/files/Reading_Group_20181204.pdf`
- MTL (Multi-Task Learning)
  `http://web.cs.ucla.edu/~patricia.xiao/files/Reading_Group_20200121.pdf`

Contribution

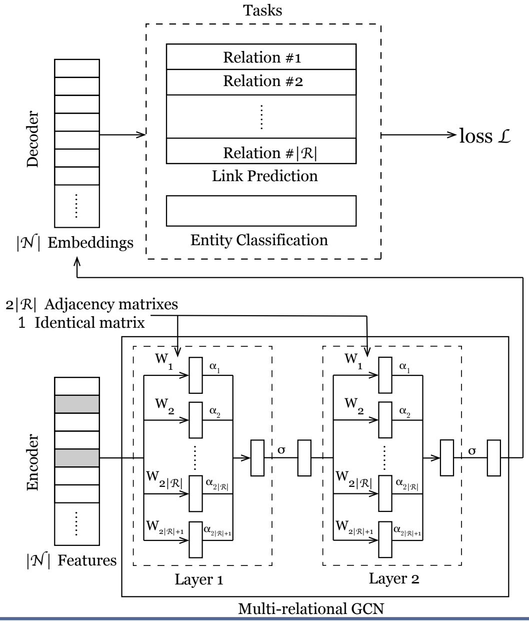
- Political-Centered Social Network Dataset
  - Described in Appendix, released with code

Contribution

- Political-Centered Social Network Dataset
  - Described in Appendix, released with code
- TIMME: learning embeddings on sparsely-labeled heterogeneous graph
  - MTL: handles the sparsity of labels

- Political-Centered Social Network Dataset
  - Described in Appendix, released with code
- TIMME: learning embeddings on sparsely-labeled heterogeneous graph
  - MTL: handles the sparsity of labels
  - *Optionally* handles incomplete input features

Model


Tasks
Relation #1
Relation #2
Relation #|ℛ|
Link Prediction
Entity Classification
loss ℒ
Decoder
|𝒩| Embeddings
2|ℛ| Adjacency matrixes
1 Identical matrix
Encoder
|𝒩| Features
W_1
W_2
$W_{2|\mathcal{R}|}$
$W_{2|\mathcal{R}|+1}$
α_1
α_2
$\alpha_{2|\mathcal{R}|}$
$\alpha_{2|\mathcal{R}|+1}$
σ
Layer 1
Layer 2
Multi-relational GCN

In homogeneous GCN layers:

$$H^{(l+1)} = \sigma\left(\hat{A} H^{(l)} W^{(l)}\right)$$

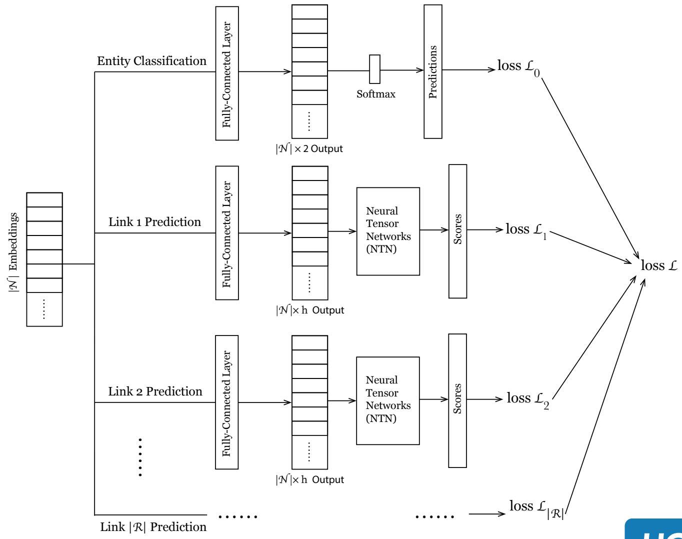
$\hat{A}$: the normalized adjacency matrix; $H^{(l+1)}$: layer-l output; $H^{(l)}$: layer-l input; $W^{(l)}$: layer-l weight.

In our heterogeneous design:

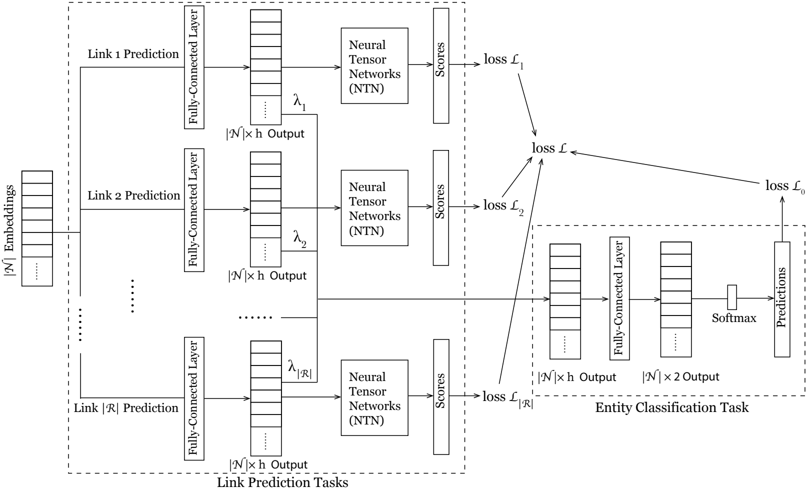
$$H^{(l+1)} = \sigma\left(\sum_{r \in \hat{R}} \alpha_r \hat{A}_r H^{(l)} W_r^{(l)}\right)$$

α_r: attention weight; $|\hat{R}| = 2R + 1$ includes R relations, R reversed relations, 1 identical matrix I.

Model: **TIMME** Decoder



Model: **TIMME-hierarchical** Decoder



TIMME and **TIMME-hierarchical**:

$$\mathcal{L} = \sum_{i=0}^{|\mathcal{R}|} \mathcal{L}_i$$

loss is the sum of losses from all $|\mathcal{R}| + 1$ tasks.

- **TIMME-hierarchical** gives us clues on each relation's importance to ideology classification via λ.

TIMME-single:

$$\mathcal{L} = \mathcal{L}_i$$

for a single task i, $i \in \{0, 1, 2, \ldots |\mathcal{R}|\}$.

- **TIMME-single** proves that multi-task version is better.

Evaluation Metrics

Evaluate on task level:

- Node-Classification Task:
  - Accuracy
  - F1-Score
- Link-Prediction Task:
  - ROC-AUC (Receiver Operating Characteristic Area Under Curve)
  - PR-AUC (Precision-Recall Area Under Curve)

Case study:

- State-level
- County-level
- Account-level

Ablation study

Results Compared with Baselines

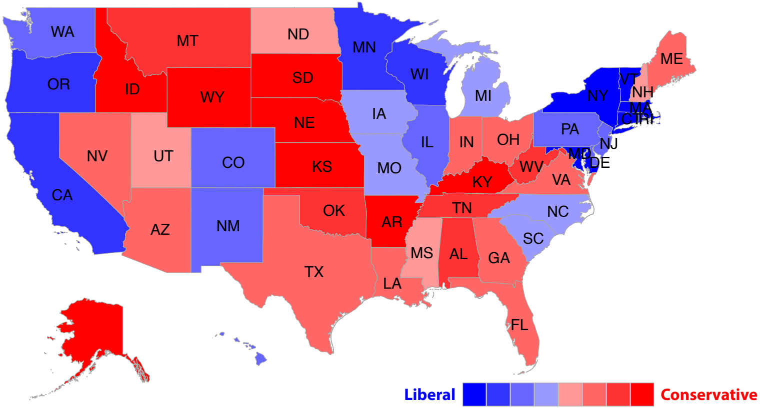
Model	PureP	P50	P20~50	P+all
GCN	**1.0000/1.0000**	0.9600/0.9600	0.9895/0.9895	0.9076/0.9083
r-GCN	**1.0000/1.0000**	0.9733/0.9733	0.9895/0.9895	0.9327/0.9333
HAN	0.9825/0.9824	0.9466/0.9467	0.9789/0.9789	0.9238/0.9250
TIMME-single	**1.0000/1.0000**	0.9733/0.9733	0.9895/0.9895	0.9333/0.9324
TIMME	0.9825/0.9824	**0.9867/0.9867**	**1.0000/1.0000**	0.9495/0.9500
TIMME-hierarchical	**1.0000/1.0000**	0.9733/0.9780	0.9895/0.9895	**0.9580/0.9583**

Table 2: Node classification measured by F1-score/accuracy.

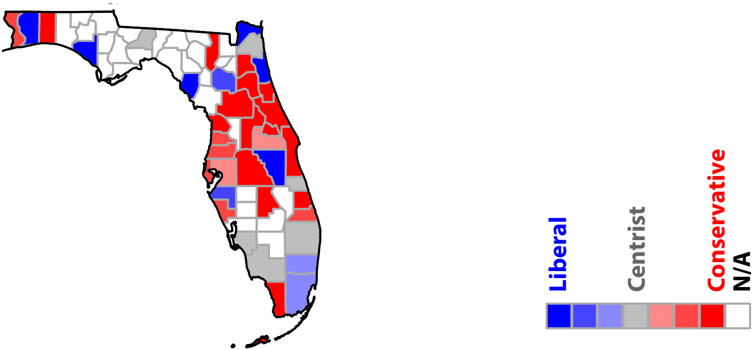
Model	PureP	P50	P20~50	P+all
		Follow Relation		
GCN+	0.8696/0.6167	0.9593/0.8308	0.9870/0.9576	0.9855/0.9329
r-GCN	0.8596/0.6091	0.9488/0.8023	0.9872/0.9537	0.9685/0.9201
HAN+	**0.8891/0.7267**	0.9598/0.8642	0.9620/0.8850	0.9723/0.9256
TIMME-single	0.8809/0.6325	0.9717/0.8792	0.9920/0.9709	0.9936/0.9696
TIMME	0.8763/0.6324	**0.9811/0.9154**	0.9945/0.9799	0.9943/0.9736
TIMME-hierarchical	0.8812/0.6409	0.9809/0.9145	**0.9984/0.9813**	**0.9944/0.9739**
		Reply Relation		
GCN+	0.8602/0.7306	0.9625/0.9022	0.9381/0.8665	0.9705/0.9154
r-GCN	0.7962/0.6279	0.9421/0.8714	0.8868/0.7815	0.9640/0.9085
HAN+	0.8445/0.6359	0.9598/0.8616	0.9495/0.8664	0.9757/0.9210
TIMME-single	0.8685/0.7018	0.9695/0.9307	0.9593/0.9070	0.9775/0.9508
TIMME	0.9077/0.8004	**0.9781/0.9417**	**0.9747/0.9347**	0.9849/0.9612
TIMME-hierarchical	**0.9224/0.8152**	0.9766/0.9409	0.9737/0.9341	**0.9854/0.9629**
		Retweet Relation		
GCN+	0.8955/0.7145	0.9574/0.8493	0.9351/0.8408	0.9724/0.9303
r-GCN	0.8865/0.6895	0.9411/0.8084	0.9063/0.7728	0.9735/0.9326
HAN+	0.7646/0.6139	0.9658/0.9213	0.9478/0.8962	0.9750/0.9424
TIMME-single	0.9015/ 0.7202	0.9754/0.9127	0.9673/0.9073	0.9824/0.9424
TIMME	0.9094/0.7285	0.9779/0.9181	**0.9772/0.9291**	0.9858/0.9511
TIMME-hierarchical	**0.9105/0.7344**	**0.9780/0.9190**	0.9766/0.9275	**0.9869/0.9543**
		Like Relation		
GCN+	0.9007/0.7259	0.9527/0.8499	0.9349/0.8400	0.9690/0.9032
r-GCN	0.8924/0.7161	0.9343/0.7966	0.9038/0.7681	0.9510/0.8945
HAN+	0.8606/0.6176	0.9733/0.8851	0.9611/0.9062	**0.9894**/0.9481
TIMME-single	0.9113/0.7654	0.9725/0.9119	0.9655/0.9069	0.9796/0.9374
TIMME	0.9249/0.7926	**0.9753**/0.9171	**0.9759/0.9292**	0.9846/0.9504
TIMME-hierarchical	**0.9278/0.7945**	0.9752/**0.9175**	0.9752/0.9271	0.9851/**0.9518**
		Mention Relation		
GCN+	0.8480/0.6233	0.9602/0.8617	0.9261/0.8170	0.9665/0.8910
r-GCN	0.8312/0.6023	0.9382/0.7963	0.8938/0.7563	0.9640/0.8902
HAN+	**0.9000/0.7206**	0.9573/0.8616	0.9574/0.8891	0.9724/0.9119
TIMME-single	0.8587/0.6502	0.9713/0.8981	0.9614/0.8923	0.9725/0.9096
TIMME	0.8684/0.6689	0.9730/0.9035	**0.9730/0.9185**	0.9839/0.9446
TIMME-hierarchical	0.8643/0.6597	**0.9732/0.9046**	0.9723/0.9166	**0.9846/0.9463**

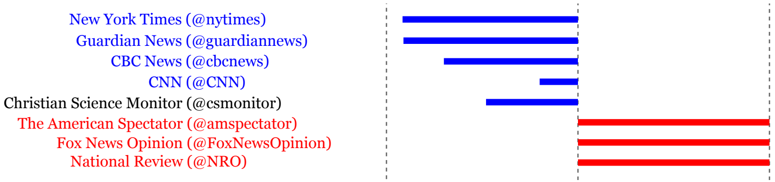
Table 3: Link-prediction measured by ROC-AUC/PR-AUC.

State-Level Ideology on Twitter

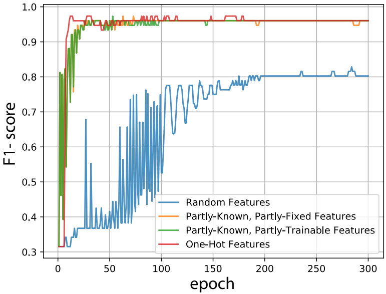


County-Level Ideology on Twitter: Florida

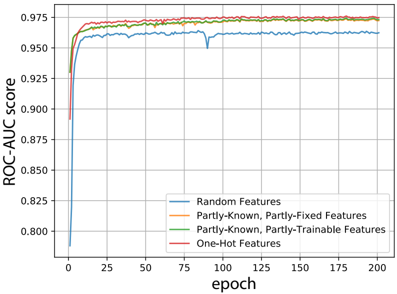



New York Times (@nytimes)
Guardian News (@guardiannews)
CBC News (@cbcnews)
CNN (@CNN)
Christian Science Monitor (@csmonitor)
The American Spectator (@amspectator)
Fox News Opinion (@FoxNewsOpinion)
National Review (@NRO)

Ablation study: Features

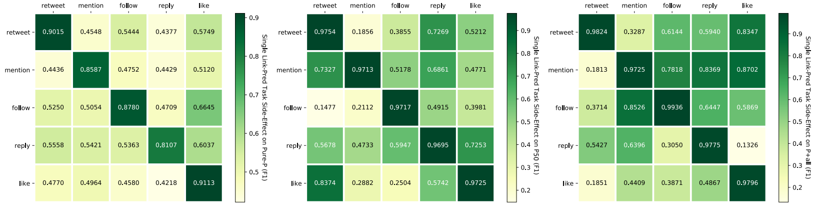


(a) Features' Impact on Node-Classification Task

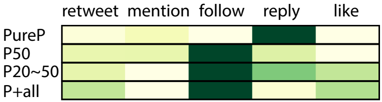


(b) Features' Impact on "Follow" Link-Prediction Task

Ablation study: Relations' Relations



TIMME-single on Pure-P (left), P50 (middle), P+all (right)



TIMME-hierarchical model's λ values learned in the end.

Resources

Resources

Code with data available on Github:

- https://github.com/PatriciaXiao/TIMME

Presentation etc. on my personal website:

- http://web.cs.ucla.edu/~patricia.xiao/timme.html

Special thanks to: Haoran Wang, Zhiwen Hu, Yupeng Gu