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

Motivation

Contribution

Data

Model

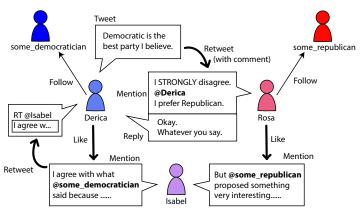
Results at a Glance

Code & Data

TIMME



Motivation



Problem: Ideology Classification on Twitter

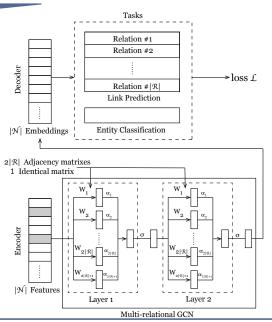


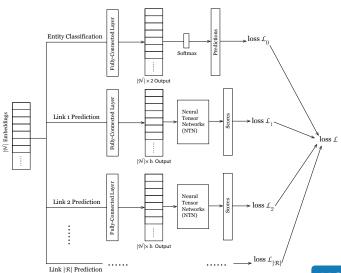
- ➤ TIMME: learning embeddings on sparsely-labeled heterogeneous graph
 - ► MTL: handles the sparsity of labels
 - ► Optionally handles incomplete input features
- ▶ Political-Centered Social Network Dataset
 - ▶ Described in Appendix, released with code

	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

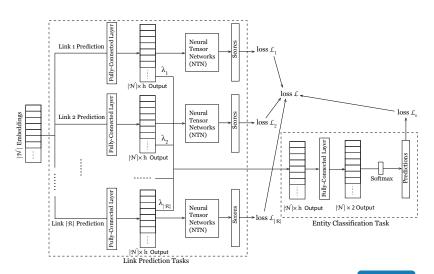
Political-Centered Social Network Dataset







UCLA





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 $i, i \in \{0, 1, 2, \dots |\mathcal{R}|\}.$

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

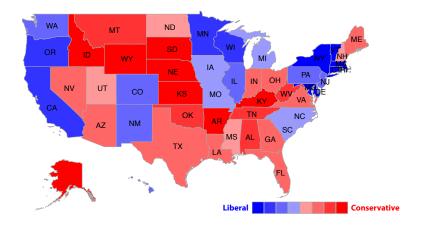
UCLA

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	
7.11 a N 1 1 10 11 11 74 /					

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

CCV. ASSESSALES (1998) SCRIPTION (1998) CMSTREAD (1998)<	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		
TOMME 07500.00 08110.00 08110.00 08110.00 08110.00 0810.00	HAN+	0.8891/0.7267	0.9598/0.8642	0.9620/0.8850	0.9723/0.9256		
CCN ASSESTANCE No-RESIDENCE C-975(10) <t< td=""><td>TIMME-hierarchical</td><td>0.8812/0.6409</td><td>0.9809/0.9145</td><td>0.9984/0.9813</td><td>0.9944/0.9739</td></t<>	TIMME-hierarchical	0.8812/0.6409	0.9809/0.9145	0.9984/0.9813	0.9944/0.9739		
		1	Reply Relation				
MMM-lender MM							
	HAN+	0.8445/0.6359	0.9598/0.8616	0.9495/0.8664	0.9757/0.9210		
	TIMME-hierarchical	0.9224/0.8152	0.9766/0.9409	0.9737/0.9341	0.9854/0.9629		
PORT		R	etweet Relation				
		0.8955/0.7145	0.9574/0.8493	0.9351/0.8408	0.9724/0.9303		
TOMME 09940723 0.97990781 0.97200929 0.98800930 MUMD-Harmach on 1906 07514 1.987004091 0.98800930 0.98800930 CN 1.986707372 1.9870144 0.989010492 0.98900930 FCN 0.986907472 0.951010492 0.98900940 0.99100940 0.99100940 FCN 0.986907472 0.951010380 0.98000940 0.99100940 0.99100940 0.99100940 TIMMS-engle 0.940907290 0.97530.911 0.96530.902 0.99800940 0.99800940 0.98800940 TIMMS-increased 0.84000221 0.98000200 0.98000200 0.98100200000 0.98100200 0.98100200 <td< td=""><td>HAN+</td><td>0.7646/0.6139</td><td>0.9658/0.9213</td><td>0.9478/0.8962</td><td>0.9750/0.9424</td></td<>	HAN+	0.7646/0.6139	0.9658/0.9213	0.9478/0.8962	0.9750/0.9424		
TMME-limethouse 0.916/0.714	TIMME-single	0.9015/ 0.7202	0.9754/0.9127	0.9673/0.9073	0.9824/0.9424		
CCN	TIMME-hierarchical	0.9105/0.7344	0.9780/0.9190	0.9766/0.9275	0.9869/0.9543		
	Like Relation						
HAN-	GCN+	0.9007/0.7259	0.9527/0.8499	0.9349/0.8400	0.9690/0.9032		
TMMS		0.8924/0.7161	0.9343/0.7966	0.9038/0.7681	0.9510/0.8945		
TMME 024907.92 97530.912 03790.9220 9890.0930 MCM 29278.07935 79720.9175 97920.922 9890.0940 New Transport CCN 084800.6231 96920.8310 20820.01940 98980.01540 96980.0180 IAN 3,9900.22 98720.01940 983980.01540 97720.0180 97720.0180 IAN 4,9900.22 98720.0180 98720.0180 97720.0181 98720.0180 98780.0180 <td>HAN+</td> <td>0.8606/0.6176</td> <td>0.9733/0.8851</td> <td>0.9611/0.9062</td> <td>0.9894/0.9481</td>	HAN+	0.8606/0.6176	0.9733/0.8851	0.9611/0.9062	0.9894/0.9481		
TMME-limber 19270/81 97520/917 97520/927 975					0.9796/0.9374		
CCN-							
CCN. 0.8801.0233 0.8020.0237 0.821.02473 0.8020.0257 CCN. 0.8112.02603 0.8902.0788 0.9802.0788 0.9802.0788 0.9802.0788 HAN- 0.9900.0226 0.95710.8016 0.95740.8891 0.97240.9819 TIMME-incl 0.88470.6862 0.9713.08881 0.9840.0838 0.9850.0989 TIMME-incl 0.8840.08697 0.97130.09816 0.97100.0916 0.9860.0848 TIMME-incl 0.8840.08678 0.97230.09016 0.97100.0916 0.9860.0848	TIMME-hierarchical	0.9278/0.7945	0.9752/ 0.9175	0.9752/0.9271	0.9851/ 0.9518		
FCD 08312/04221 0.9882/07528 0.8983/07528 0.964002.8992 HMNH-single 0.9501/0.0510 0.9571/0.0510 0.9674/0.8910 0.9724/0.0710 TIMME-single 0.8587/0.6502 0.9713/0.0891 0.9614/0.8923 0.9725/0.0906 TIMME-single 0.8483/0.0879 0.9724/0.0916 0.9728/0.0916 0.9738/0.0916 MOME-particular 0.8483/0.0879 0.9724/0.0916 0.9728/0.0916 0.9728/0.0916	Mention Relation						
HAN* 0.9900,7206 0.9573/0.861 0.9574/0.891 0.9724/0.919 TIMME-single TIMME 0.85870/6.902 0.9713/0.8981 0.9614/0.8923 0.9725/0.096 0.8584/0.669 0.9730/0.996 0.9732/0.996 0.9723/0.916 0.9846/0.9464 TIMME-hierarchical 0.8643/0.6597 0.9732/0.9966 0.9723/0.916 0.9846/0.9464	GCN+	0.8480/0.6233	0.9602/0.8617	0.9261/0.8170	0.9665/0.8910		
TIMME-single 0.8587/0.5502 0.9713/0.8981 0.9614/0.8923 0.9725/0.9096 TIMME 0.8584/0.6699 0.9730/0.9035 0.9730/0.9185 0.9839/0.9446 TIMME-hierarchical 0.8643/0.6597 0.9732/0.9046 0.9723/0.9165 0.9846/0.9463							
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	HAN+	0.9000/0.7206	0.9573/0.8616	0.9574/0.8891	0.9724/0.9119		
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	TIMME-single	0.8587/0.6502	0.9713/0.8981	0.9614/0.8923	0.9725/0.9096		
	TIMME		0.9730/0.9035				
Table 3: Link-prediction measured by ROC-AUC/PR-AUC.	TIMME-hierarchical	0.8643/0.6597	0.9732/0.9046	0.9723/0.9166	0.9846/0.9463		
	Table 3: Link-	prediction :	measured b	y ROC-AUG	C/PR-AUC.		





Code & Data



Code with data available on Github:

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

All information included in readme.

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