A Progressive Framework for Role-Aware Rumor Resolution

Lei Chen[1], Guanying Li[2], Zhongyu Wei[1,3]*, Yang Yang[4], Baohua Zhou[2], Qi Zhang[5], Xuanjing Huang[5]

[1] School of Data Science, Fudan University, China
[2] School of Journalism, Fudan University, China
[3] Research Institute of Intelligent and Complex Systems, Fudan University, China
[4] College of Computer Science and Technology, Zhejiang University, China
[5] School of Computer Science, Fudan University, China
[1,3,5]{chenl18,zywei,qi_zhang,xjhuang}@fudan.edu.cn
[2]zhoubaohua@yeah.net;[4]yangya@zju.edu.cn

COLING2022

https://github.com/lchen96/trigger_identification

Reported by Xiaoke Li

Introduction

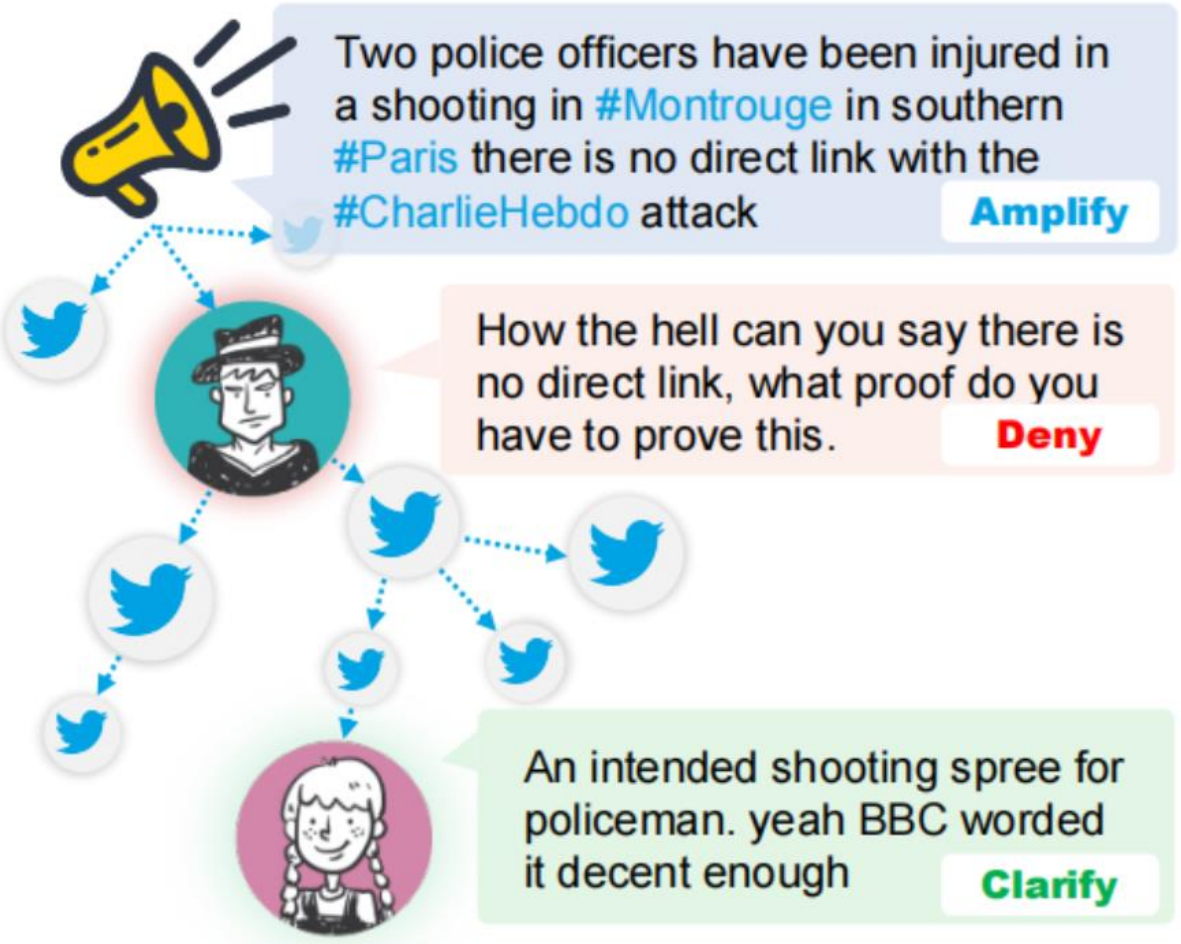


Figure 1: An illustration of rumor cascades and typical roles of messages helpful for rumor verification.

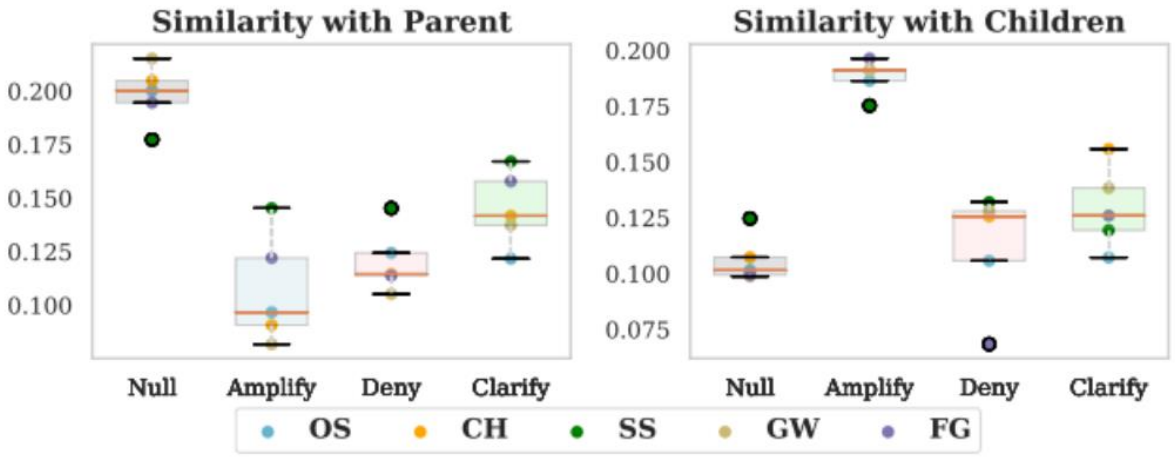


Figure 2: Context similarity for different types of triggers in different events. Scatter points represent the averaged context similarity for a certain kind of trigger in a specific event. Shapes of boxes depict the degree of trigger assimilation for different events.

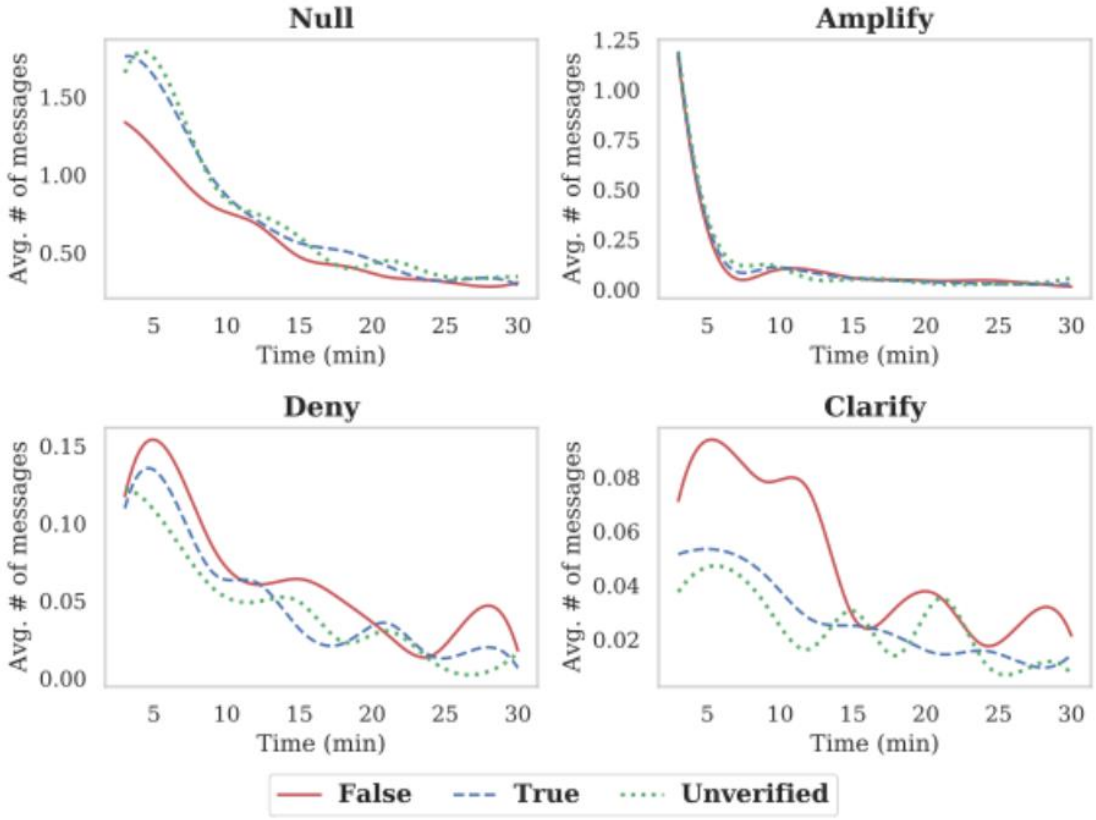


Figure 3: Temporal variation of trigger distribution. Each subgraph represents a certain kind of trigger. Different line styles stand for the category of rumors.

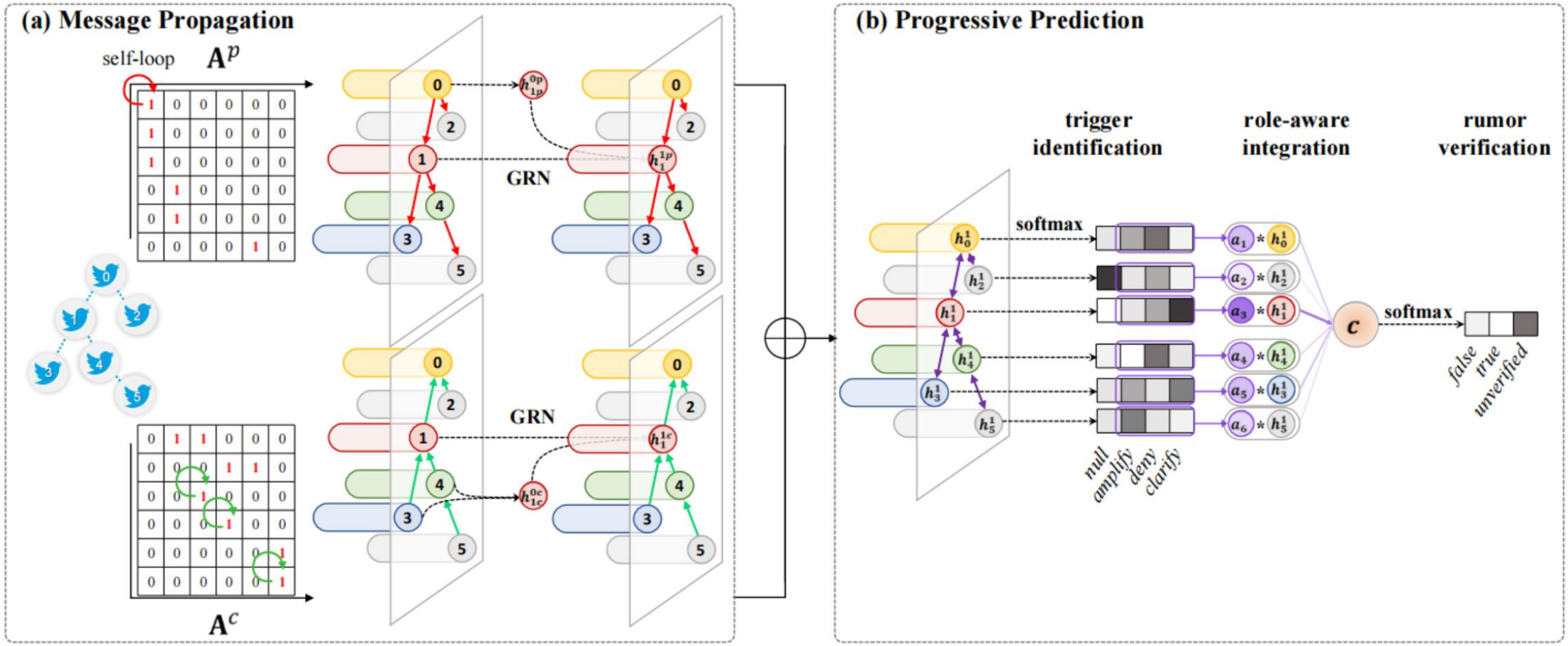
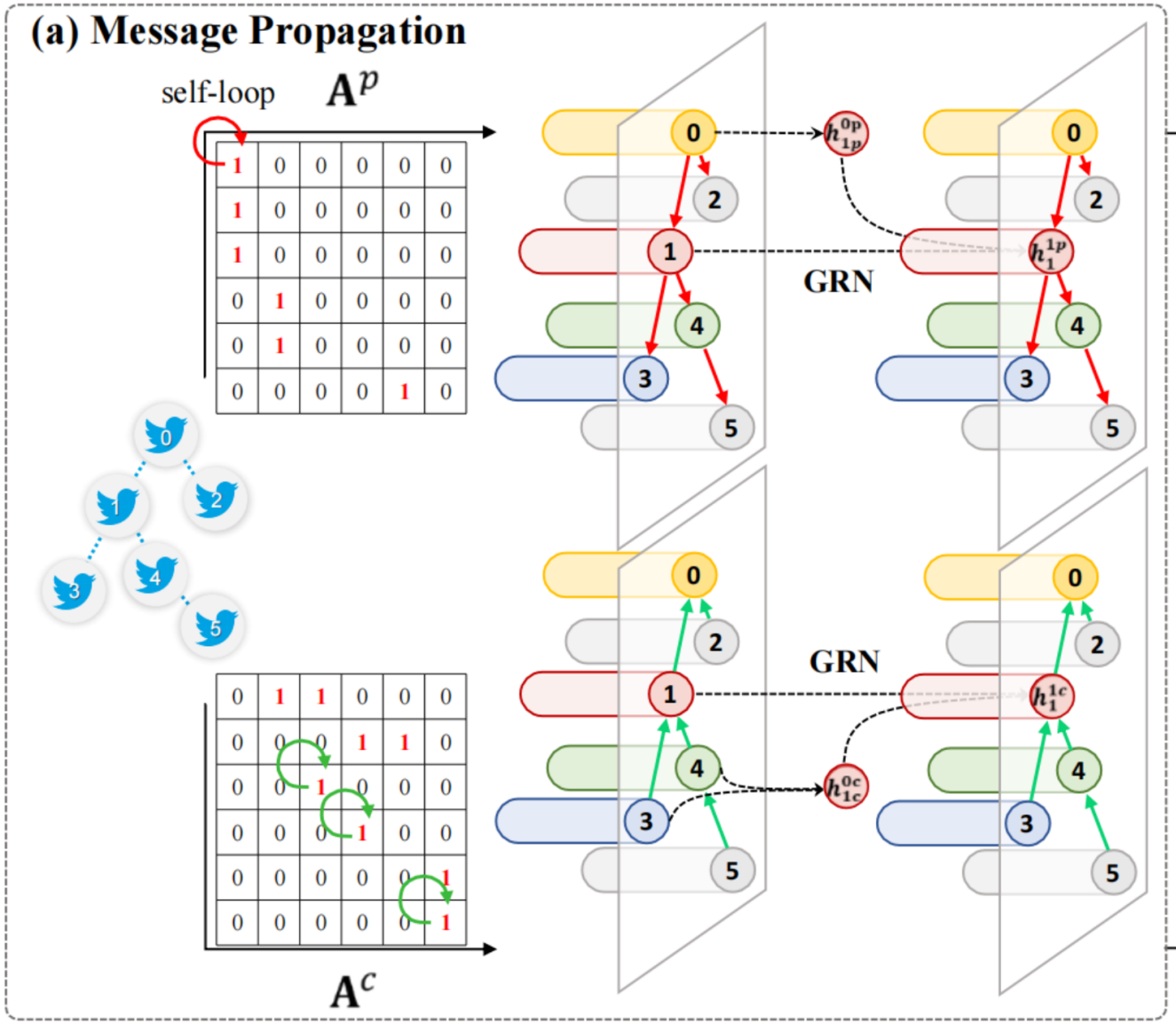


Figure 4: Overall architecture of our proposed model. The two squares on the left represent the decomposed adjacency matrices $\mathbf{A}^p$ and $\mathbf{A}^c$ that control the direction of information flow. Both of the two tasks share the unsymmetric GRN Layers. The updated node representation is used to predict trigger labels. Role-aware integration mechanism is then applied to acquire cascade representation and produce verification prediction.



$$\hat{\mathbf{A}}^p = (\mathbf{D}^p)^{-1}\mathbf{A}^p \quad \hat{\mathbf{A}}^c = (\mathbf{D}^c)^{-1}\mathbf{A}^c$$

$$\boldsymbol{h}_{i,p}^{(l-1),p} = \sum_{j \in \{j|\hat{\mathbf{A}}_{ij}^p \neq 0\}} \hat{\mathbf{A}}_{ij}^p \boldsymbol{h}_j^{(l-1),p} \tag{1}$$

$$\boldsymbol{f}_i^l = \sigma_g\left(\boldsymbol{W}_f \boldsymbol{h}_i^{(l-1),p} + U_f \boldsymbol{h}_{i,p}^{(l-1),p} + \boldsymbol{b}_f\right) \tag{2}$$

$$\boldsymbol{i}_i^l = \sigma_g\left(\boldsymbol{W}_i \boldsymbol{h}_i^{(l-1),p} + \boldsymbol{U}_i \boldsymbol{h}_{i,p}^{(l-1),p} + \boldsymbol{b}_i\right) \tag{3}$$

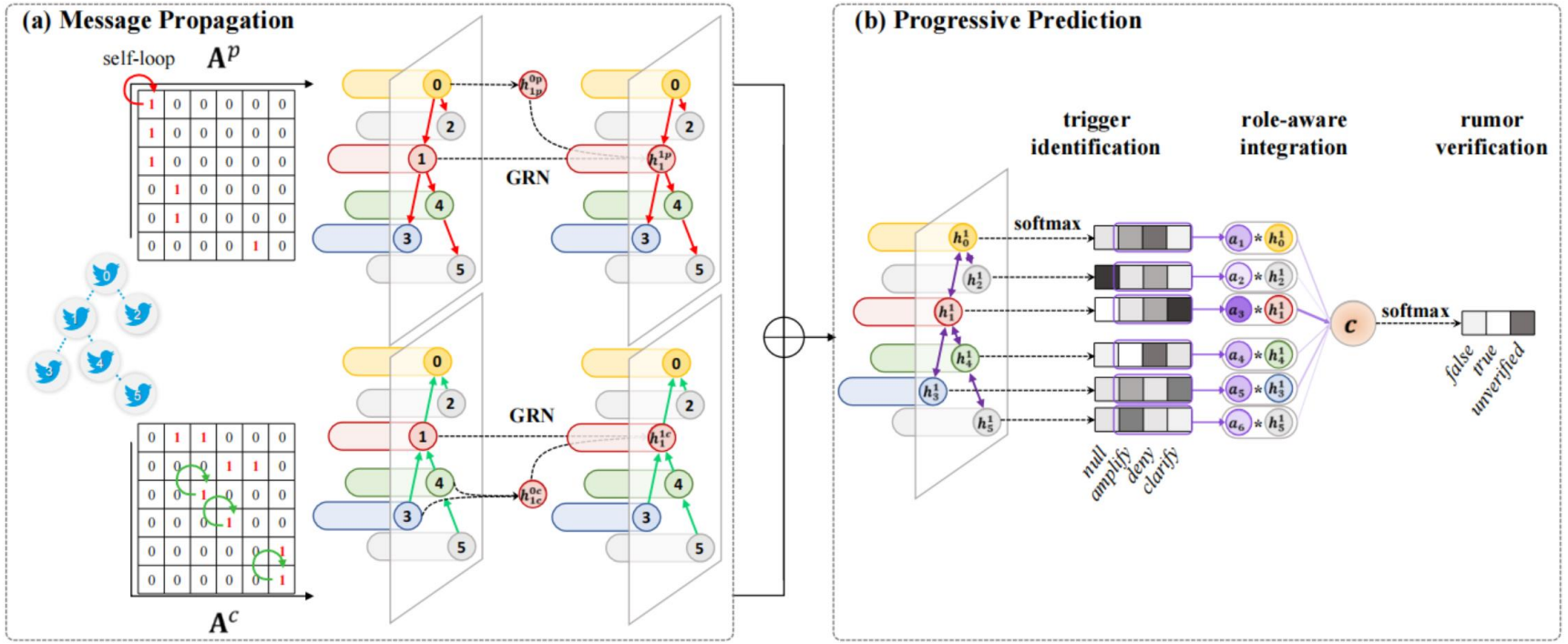
$$\boldsymbol{o}_i^l = \sigma_g\left(\boldsymbol{W}_o \boldsymbol{h}_i^{(l-1),p} + \boldsymbol{U}_o \boldsymbol{h}_{i,p}^{(l-1),p} + \boldsymbol{b}_o\right) \tag{4}$$

$$\tilde{\boldsymbol{c}}_i^l = \sigma_c\left(\boldsymbol{W}_c \boldsymbol{h}_i^{(l-1),p} + \boldsymbol{U}_c \boldsymbol{h}_{i,p}^{(l-1),p} + \boldsymbol{b}_c\right) \tag{5}$$

$$\boldsymbol{c}_i^l = \boldsymbol{f}_i^l \circ \boldsymbol{c}_i^{(l-1)} + \boldsymbol{i}_i^l \circ \tilde{\boldsymbol{c}}_i^l \tag{6}$$

$$\boldsymbol{h}_i^{l,p} = \boldsymbol{o}_i^l \circ \sigma_c\left(\boldsymbol{c}_i^l\right) \tag{7}$$

Method



$$\mathcal{Y}_i^t = \text{softmax}(\text{FFN}(\boldsymbol{h}_i^1)) \in \mathbb{R}^4 \quad (8)$$

$$\mathcal{L}_t = -\frac{1}{|\mathcal{V}|}\sum_{i}^{|\mathcal{V}|}\sum_{j}^{L_t}\mathcal{Y}_i^{t,j}\log\hat{\mathcal{Y}}_i^{t,j} \quad (9)$$

$$a_i = \begin{bmatrix}0 & 1 & 1 & 1\end{bmatrix}\mathcal{Y}_i^t \quad (10)$$

$$\boldsymbol{c} = \sum_{i}^{|\mathcal{V}|} a_i \boldsymbol{h}_i^1 \quad (11)$$

$$\mathcal{Y}^v = \text{softmax}(\text{FFN}(\boldsymbol{c})) \in \mathbb{R}^3 \quad (12)$$

$$\mathcal{L}_v = -\sum_{j}^{L_v}\mathcal{Y}^{v,j}\log\hat{\mathcal{Y}}^{v,j} \quad (13)$$

$$\mathcal{L} = \mathcal{L}_t + \mathcal{L}_v \quad (14)$$

Experiments

event	# of cas.	# of mes.	verify dist. (F:T:U)	trigger dist. (N:A:C:D)
CH	449	6110	114:187:148	4705:915:271:219
OS	467	6036	72:327:68	4793:868:254:121
SS	508	7832	76:378:54	5868:1050:471:443
FG	268	4516	8:9:251	3679:527:181:129
GW	237	2377	111:94:32	1762:388:147:80
All	1929	26871	381:995:553	20807:3748:1324:992

Table 1: Statistics of extended PHEME dataset. The abbreviation of different events is in short of *Charlie Hebdo*, *Ottawa Shooting*, *Sydney Siege*, *Ferguson Unrest*, *Germanwings Crash* respectively. The next two columns represents the amount of cascades and messages involved in different events. As for distribution of verification and trigger labels, capital letters stand for possible categories (F: *false*, T: *true*, U: *unverified*, N: *null*, A: *amplify*, C: *clarify*, D: *deny*).

Method	Trigger		Verify	
	Random	LOEO	Random	LOEO
CNN	0.524	0.501	0.741	0.308
RNN	0.562	0.560	0.785	0.314
TreeLSTM	0.538	0.514	0.710	0.317
TreeTrans	0.541	0.511	0.714	0.314
GCN	0.548	0.542	0.772	0.322
GraphSage	0.549	0.561	0.781	0.304
UGRN	**0.574**	**0.570**	**0.819**	**0.346**

Table 2: Results of trigger identification and rumor verification. All the numerical values represent macro F1-score when adopting random or LOEO cross validation. The result of LOEO validation is the average of 5 folds. **Bold**: the best performance in each column.

Experiments

Component	Trigger		Verify	
	Random	LOEO	Random	LOEO
GRN	0.531	0.514	0.754	0.324
UGRN-c	0.541	0.522	0.768	0.321
UGRN-p	0.552	0.541	0.778	0.334
UGRN	**0.574**	**0.570**	**0.819**	**0.346**

Table 3: Ablation study on key components of UGRN. Presentation of result is the same with Table 2.

Experiments

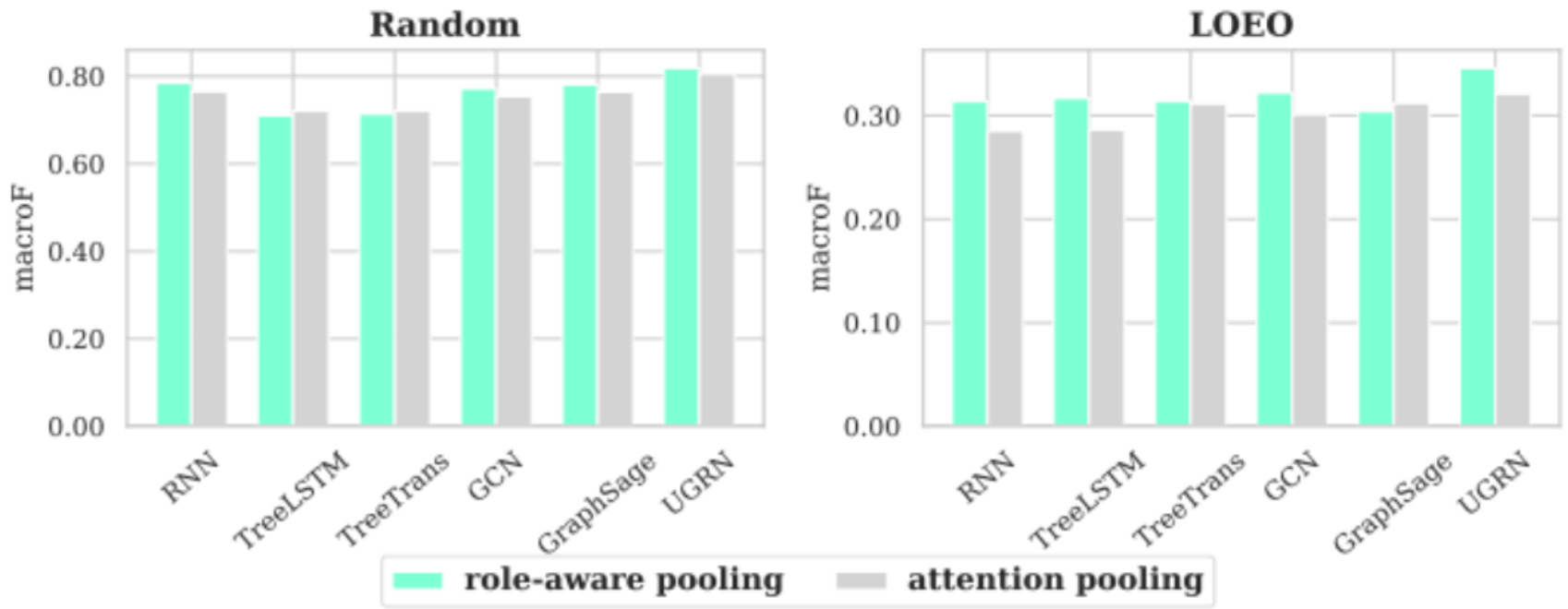


Figure 5: The effect of role-aware integration.

Experiments

Task	Trigger		Verify	
	Random	LOEO	Random	LOEO
Trigger	0.568	0.558	-	-
Verify	-	-	0.795	0.286
Multi-Task	**0.574**	**0.570**	**0.819**	**0.346**

Table 4: The effect of multi-task learning framework. Presentation of result is the same with Table 2.

Experiments

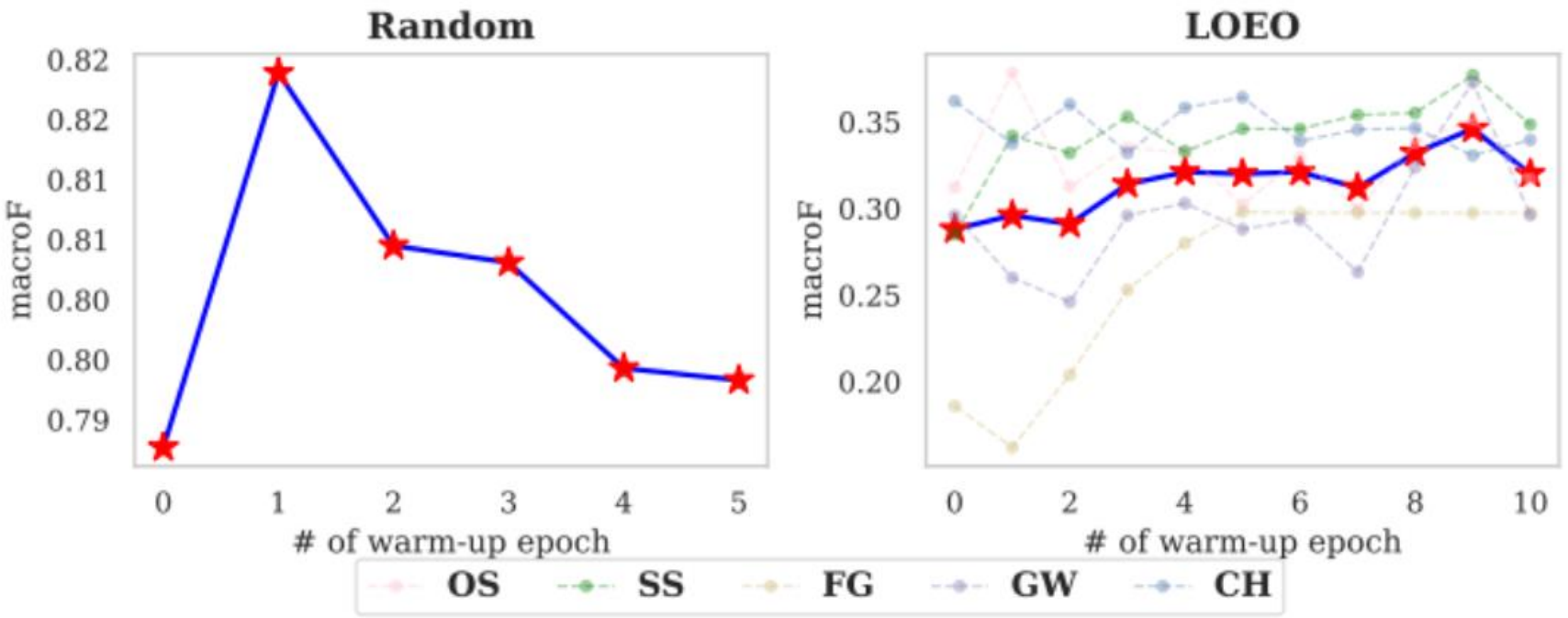


Figure 6: The effect of trigger warm-up strategy. Blue solid lines represent the averaged result and the dashed lines stand for results with different test event.

Thanks