Context-Aware Event Forecasting via Graph Disentanglement

Yunshan Ma[1*], Chenchen Ye[1*], Zijian Wu[1], Xiang Wang[2], Yixin Cao[3], Tat-Seng Chua[1]

* denotes equal contribution; 1. National University of Singapore; 2. University of Science and Technology of China; 3. Singapore Management University

09 August, 2023

Outline

- Background
- Motivation
- Proposed Method
- Experiments and Analysis
- Conclusion

Background

Why event forecasting?

Pandemic Outbreak

- **SARS-2003**: infected over 8,000 people and caused 774 deaths.
- **H1N1-2009**: low mortality rate than anticipated, but Singapore's overaction leads to large economic losses.
- **COVID-2019**: infected billions globally, resulting in millions of deaths.

What if…

- Pandemic severity is predicted to avoid loss of life or economic loss.
- Spread ability is estimated to help the government formulate lockdown policies in advance.

Civil Unrest

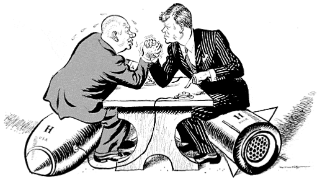
- **Arab Spring**: Protests and uprisings in the Middle East and North Africa, leading to regime changes.
- **Hong Kong Protests**: Pro-democracy demonstrations against a proposed extradition bill, leading to tensions between HK and mainland China.

What if…

- Riot signs are detected to implement policies before they escalate.
- Protest's magnitude is estimated to prepare businesses and infrastructure for potential disruptions.

International Conflict

- **Gulf War**: A conflict initiated by Iraq's invasion of Kuwait, leading to a US-led coalition intervention.
- **Arab-Israeli War**: A series of conflicts between Israel and neighboring Arab countries, primarily involving Egypt, Jordan and Syria.



What if…

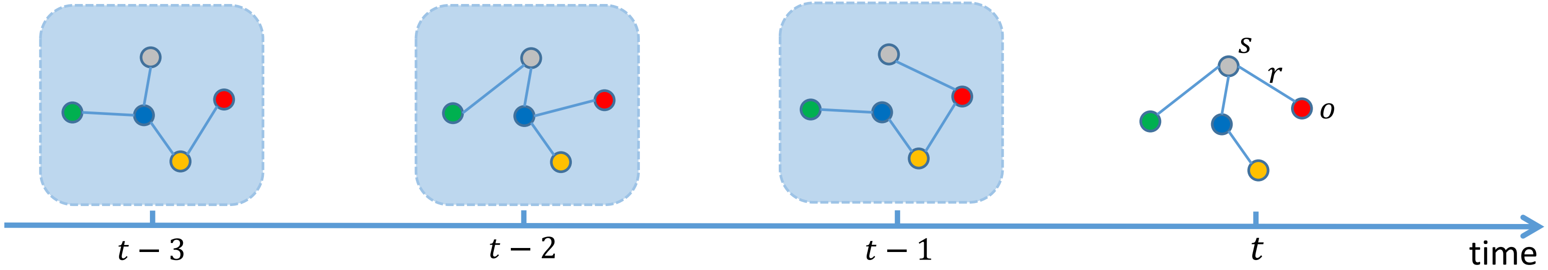
- Regions are identified to enable targeted interventions and monitoring.
- Preventive diplomacy is implemented to engage in early dialogue and mediation to avert conflicts.

Background

What is event forecasting?

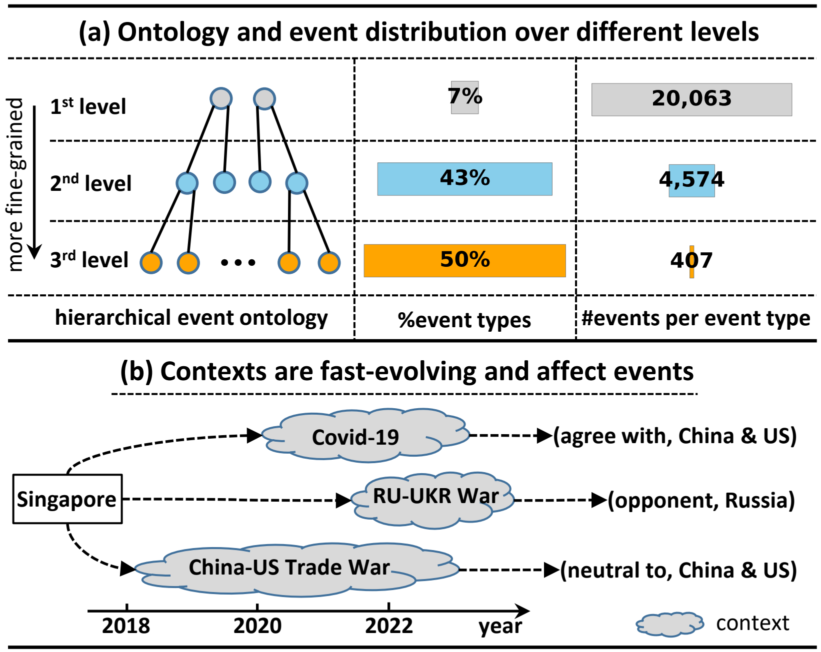
Task: Predict which entity will have a given relation together with a given entity at a certain future timestamp

Each event is represented as a quadruple (subject, event type, object, timestamp), all the events along the timeline demonstrate a a temporal knowledge graph.



Formal definition: Given query $(subject, event\ type, ?)$ at current timestamp t, and history graph sequence $\{G_{t-k}, \ldots, G_{t-2}, G_{t-1}\}$, predict $object$ related to this query.

Motivation

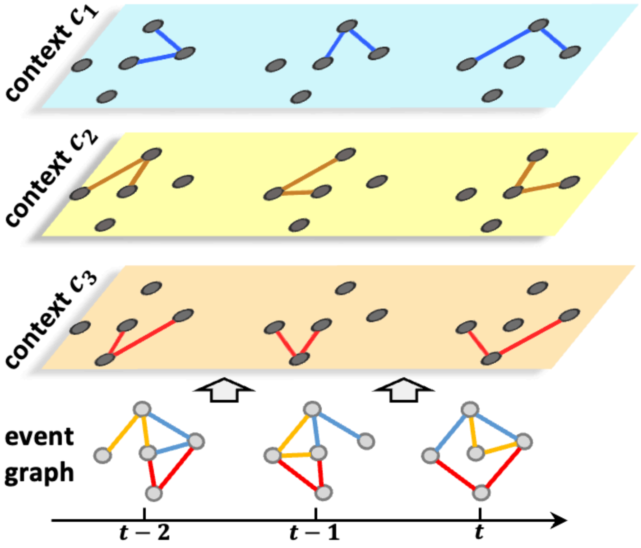


- Most current events fall in the coarse-grained and higher level types of the ontology, while more informative fine-grained events are fewer.
- Out-of-ontology and diverse contexts affect events. Context can provide more fine-grained information to enhance the event forecasting performance.

Motivation

Use context information to disentangle the event graph

Data in Context-aware Event Forecasting Tasks:



Context-aware Event Forecasting Methods:

Predict using Temporal Knowledge Graphs.

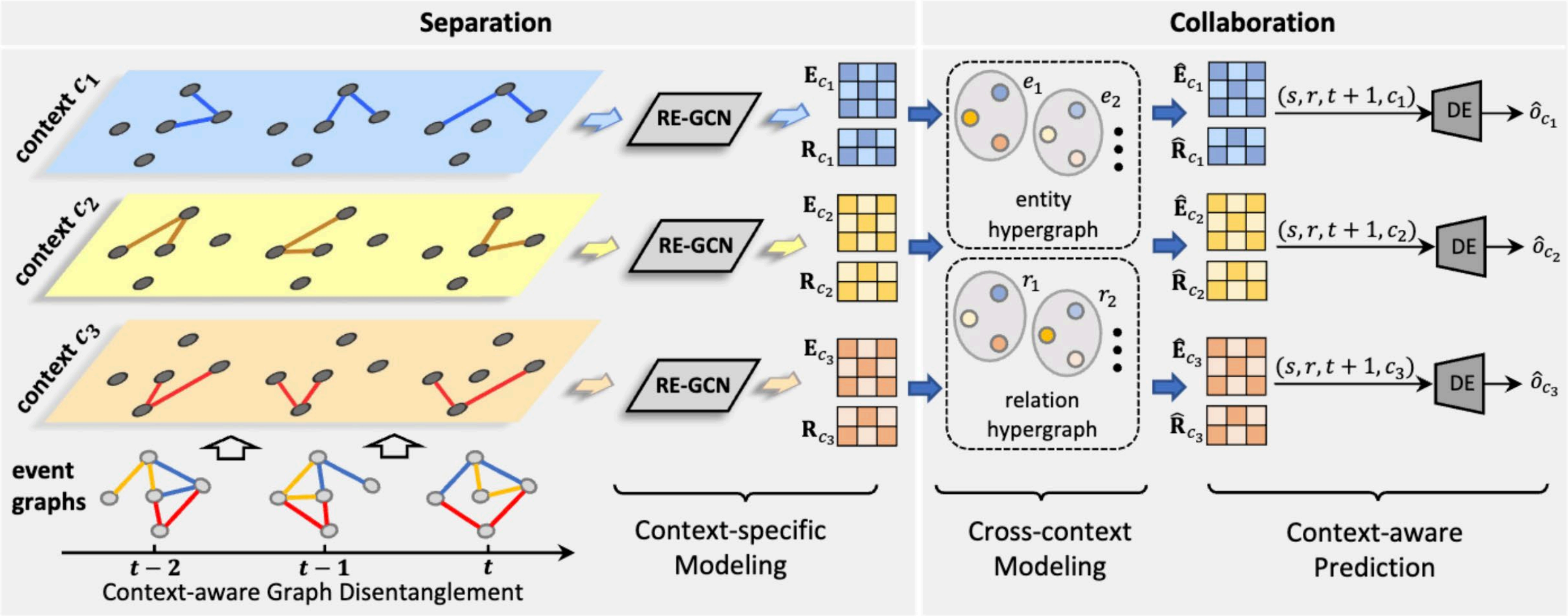
Enhance prediction with Associated Context Data.

New problem definition: Given query $(subject, relation, ?, \boldsymbol{context})$ at current timestamp t, and history graph sequence $\{G_{t-k}, \ldots, G_{t-2}, G_{t-1}\}$, predict $object$ related to this query.

Proposed Method: Overview

SeCoGD: Context-aware Event Forecasting via Graph Disentanglement

Motivation: Utilizing the out-of-ontology context to disentangle representations.



Proposed Method

Stage 1 - Separation

Use context as a prior guide to disentangle the event graph

Context-aware Graph Disentanglement:

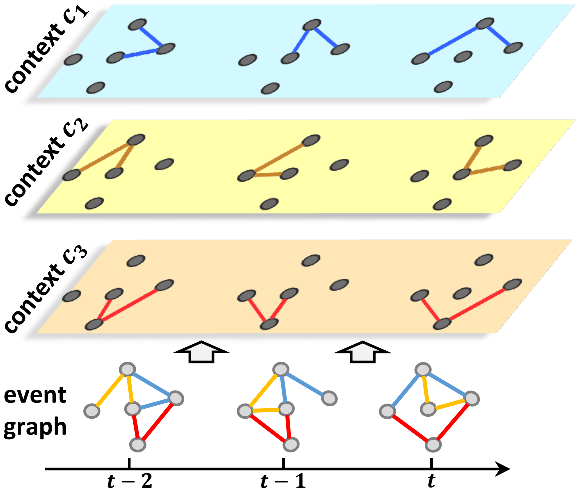
(a) Given K context topics, separate the original graph G_t into K subgraphs using LDA (Latent Dirichlet Allocation) topic model:

$$G_t \rightarrow \{G_t^{c_1}, \ldots, G_t^{c_k}, \ldots, G_t^{c_K}\}$$

(b) Each subgraph can be denoted as:

$$G_t^{c_k} = \{(s_n, r_n, o_n, c_k, t)\}_{n=1}^{N_t^{c_k}},$$

where $N_t^{c_k}$ is the number of events in timestamp t with in the context c_k.



Proposed Method

Stage 1 - Separation

Learn evolving patterns under corresponding context

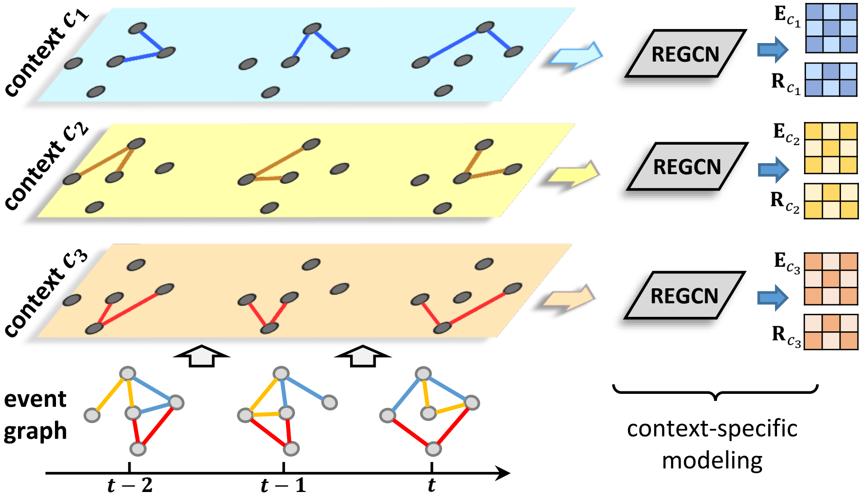
Context-specific Modeling:

(a) Concurrent event modeling:

$$\mathbf{e}_o^l = f\left(\frac{1}{|\mathcal{E}_o|}\sum_{(s,r)\in\mathcal{E}_o}\mathbf{W}_1^l(\mathbf{e}_s^{l-1}+\mathbf{r})+\mathbf{W}_2^l\mathbf{e}_o^{l-1}\right)$$

(b) Temporal pattern modeling:

$$\begin{cases} \mathbf{E}_{t,c} = \mathbf{U}_{t,c}\mathbf{E}'_{t,c} + (1-\mathbf{U}_{t,c})\mathbf{E}_{t-1,c} \\ \mathbf{U}_{t,c} = \sigma(\mathbf{W}_4\mathbf{E}_{t-1,c}+\mathbf{b}) \end{cases}$$



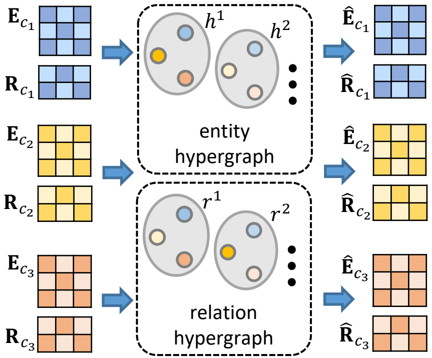
Proposed Method

Stage 2 - Collaboration

Even though the same entity demonstrates different characteristics in various contexts, these contexts are not independent but correlated with each other.

Cross context Modeling:

Hypergraph propagation:



Multi-layer message passing

$$\hat{\mathbf{e}}^{p}_{v,c} = \frac{1}{|C_v| - 1} \sum_{i \in C_v \setminus \{c\}} \hat{\mathbf{e}}^{p-1}_{v,i}$$

$$\hat{\mathbf{e}}_{v,c} = \sum_{p=0}^{P} \hat{\mathbf{e}}^{p}_{v,c},$$

C_v are all the contexts that the entity v has been in

Proposed Method

Stage 2 - Collaboration

For context-aware prediction, we utilize ConvTransE to capture cross-context knowledge:

$$\hat{\mathbf{p}}(\mathcal{E}|s,r,c,G_{\leq t}) = \text{softmax}\big(\hat{\mathbf{E}}_c \text{ConvTransE}(\hat{\mathbf{e}}_{s,c}, \hat{\mathbf{r}}_c)\big)$$

$\hat{\mathbf{e}}_{s,c}, \hat{\mathbf{r}}_c$ are the representation for query s and r

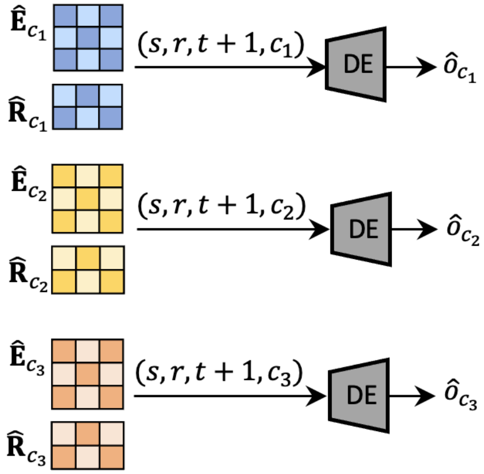
The predicted object is presented as:

$$\hat{o}_{(s,r,t+1,c)} = \arg\max_{\mathcal{E}} \hat{p}(\mathcal{E}|s,r,c,G_{\leq t})$$

We employ cross-entropy loss to optimize the whole framework in an end-to-end fashion:

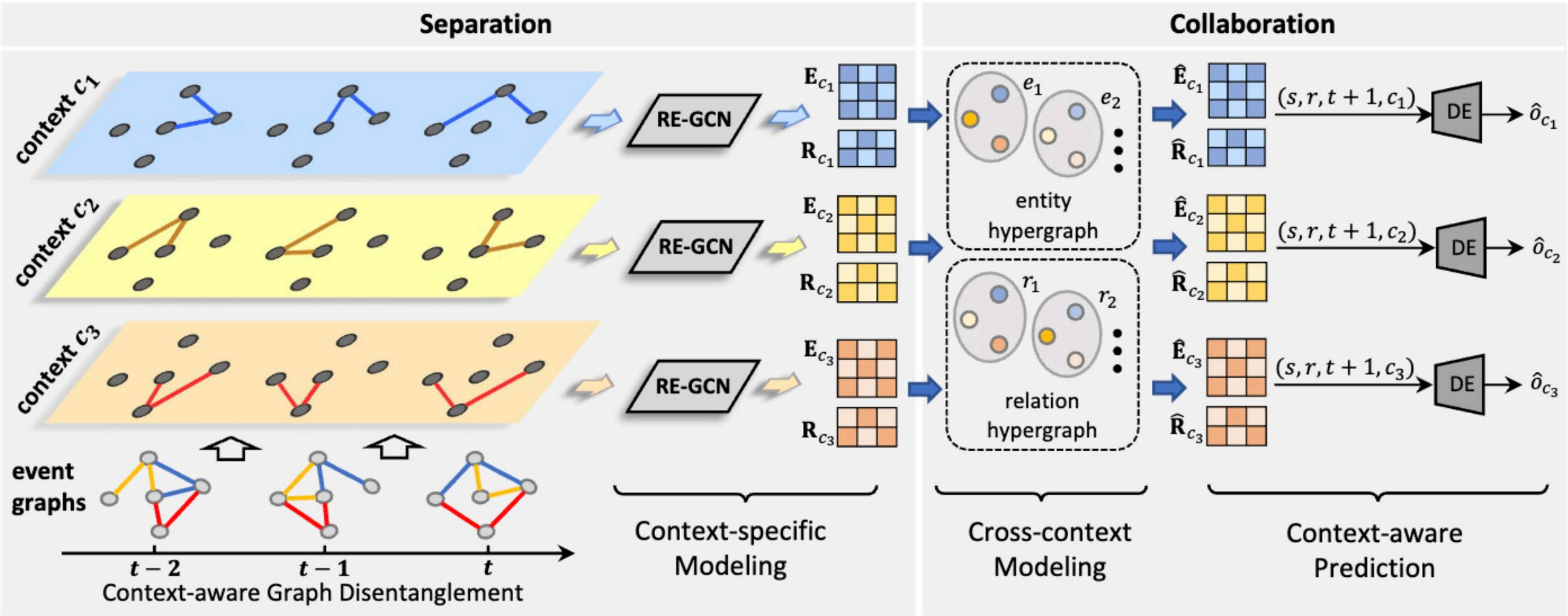
$$\mathcal{L} = \sum_{t=0}^{T-1} \sum_{c \in C} \sum_{(s,r) \in G_{t+1}^{c}} \mathbf{y}_{(s,r,t+1,c)} \log \hat{\mathbf{p}}(\mathcal{E}|s,r,c,G_{\leq t})$$

Where T is the total number of timestamps in the training set, and $y_{s,r,t+1,c}$ is the one-hot representation of ground-truth object o



Proposed Method: Overview

The overall framework of SeCoGD:



Experiments and Analysis

Datasets construction

- Crop three subsets of GDELT according to the regions of the events, i.e., Egypt (EG), Iran (IR), and Israel (IS), spanning from February 2015 to March 2022.
- Keep events with valid news URLs and from famous news agencies.
- We take the one-day time interval and collapse the 15 minutes-level timestamps of events on the same day to the day-level timestamp.

Datasets statistics

	$\lvert\mathcal{V}\rvert$	$\lvert\mathcal{E}\rvert$	#urls	#days	#train	#valid	#test
EG	2,594	225	96,081	2,584	377,430	36,588	28,644
IR	2,988	236	223,616	2,584	973,752	69,827	76,239
IS	3,456	238	345,611	2,584	1,430,389	171,518	156,695

The three datasets are diverse in both international roles and statistical characteristics.

Experiments and Analysis

Performance Comparison

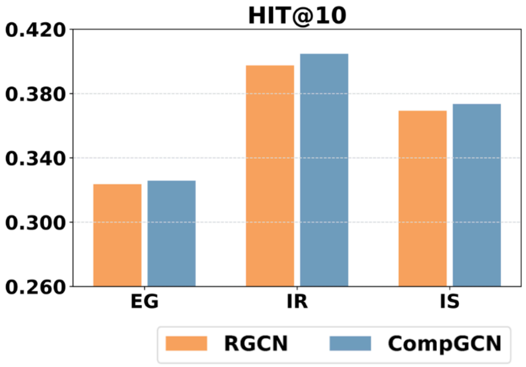
	Model	EG				IR				IS			
		MRR	HIT@1	HIT@3	HIT@10	MRR	HIT@1	HIT@3	HIT@10	MRR	HIT@1	HIT@3	HIT@10
Static KG Methods	**DistMult** [49]	0.1164	0.0344	0.1214	0.2927	0.1349	0.0392	0.1468	0.3379	0.1031	0.0223	0.0929	0.2950
	ConvE [11]	0.1151	0.0312	0.1272	0.2882	0.1365	0.0409	0.1485	0.3400	0.1060	0.0251	0.0984	0.2935
	ConvTransE [36]	0.1205	0.0377	0.1305	0.2921	0.1405	0.0462	0.1529	0.3412	0.1079	0.0287	0.0994	0.2930
	RotatE [38]	0.0892	0.0125	0.0772	0.2748	0.1055	0.0125	0.1074	0.3152	0.0879	0.0132	0.0714	0.2638
	RGCN [35]	0.0974	0.0279	0.1046	0.2377	0.1185	0.0366	0.1301	0.2860	0.0861	0.0242	0.0652	0.2307
Temporal KG Methods	**TANGO** [15]	0.1043	0.0240	0.1106	0.2761	0.1249	0.0281	0.1367	0.3314	0.0972	0.0171	0.0852	0.2889
	RE-NET [21]	0.1212	0.0413	0.1224	0.2932	0.1401	0.0451	0.1501	0.3452	0.1064	0.0263	0.1016	0.2894
	RE-GCN [29]	0.1245	0.0352	0.1366	0.3101	0.1647	0.0622	0.1796	0.3838	0.1301	0.0408	0.1281	0.3346
	EvoKG [34]	0.0797	0.0012	0.0775	0.2529	0.0892	0.0011	0.0767	0.3120	0.0779	0.0008	0.0518	0.2789
	HiSMatch [28]	0.1126	0.0275	0.1279	0.2906	0.1469	0.0496	0.1599	0.3572	0.1283	0.0434	0.1248	0.3017
Text-based Methods	**CMF**$_{ont}$ [9]	0.1206	0.0348	0.1298	0.3015	0.1527	0.0529	0.1643	0.3673	0.1248	0.0368	0.1224	0.3256
	CMF$_{art}$ [9]	0.1202	0.0345	0.1293	0.3027	0.1510	0.0496	0.1636	0.3716	0.1263	0.0382	0.1236	0.3261
Disentangled Methods	**DisenGCN** [31]	0.0849	0.0196	0.0805	0.2198	0.1084	0.0275	0.1096	0.2793	0.0833	0.0162	0.0633	0.2427
	DisenKGAT [47]	0.0801	0.0083	0.0822	0.2382	0.0895	0.0059	0.0977	0.2744	0.0724	0.0106	0.0429	0.2322
	SeCoGD(ours)	**0.1464**	**0.0593**	**0.1605**	**0.3236**	**0.1757**	**0.0724**	**0.1902**	**0.3975**	**0.1552**	**0.0595**	**0.1588**	**0.3693**
	%Improv.	17.59	57.29	17.50	4.35	6.68	16.40	5.90	3.57	19.29	37.10	23.97	10.37

- RE-GCN is the strongest baseline, even better than temporal event forecasting models with text information.
- Graph disentangle models do not perform well, indicating the importance of using context information as a prior guide.
- Our proposed SeCoGD performs better than RE-GCN by a large margin.

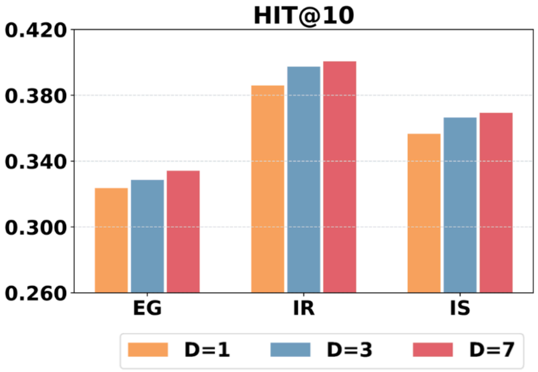
Experiments and Analysis

Ablation Study

Study of the separation stage:



Results of using different graph kernels.



Results of with different historical length D.

- CompGCN and RGCN perform similarly to each other on the three datasets, showing that our framework is not sensitive to relational modeling models.
- Longer historical length can yield better performance but take extra computational costs.

Experiments and Analysis

Ablation Study

Study of the collaboration stage:

Model	EG		IR		IS	
	MRR	H@10	MRR	H@10	MRR	H@10
SeCoGD	0.146	0.324	0.176	0.397	0.155	0.369
w/o Ent HG	0.139	0.315	0.168	0.391	0.147	0.359
w/o Rel HG	0.143	0.331	0.170	0.400	0.147	0.362
w/o Ent or Rel HG	0.138	0.315	0.163	0.386	0.144	0.355
Avr. Context	0.130	0.309	0.163	0.373	0.129	0.331

Study of the cross-context modelling and context-aware prediction

- Remove either relation or entity hypergraph are worse than SeCoGD but better than that of removing both, demonstrating the efficacy of both hypergraphs.
- The performance drop of removing the entity hypergraph is generally larger than that of removing the relation hypergraph, implying that the collaboration of entities is more valuable.

Experiments and Analysis

Ablation Study

Study of the collaboration stage:

Model	EG		IR		IS	
	MRR	H@10	MRR	H@10	MRR	H@10
RE-GCN	0.125	0.310	0.165	0.384	0.130	0.335
K-means	0.139	0.314	0.169	0.388	0.145	0.352
GMM	0.139	0.316	0.165	0.375	0.134	0.339
LDA(SeCoGD)	0.146	0.324	0.176	0.397	0.155	0.369

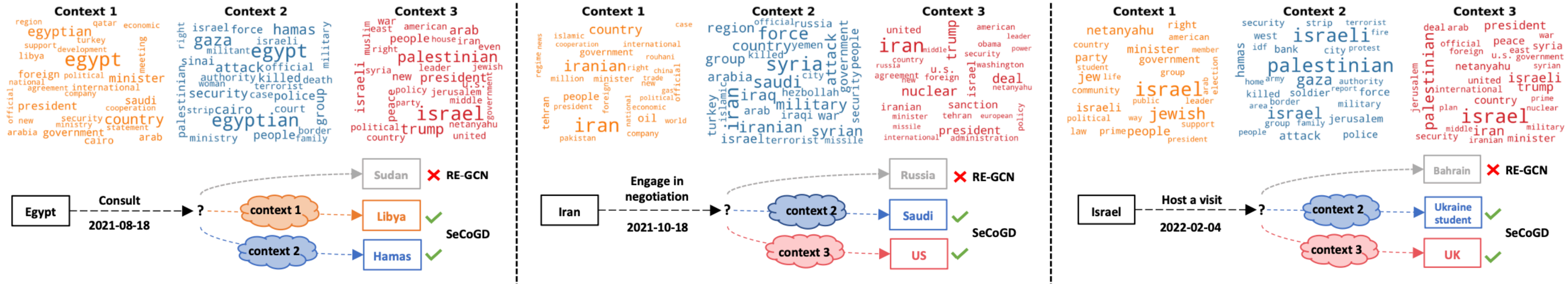
Study of alternative context generation methods.

- RE-GCN performs much worse than SeCoGD.
- The specification of the proper context during inference is crucial to SeCoGD, justifying our hypothesis that the context plays a pivotal role in accurate event forecasting.

Experiments and Analysis

Case Study

In each sub-figure, the context number K is set as three, the top shows the word cloud of each context, and the bottom illustrates several exemplar forecasting results by SeCoGD and RE-GCN.



(a) Dataset EG (b) Dataset IR (c) Dataset IS

- Each context in the EG dataset covers background information such as popular actors, important cities, and critical actions.
- Topics are prone to economic, military, and political events, respectively.
- The case study demonstrates the flexibility in depicting the event by context.

Conclusion

Contributions:

- We highlight the importance of diverse contexts in event forecasting and propose a novel task of context-aware event forecasting.
- We build a novel framework SeCoGD, and the two-stage design of separation and collaboration is effective in capturing the complex patterns in the multi-context scenario.
- We build three datasets based on GDELT to facilitate current and future studies for context-aware event forecasting. Our method significantly outperforms SOTA methods on the three datasets.

Future Works:

- Investigate the use of human-generated contexts (e.g., tags and categories) and more effective approaches for mining beneficial patterns from raw texts to improve context generation.
- Explore advanced graph disentanglement methods to enhance the performance of the SeCoGD framework and better separate event graphs and collaborate among contexts.

THANK YOU

Please contact the author via email to access the code and dataset or ask any questions.

Yunshan Ma: yunshan.ma@u.nus.edu
Chenchen Ye: chenchenye.ccye@gmail.com