中文信息处理实验室-让机器理解中文
Chinese Information Processing Laboratory

Honey or Poison? Solving the Trigger Curse in Few-shot Event Detection via Causal Intervention

Jiawei Chen[1,3], Hongyu Lin[1], Xianpei Han[1,2], Le Sun[1,2]

[1]Chinese Information Processing Laboratory [2]State Key Laboratory of Computer Science
Institute of Software, Chinese Academy of Sciences
[3]University of Chinese Academy of Sciences

Few-shot Event Detection

■ **Few-shot event detection (FSED)**

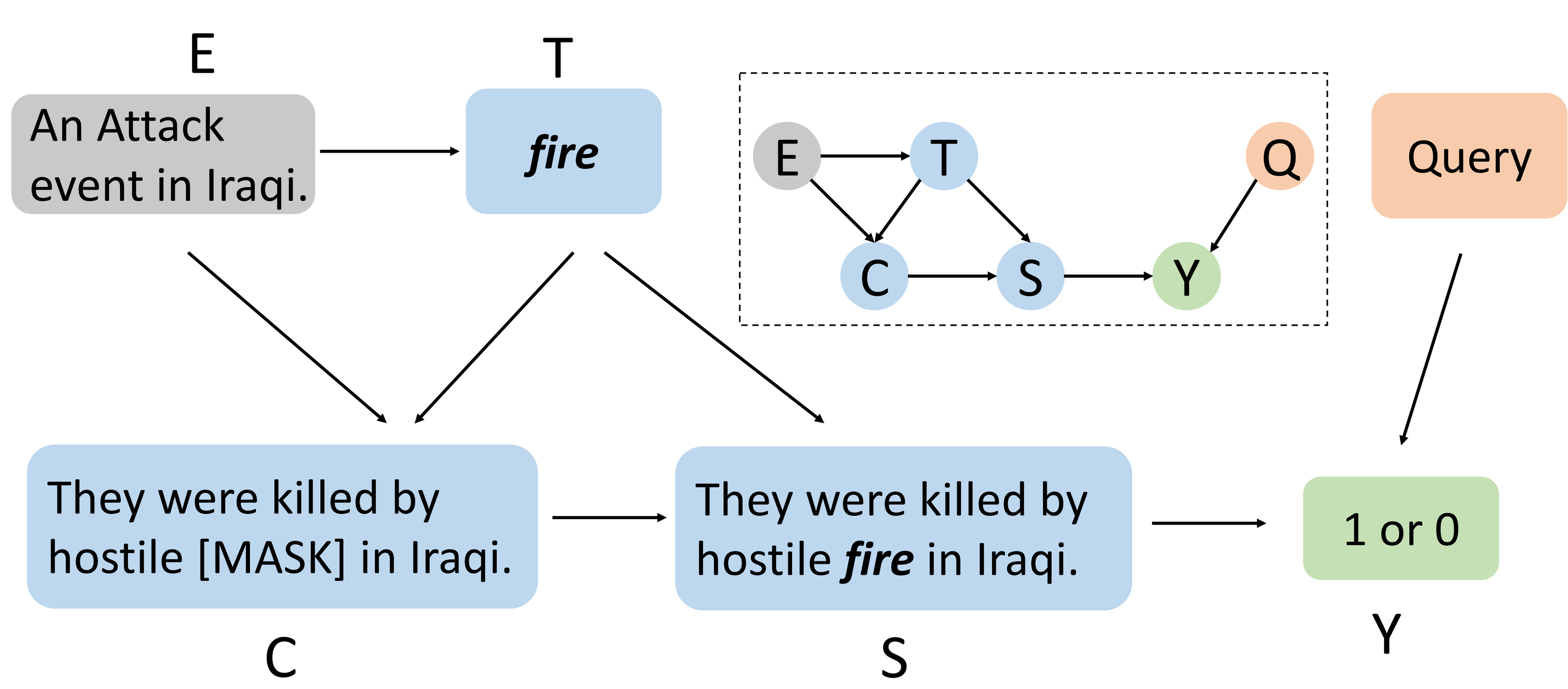
- Detect new event with a few labeled data.

■ **Challenge: Trigger Curse**

- Overfitting the trigger will harm the generalization ability.
- Underfitting the trigger will hurt the detection performance.

Trigger-centric structural causal model

■ We formulate the data distribution of FSED using a trigger-centric structural causal model (SCM).



■ **Nodes in the SCM**:

- E : Event
- T : Trigger set of support set
- C : Context set of support set
- S : Support set instances
- Y : Predicted result

■ **Edges in the SCM**:

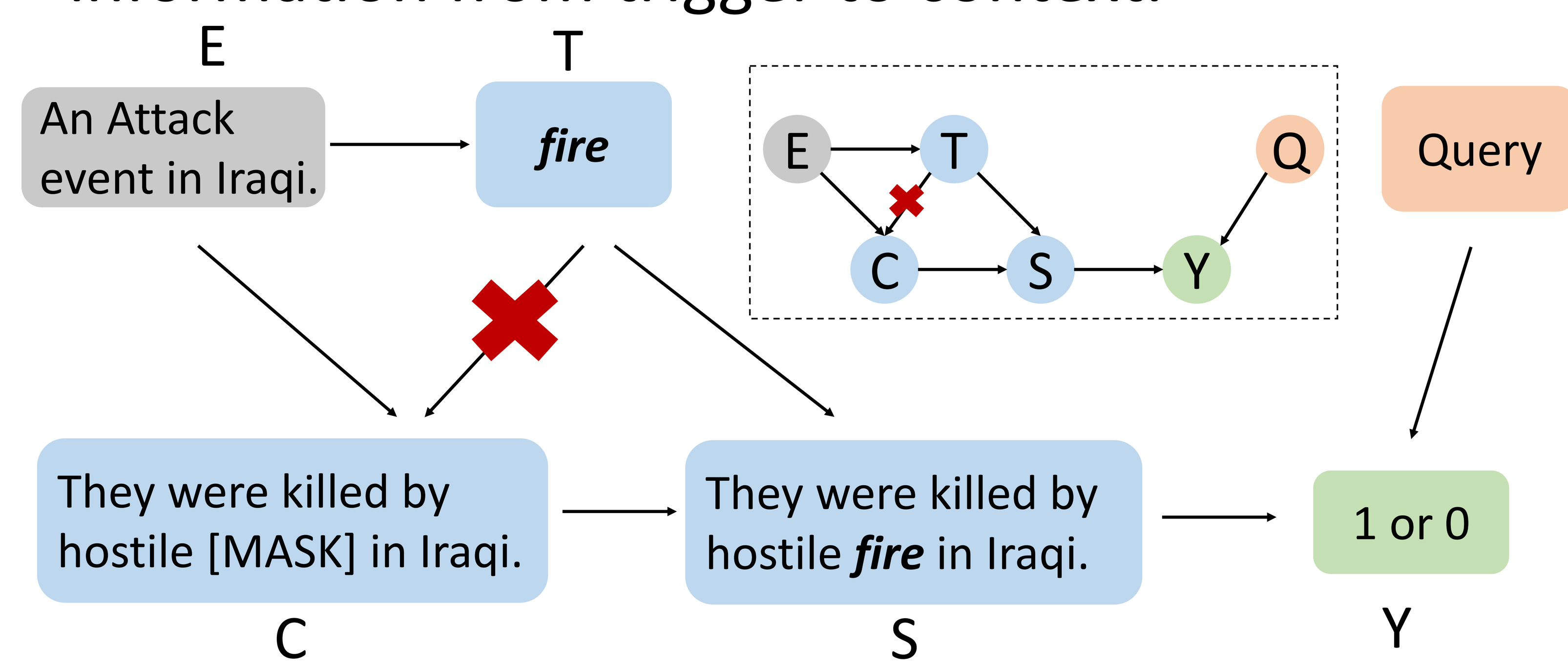
- $E \rightarrow T$: Trigger decision process.
- $E \rightarrow C \leftarrow T$: Context is generated depending on both the event and the trigger.
- $C \rightarrow S \leftarrow T$: Event instances are generated by combining context and the trigger.
- S→ $Y \leftarrow Q$: The result is predicted by the instances in support set and instances in query.

■ **There exists a backdoor path $C \leftarrow T \rightarrow Y$**

- Trigger is the confounder of context and predicted result.
- The backdoor misleads the conventional learning procedure to mistakenly regard effects of triggers as the effects of contexts.

Causal Intervention for solving Trigger Curse

■ We propose to intervene on context to block the information from trigger to context.



- We use Interventional distribution to optimizing causality.
- Backdoor adjustment is used to to estimate the interventional distribution.

$$P(Y|do(C=\mathrm{C}), E=e, Q=q)$$
$$= \sum_{s \in S} \sum_{t \in T} P(Y|s,q) P(s|\mathrm{C},t) P(t|e)$$

Experimental Results

	Model	ACE05		MAVEN		KBP17	
		Macro	Micro	Macro	Micro	Macro	Micro
Finetuing-based	Finetune	51.0±1.4	58.2±1.6	30.7±1.5	31.6±2.3	59.4±1.9	62.7±1.8
	Finetune*	39.9±1.1	45.5±0.7	20.8±1.0	20.6±0.8	45.0±0.7	47.3±0.6
	Pretrain+Finetune	22.9±6.0	20.3±4.3	20.9±4.6	16.9±5.2	35.1±5.9	30.1±5.5
	Pretrain+Finetune*	14.6±3.3	15.6±3.4	12.5±3.8	14.9±4.0	23.4±6.8	25.8±6.3
Prototypical Net	FS-Base	63.8±2.8	67.3±2.7	44.7±1.4	44.5±2.0	65.5±2.7	67.3±3.1
	FS-LexFree	52.7±2.9	53.9±3.2	25.6±1.0	21.8±1.4	60.7±2.5	61.4±2.8
	FS-ClusterLoss	64.9±1.5	69.4±2.0	44.2±1.2	44.0±1.2	65.5±2.3	67.1±2.4
	FS-Causal (Ours)	**73.0**±2.2	**76.9**±1.4	**52.1**±0.2	**55.0**±0.4	**70.9**±0.6	**73.2**±0.9
Relation Net	FS-Base	65.7±3.7	68.7±4.5	52.4±1.4	56.0±1.4	**67.2**±1.5	71.2±1.4
	FS-LexFree	59.3±3.5	60.1±3.9	43.8±1.9	45.9±2.4	61.9±2.4	65.4±2.8
	FS-ClusterLoss	57.6±2.3	60.2±3.2	46.3±1.1	51.8±1.4	56.8±3.0	62.1±2.5
	FS-Causal (Ours)	**67.2**±1.4	**71.8**±1.9	**53.0**±0.5	**57.0**±0.9	66.4±0.4	**72.0**±0.6

■ By intervening on the context in SCM and using backdoor adjustment during training, our method can effectively learn FSED models.

■ The causal theory is a promising technique for resolving the trigger cruse problem.

■ Our method can achieve state-of-the-art FSED performance.

Conclusions

■ Trigger Curse is a challenge in Few-shot Event Detection.

■ Identify trigger curse using structural causal model.

■ Causal intervention via backdoor adjustment can be used to solve trigger curse.