Document-Level Multi-Event Extraction with Event Proxy Nodes and Hausdorff Distance Minimization

Xinyu Wang, Lin Gui, Yulan He

Document-Level Multi-Event Extraction

Input:

... [5] Shenkai Petrochemical Co., Ltd. received the receipt from the company's shareholder, **Yexiang Investment Management Co., Ltd.** on the evening of November 15, 2016 Regarding the notice of the shares being frozen. ... [8] On November 14, 2016, Yexiang Investment received the Notice of Litigation Preservation from the **People's Court of Binjiang District**, and granted a total of **47,577,481 shares** held by Yexiang Investment will be frozen, and the freezing period is from **October 31, 2016** to **October 30, 2019** ... [10] Yexiang Investment is … holding **47,577,481 shares** of the company, accounting for **13.07%** of the company's total share capital. ... [12] On **February 2, 2016**, the **42,000,000 shares** held by it are pledged to **Haitong Securities Co., Ltd.,** and the repurchase transaction date was **February 1, 2017**. ...

Output:

Event #1: **Equity Pledge**
Pledger: **Yexiang Investment Management Co., Ltd.**
Pledgee: **Haitong Securities Co., Ltd.**
TotalHoldingShares: **47,577,481 shares**
TotalHoldingRatio: **13.07%**
PledgedShares: **42,000,000 shares**
StartDate: **February 2, 2016**
EndDate: **February 1, 2017**

Event #2: **Equity Freeze**
EquityHolder: **Yexiang Investment Management Co., Ltd.**
LegalInstitution: **People's Court of Binjiang District**
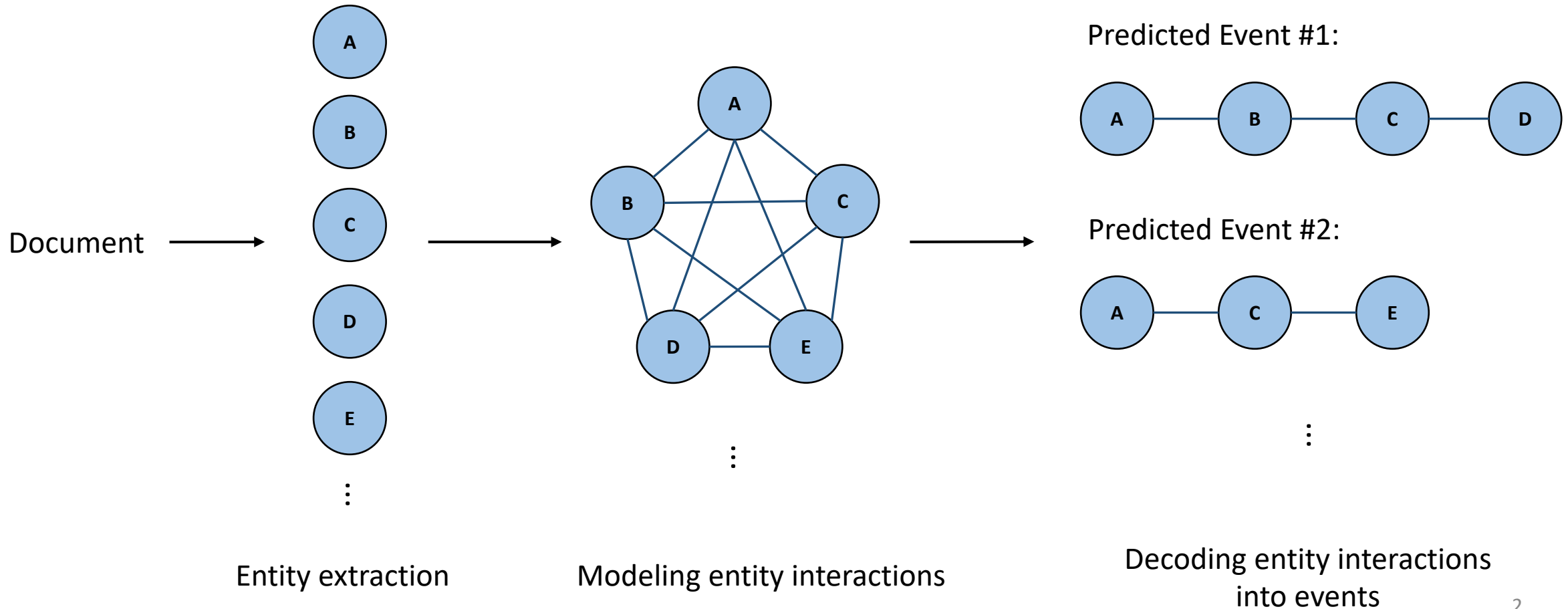TotalHoldingRatio: **13.07%**
FrozeShares: **47,577,481 shares**
StartDate: **October 31, 2016**
EndDate: **October 30, 2019**

Existing methods

The idea of existing method is modeling entity relations and decoding entity relations into events:

Entity extraction

Modeling entity interactions

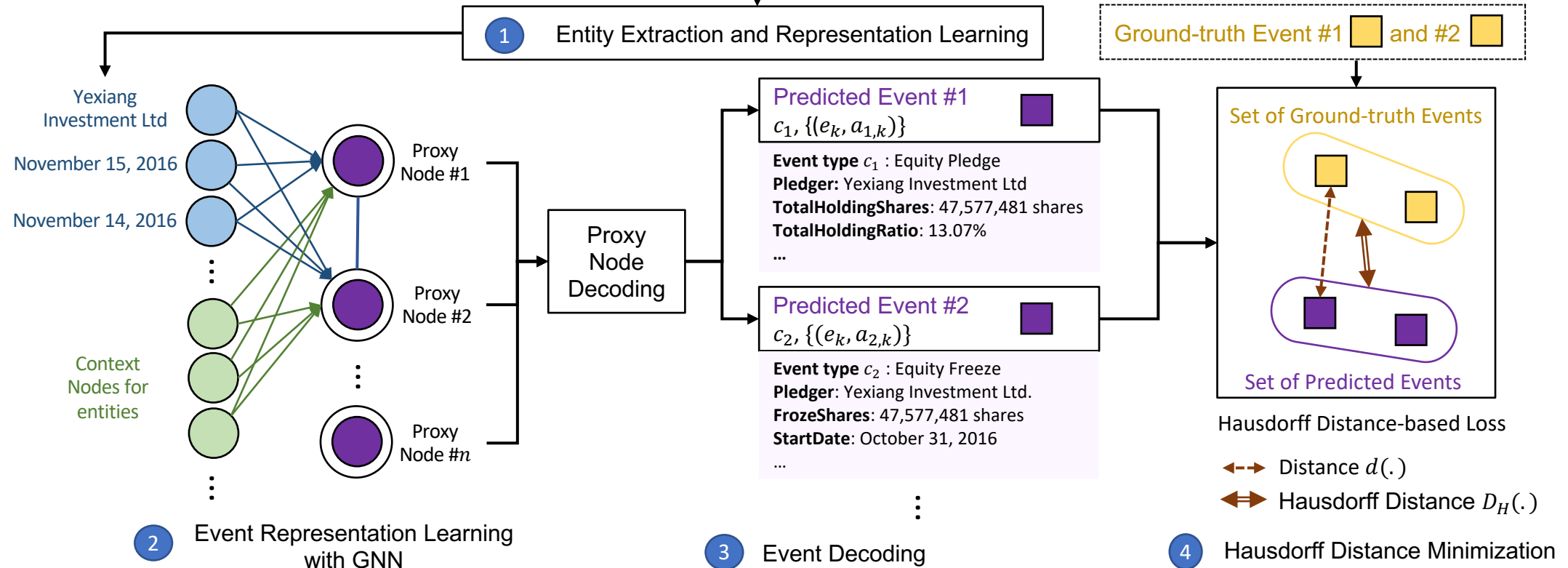Decoding entity interactions into events

limitations

- They decode events from entity information and tend to produce local optimal results without considering the interdependency of multiple events globally in a document.

- Event decoding by iteratively merging entities suffers from error propagation that an event type or an entity that has been incorrectly classified cannot be corrected later.

- Every decoding decision requires iterating all entity mentions in a document, which is computationally inefficient

Our method

We propose Proxy Nodes Clustering Network (ProCNet) to capture the global information among events with proxy nodes and Hausdorff distance in a new event-level metric space.


... the receipt from the company's shareholder, e_1:Yexiang Investment Management Co., Ltd. on the evening of e_2:November 15, 2016 Regarding the notice of the shares being frozen. ... On e_3:November 14, 2016, Yexiang Investment received the Notice of Litigation Preservation from the e_4:People's Court of Binjiang District, and granted a total of e_5:47,577,481 shares held by Yexiang Investment will be frozen, and the freezing period is from e_6:October 31, 2016 to e_7:October 30, 2019 ... Yexiang Investment is ... holding e_5':47,577,481 shares of the company, accounting for e_8:13.07% of the company's total share capital. ... On e_9:February 2, 2016, ...

① Entity Extraction and Representation Learning

Ground-truth Event #1 ▦ and #2 ▦

Yexiang Investment Ltd
November 15, 2016
November 14, 2016
Proxy Node #1
Proxy Node #2
Proxy Node #n
Context Nodes for entities

Proxy Node Decoding

Predicted Event #1
$c_1, \{(e_k, a_{1,k})\}$
Event type c_1 : Equity Pledge
Pledger: Yexiang Investment Ltd
TotalHoldingShares: 47,577,481 shares
TotalHoldingRatio: 13.07%
…

Predicted Event #2
$c_2, \{(e_k, a_{2,k})\}$
Event type c_2 : Equity Freeze
Pledger: Yexiang Investment Ltd.
FrozeShares: 47,577,481 shares
StartDate: October 31, 2016
…

Set of Ground-truth Events

Set of Predicted Events

Hausdorff Distance-based Loss

◀--▶ Distance $d(.)$
◀▶ Hausdorff Distance $D_H(.)$

② Event Representation Learning with GNN

③ Event Decoding

④ Hausdorff Distance Minimization

4

Entity Extraction and Representation Learning

... the receipt from the company's shareholder, e_1 :Yexiang Investment Management Co., Ltd. on the evening of e_2:November 15, 2016 Regarding the notice of the shares being frozen. ... On e_3 :November 14, 2016, Yexiang Investment received the Notice of Litigation Preservation from the e_4:People's Court of Binjiang District, and granted a total of e_5:47,577,481 shares held by Yexiang Investment will be frozen, and the freezing period is from e_6 :October 31, 2016 to e_7 :October 30, 2019 ... Yexiang Investment is … holding e_5' :47,577,481 shares of the company, accounting for e_8 :13.07% of the company's total share capital. ... On e_9:February 2, 2016, ...

Yexiang Investment Ltd

November 15, 2016

November 14, 2016

Context Nodes for entities

Pretrained Language Model
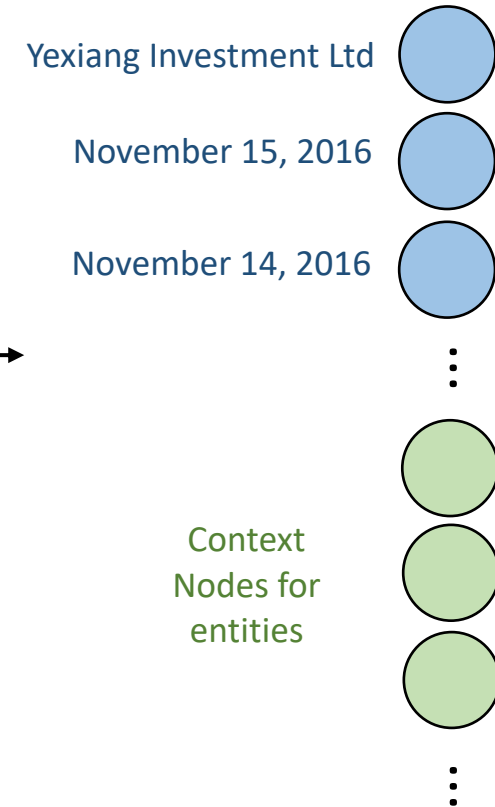
① Entity Extraction and Representation Learning
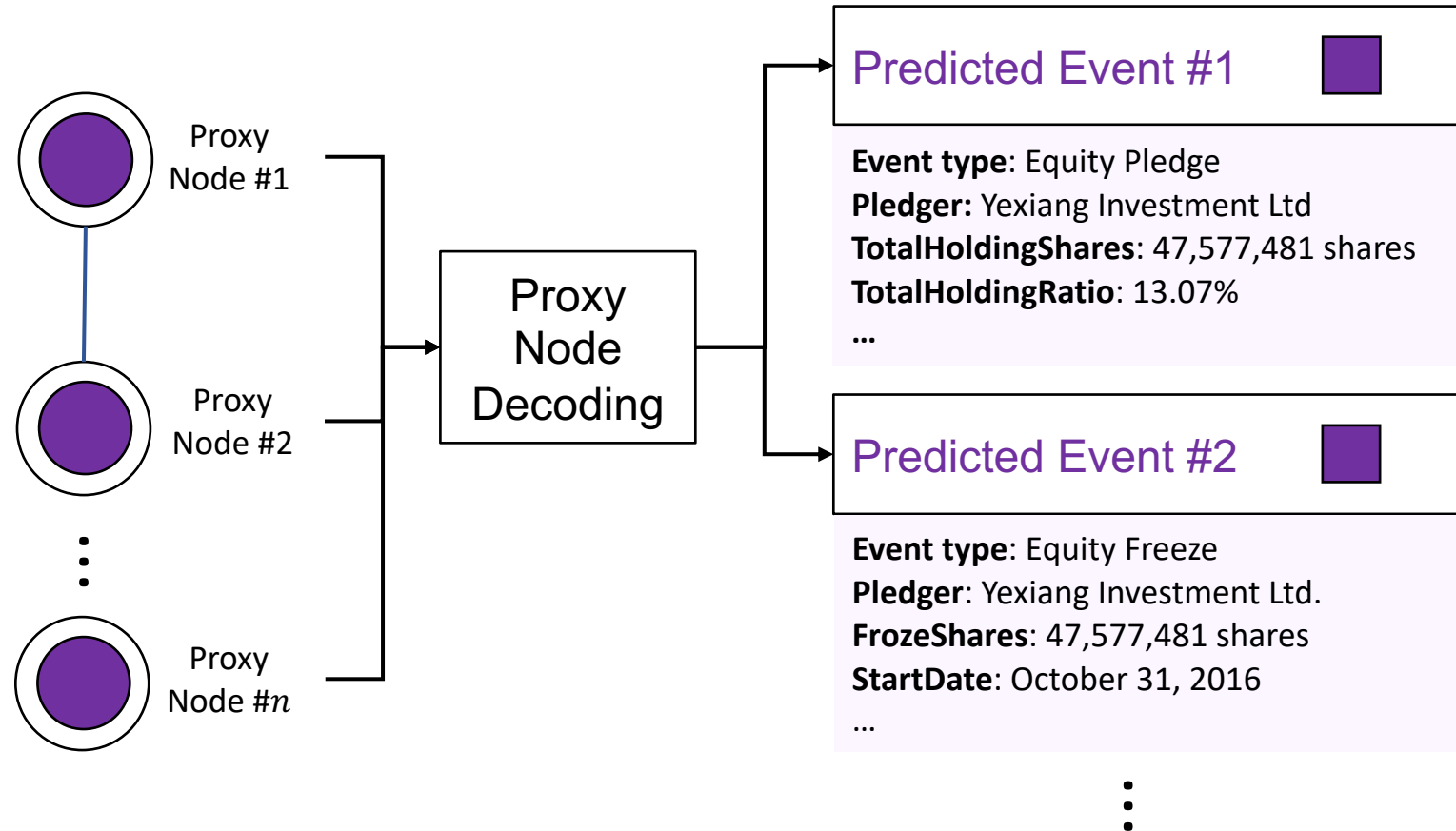
Event Representation Learning

- Each proxy node represents a pseudo-event.

- Proxy nodes are interconnected to allow information exchange among potential events.

- Learning with Hypernetwork Graph Neural Network.

- Each proxy node resides in a new event-level metric space by aggregating information from the entity-level space.

Proxy Node #1

Proxy Node #2
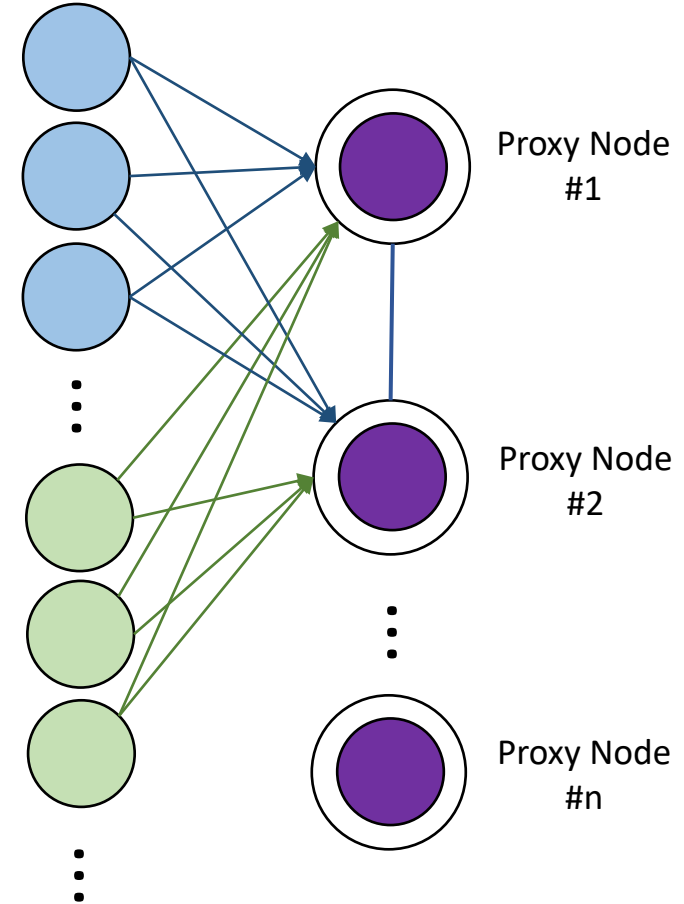
Proxy Node #n

2 Event Representation Learning with GNN

Event Decoding

Decode proxy node into events.

Proxy Node #1

Proxy Node #2

Proxy Node #n

Proxy Node Decoding

Predicted Event #1

Event type: Equity Pledge
Pledger: Yexiang Investment Ltd
TotalHoldingShares: 47,577,481 shares
TotalHoldingRatio: 13.07%
...

Predicted Event #2

Event type: Equity Freeze
Pledger: Yexiang Investment Ltd.
FrozeShares: 47,577,481 shares
StartDate: October 31, 2016
...

3   Event Decoding

Hausdorff Distance Minimization

- We treat the predicted events as a set and the ground-truth events as another set.

- Hausdorff distance measures the distance of the two sets and simultaneously consider all events and all their arguments.

- Minimize the Hausdorff distance to train the model to directly produce a globally optimal solution without the need of using decoding strategies as in existing approaches.
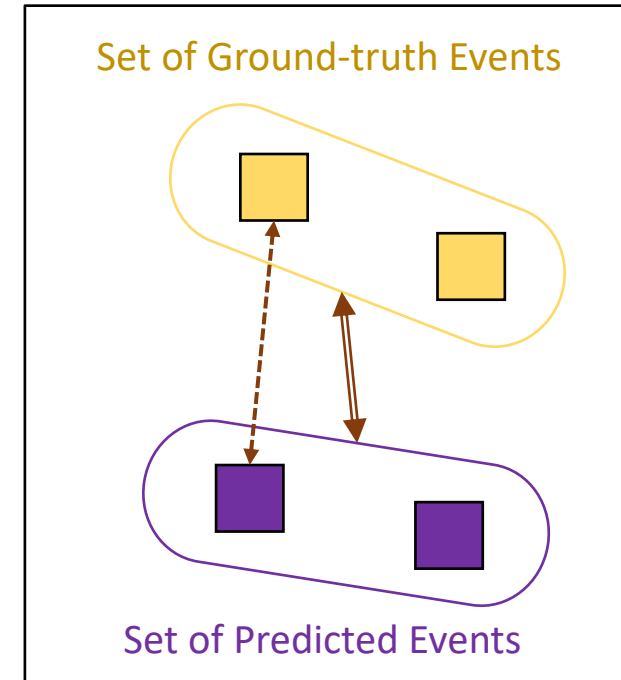
Set of Ground-truth Events
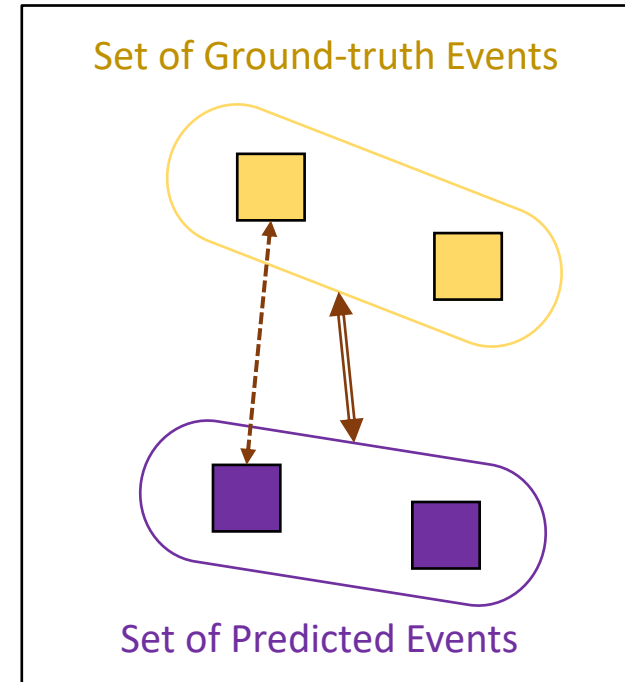
Set of Predicted Events

Distance $d(.)$

Hausdorff Distance $D_H(.)$

4 Hausdorff Distance Minimization

Hausdorff Distance Minimization

$$d(\mu_{z_i}, \mu_{y_j}) = \mathbf{CE}(\boldsymbol{p}_{\hat{c}_i}, c_j) + \frac{1}{|\bar{e}|} \sum_{k=1}^{|e|} \mathbf{CE}(\boldsymbol{p}_{\hat{a}_{i,k}}, a_{j,k})$$

We define distance between a pseudo-event (proxy node) u_z and a ground-truth event u_y as $d(.)$ with their cross-entropy of event type and argument.

Set of Ground-truth Events
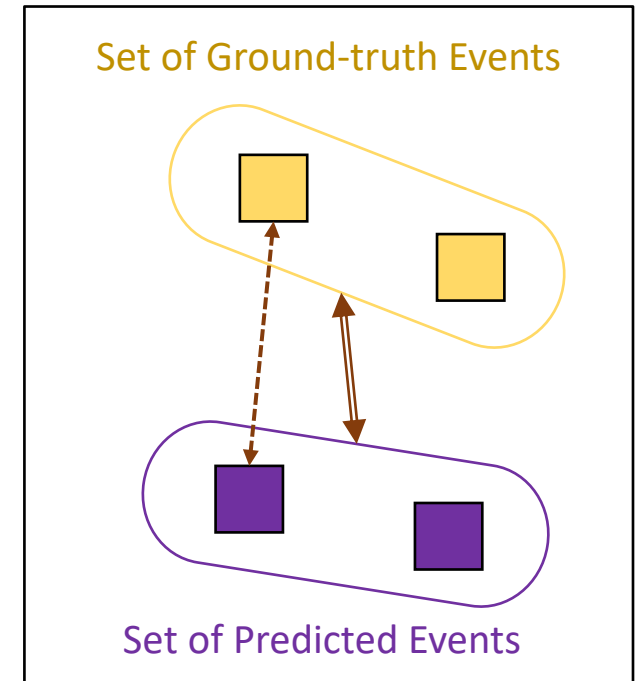
Set of Predicted Events

- - - → Distance $d(.)$

◄──► Hausdorff Distance $D_H(.)$

④ Hausdorff Distance Minimization

Hausdorff Distance Minimization

$$D_H(\mathcal{U}_z, \mathcal{U}_y) = \frac{1}{|\mathcal{U}_z|} \sum_{\mu_z \in \mathcal{U}_z} \min_{\mu_y \in \mathcal{U}_y} d(\mu_z, \mu_y)$$

$$+ \frac{1}{|\mathcal{U}_y|} \sum_{\mu_y \in \mathcal{U}_y} \min_{\mu_z \in \mathcal{U}_z} d(\mu_z, \mu_y)$$

We minimize an approximated average Hausdorff distance with defined $d(.)$.

Set of Ground-truth Events

Set of Predicted Events

- - - → Distance $d(.)$

◄——► Hausdorff Distance $D_H(.)$

④ Hausdorff Distance Minimization

Experiments – Dataset

ChFinAnn Dataset

Train set	25,632
Development set	3,204
Test set	3,204
Total	32,040

DuEE-Fin Dataset

Train set	6,515
Development set	500
Test set	1,171
Total	8,168

Event Type	Distribution
Equity Freeze	4.2%
Equity Repurchase	9.5%
Equity Underweight	16.0%
Equity Overweight	18.3%
Equity Pledge	52.0%

Event Type	Distribution
Win Bidding	9.5%
Financial Loss	11.1%
Business Acquisition	9.7%
Business Bankruptcy	2.5%
CCorporate Financing	5.5%
Companies Listing	5.1%
Shareholders Holdings Decrease	9.3%
Shareholders Holdings Increase	3.5%
Share Repurchase	14.1%
Regulatory Talk	1.8%
Pledge Release	7.7%
Pledge	10.8%
Executive Change	9.4%

Experiments – Overall Results

Overall F1 Scores
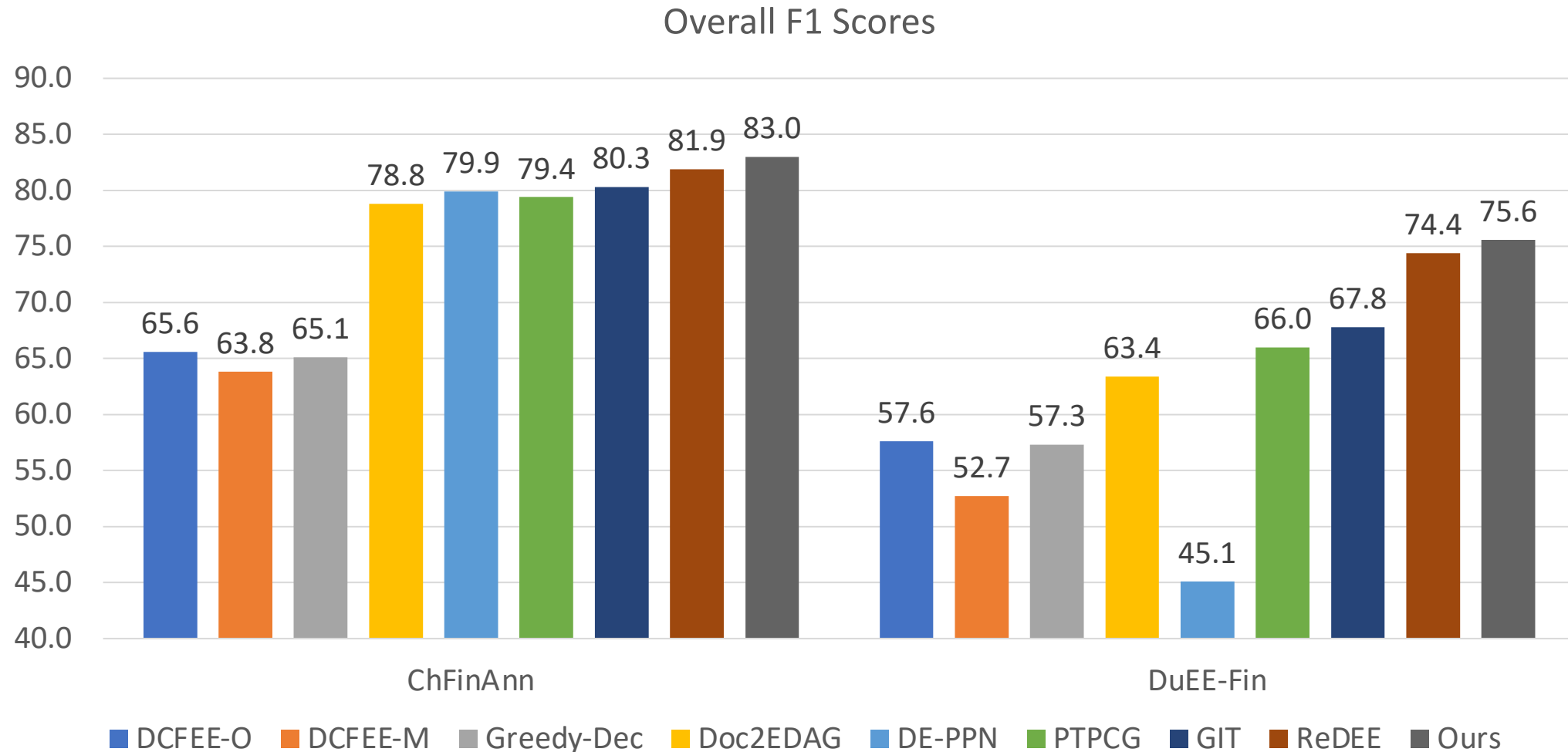
Experiments – Training Time

Training Time
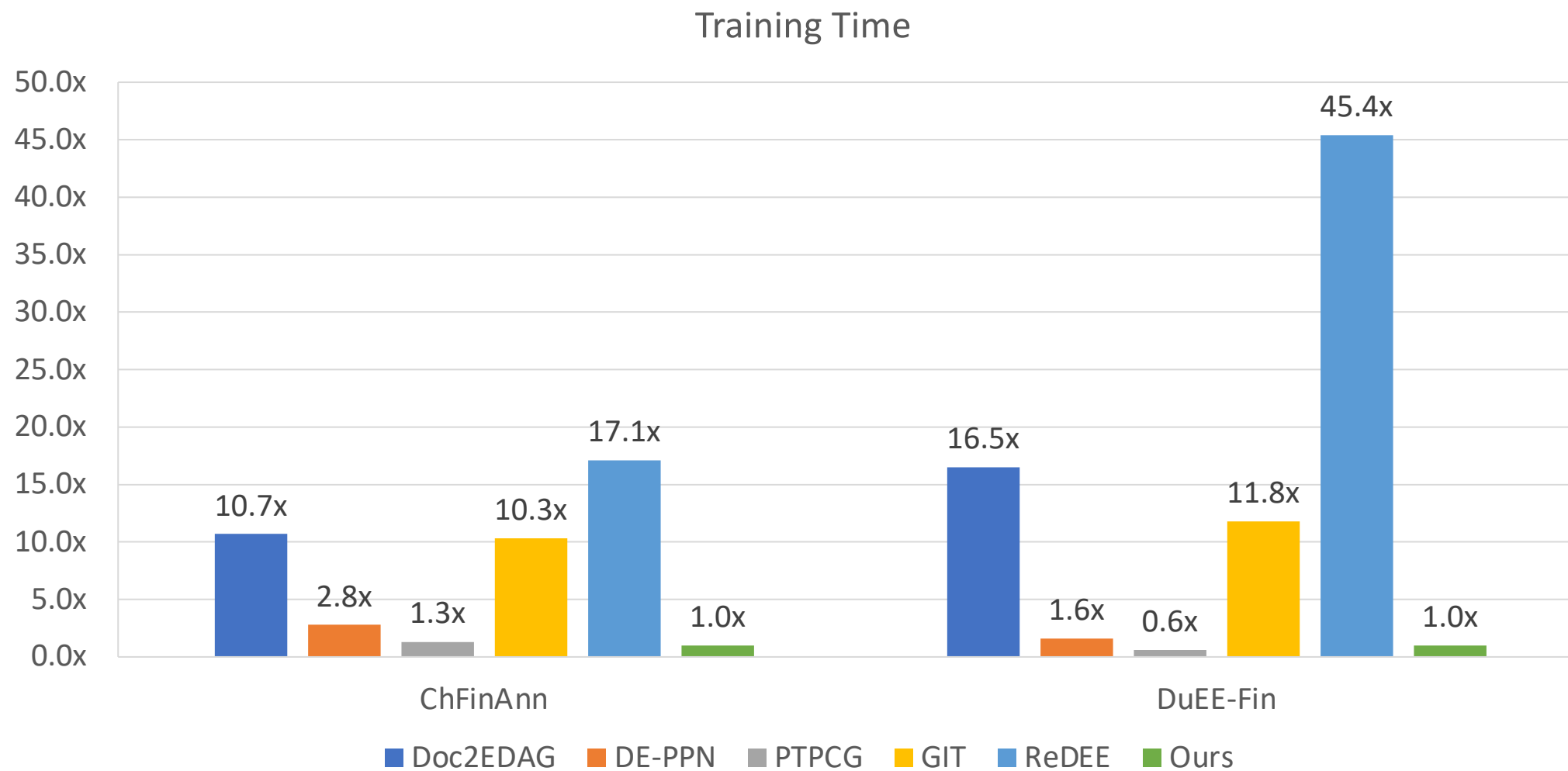
Error Case Study

... [9] Name of shareholder: Guangxin Holding Group Co., Ltd. ... [11] Guangxin Group disclosed its share increase plan on **October 23, 2018**. ... [21] As of **November 19, 2018**, Guangxin Group has accumulated 1653640 shares of the company through centralized bidding through the Shanghai Stock Exchange trading system, accounting for 0.08% of the company's total share capital, with an average increase price of 9.59 yuan per share, the increase amount is 15854061.0 yuan,... [22] On **November 20, 2018**, Guangxin Group increased its holdings of the company's shares by 1788000 shares through centralized bidding through the Shanghai Stock Exchange trading system, accounting for 10% of the company's total share capital. 0.08%, the average increase price is 9.36 yuan per share, and the increase amount is 16740153.0 yuan. ... [24] After the above-mentioned increase in holdings, Guangxin Group directly held **264558774 shares** of the company, accounting for 12.49% of the company's total share capital ...

True Events:

Event #1: Equity Overweight
EquityHolder: Guangxin Holding Group Co., Ltd.
TradedShares: 1653640 shares
StartDate: October 23, 2018
EndDate: November 19, 2018
LaterHoldingShares: null
AveragePrice: 9.59 yuan per share

Event #2: Equity Overweight
EquityHolder: Guangxin Holding Group Co., Ltd.
TradedShares: 1788000 shares
StartDate: November 20, 2018
EndDate: November 20, 2018
LaterHoldingShares: 264558774 shares
AveragePrice: 9.36 yuan per share

Predicted Events:

Event #1: Equity Overweight
EquityHolder: Guangxin Holding Group Co., Ltd.
TradedShares: 1653640 shares
StartDate: **null**
EndDate: **November 20, 2018**
LaterHoldingShares: null
AveragePrice: 9.59 yuan per share

Event #2: Equity Overweight
EquityHolder: Guangxin Holding Group Co., Ltd.
TradedShares: 1788000 shares
StartDate: **null**
EndDate: November 20, 2018
LaterHoldingShares: **null**
AveragePrice: 9.36 yuan per share

Conclusion

- We propose a new framework for document-level multi-event extraction by learning event proxy nodes in a new event-level metric space to better model the interactions among events.

- We utilize the Hausdorff distance to optimize the difference between the generated events and the gold standard events directly, considering all events simultaneously to train model towards the global minimum.