Forecasting Firm Material Event Sequences from SEC 8-K Reports

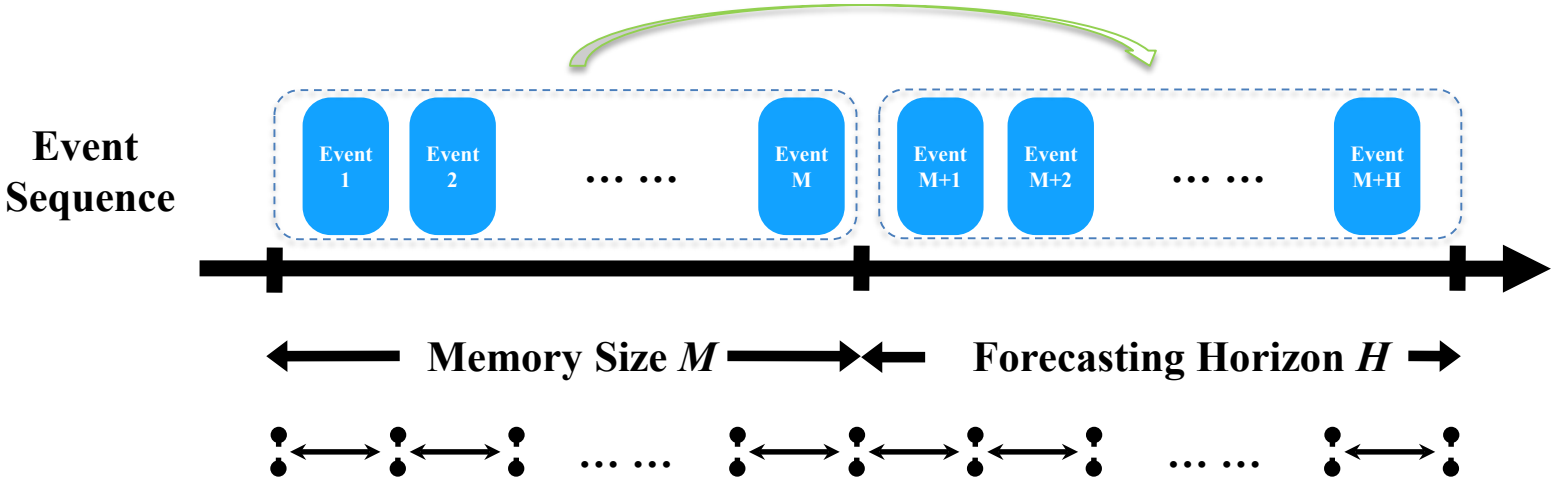
Shuang (Sophie) Zhai, Zhu (Drew) Zhang

IOWA STATE UNIVERSITY
COLLEGE OF BUSINESS

Background and Motivation

- SEC-required Form 8-K (also called 'Material Event Report' or 'Current Report')
- Corporate Event Series/Sequences (CES) embody corporate strategy and business processes
- Lack of study in previous finance literature
- Textual data presents challenges to traditional time series models

Problem Formulation



Material Event Types

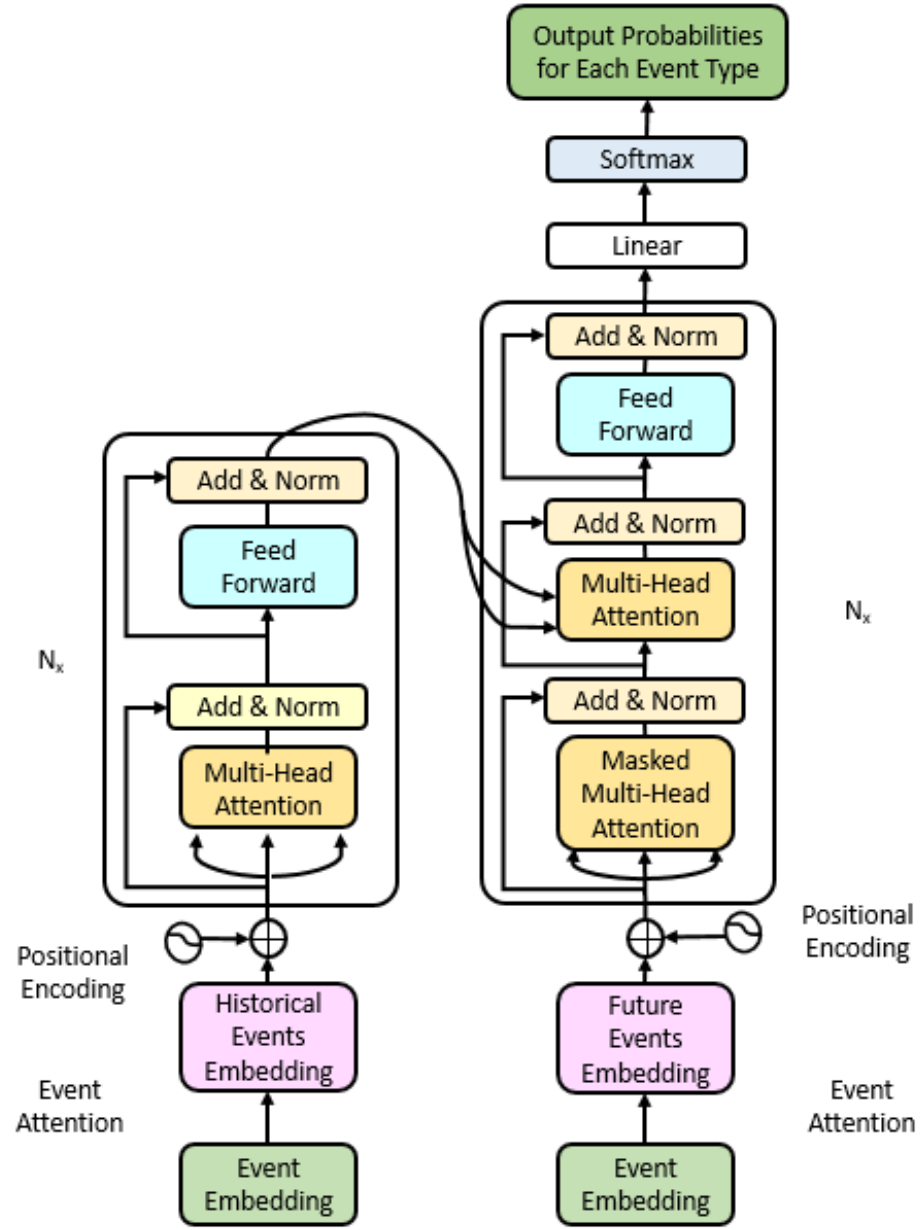
ID	Event Type	Code	Report Item	Examples
1	Business combination and restructuring	BC	**1.01, 1.02**, 2.01, **7.01, 8.01**	merger, acquisition, join venture, separation, spin-off
2	Financial activities	FN	**1.01, 1.02**, 2.03, 2.04, 2.05, 2.06, 3.02, 6.01, 6.02, 6.03, 6.04, 6.05, **7.01, 8.01**	lend, borrow, loan, Notes, payment, debt, stock, repurchase, dividend, asset-backed securities (ABS)
3	Operation activities	OA	**1.01, 1.02, 7.01, 8.01**	operation, contract, consulting, service, product, supply
4	Senior personnel change	PC	**1.01, 1.02**, 5.02, **7.01, 8.01**	executive officer/director, retire, leave, appointment
5	Information disclosure	ID	2.02, 4.01, 4.02, 5.07, 5.08, **7.01, 8.01**	conference, presentation, statement, exhibit
6	Document update	DU	3.03, 5.01, 5.03, 5.05, 5.06, **7.01, 8.01**	by-laws, code of ethics
7	Intellectual property activities	IP	**1.01, 1.02, 7.01, 8.01**	intellectual property, patent approval
8	Litigation and lawsuit	LL	**1.01, 1.02, 7.01, 8.01**	settlement, litigation, lawsuit
9	Delisting, trading suspension	DL	3.01, 5.04, **7.01, 8.01**	delisting, trading suspension
10	Bankruptcy	BK	1.03, **7.01, 8.01**	bankruptcy
11	None	None		no material event happend

Corporate Event Sequence Transformer (CEST) Model

$$E_{(i)j} = g_{k=0}^{|S|}(S_{(i)jk})$$

$$y_{(i)q} = f_{j=0}^{M}(E_{(i)j})$$

$$y_{(i)q} \in Y_{(i)}^{H}$$



Models in Comparison

- **Markovian Baseline**: Markov Chain Monte Carlo (MCMC) simulation.
- **GRU**: classical Seq2Seq model.
- **Corporate Event Sequence Transformer (CEST):** the proposed Transformer model.

Overall Evaluation

(M=36, H=12)

	MCMC Baseline	GRU	CEST
Cross Entropy	4.72	3.97	**3.65**
Perplexity	26.35	15.67	**12.55**

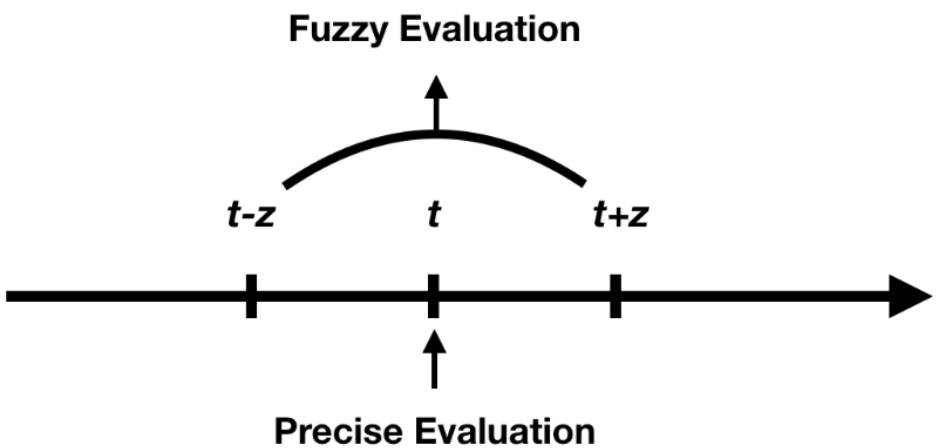
$$CCE = -\frac{1}{H}\sum_{t=1}^{H}\sum_{ev=1}^{|Ev|} y_{ev,t} log\hat{P}(y_{ev,t})$$

$$Perplexity = 2^{CCE}$$

Event-Type-Specific Evaluation

$$Eval_{ev,t} = \begin{cases} 1, if\ y_{ev,t} > threshold \\ 0, otherwise \end{cases}$$

- **Precise Evaluation** and **Fuzzy Evaluation**:



Model Performance

(M=36, H=12, precise evaluation on time t, F1%)

Event Type	MCMC	GRU	CEST
Business Combination and Restructuring	**8.54%**	3.89%	7.35%
Financial Activities	19.66%	36.23%	**39.56%**
Operation Activities	1.42%	17.91%	**25.56%**
Senior Personnel Change	20.86%	20.37%	**26.07%**
Information Disclosure	34.59%	74.77%	**82.35%**
Document Updates	**4.30%**	0.21%	1.72%
Intellectual Property	0.46%	13.66%	**22.66%**
Litigation and Lawsuit	1.26%	0.00%	**5.22%**
Delisting	**0.48%**	0.00%	0.00%
Bankruptcy	0.00%	0.00%	0.00%
Micro-Averaging	30.40%	51.22%	**58.68%**

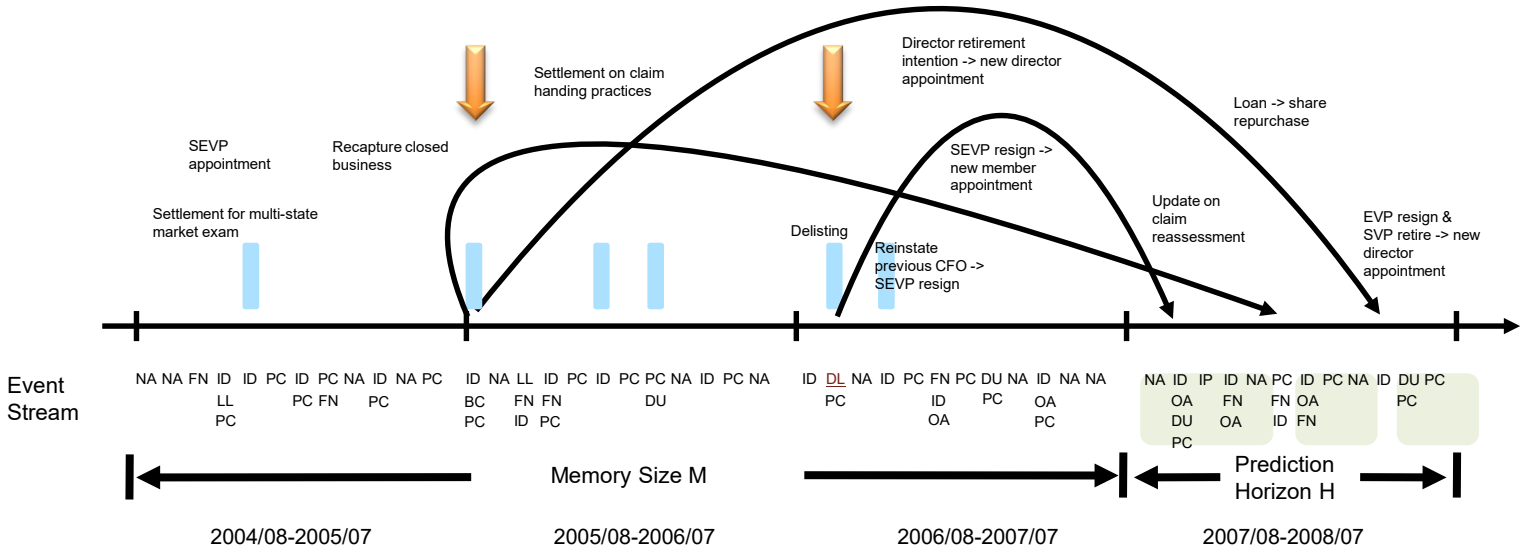
Model Performance

(M=36, H=12, fuzzy evaluation on time t±3, F1%)

Event Type	MCMC	GRU	CEST
Business Combination and Restructuring	**35.76%**	21.03%	29.79%
Financial Activities	59.37%	74.29%	**74.29%**
Operation Activities	8.22%	55.28%	**63.13%**
Senior Personnel Change	**61.20%**	56.46%	59.28%
Information Disclosure	76.03%	92.26%	**94.49%**
Document Updates	**21.22%**	1.22%	6.79%
Intellectual Property	2.65%	**45.87%**	45.67%
Litigation and Lawsuit	7.01%	0.00%	**15.41%**
Delisting	**2.44%**	0.00%	0.00%
Bankruptcy	0.00%	0.00%	0.00%
Micro-Averaging	72.37%	82.36%	**85.15%**

Explainability Example

Unum Group (UNM)



Conclusion and Future Work

- A Transformer-based model for event sequence forecasting
- Impressive performance
- Explainable insights
- Many potential business applications
- Model and data to be enriched

Questions/Comments

szhai@iastate.edu

zhuzhang@iastate.edu