dataset requires heavy human labeling and it only includes TE-level questions. In contrast, structured TE formulations mainly include TKG [3] and TCE with schema [23, 42]. However, TKG limits event analysis and forecasting at the TE level and therefore fails to capture multiple actors, relations, and timelines at the CE level. While TCE with schema models both CE and TE, its dependency on schema induction leads to unflexible event representation and restricts TEs to a relative temporal ordering instead of having complete timestamps. To the best of our knowledge, none of the previous works satisfies all the three properties of structured, complex, and time-complete, and in this work, our SCTc-TE formulation highlights these three critical properties for TE and CE forecasting.

Based on our SCTc-TE, the most promising methods are methods proposed for TKG forecasting. Static KG methods [10, 36] treat event forecasting as a link prediction task on a static event graph while ignoring event timestamps. TKG-based methods consider temporal information aside from relational information. For example, RE-NET [19], RE-GCN [26] and EvoKG [33] use GNN to aggregate the relational information at each timestamp and then use RNN to propagate this information over time. Some work also tries to map the discrete timestamps in a continuous space. For example, Know-Evolve [40] models event time as the Temporal Point Process (TPP), and TANGO [11] uses Neural Ordinary Differential Equation (ODE). Efforts have also been made to identify related historical events or a reasoning path to enhance the event forecasting performance. For example, CyGNet [48] retrieves relevant TE, and CluSTeR [25] and TimeTraveler [37] search for event evidence chain. Some TKG-based methods also try to incorporate textual information in addition to structural information into the event forecasting model. For example, Glean [8] uses word graph embedding and CMF [9] uses text embedding of the event relation ontology to enrich the information stored in each relational link in TKG. In parallel, multiple efforts have been devoted to schema-guided TCE forecasting [23, 42], however, such methods operate on a more complicated event graph definition, which cannot be easily adapted to our setting.

6 CONCLUSION AND FUTURE WORK

In this work, we highlighted three key properties of temporal events, structured, complex, and time-complete, and presented a new formulation of SCTc-TE. Thereby, we developed a fully automated pipeline that employs both LLM and time-aware clustering to construct SCTc-TEs from a large amount of news articles, and we constructed two large-scale datasets MidEast-TE and GDELT-TE. Finally, we proposed a novel forecasting method LoGo that outperforms SOTA methods by a large margin on both datasets.

Several future directions warrant attention: improving the zero-shot event extraction by LLMs, exploring automatic schema induction from data, and developing advanced graph models for SCTc-TE forecasting are key areas for further research.

REFERENCES

[1] Daniel M Benjamin, Fred Morstatter, Ali E Abbas, Andres Abeliuk, Pavel Atanasov, Stephen Bennett, Andreas Beger, Saurabh Birari, David V Budescu, Michele Catasta, et al. 2023. Hybrid forecasting of geopolitical events. *AI Magazine* (2023).

[2] Elizabeth Boschee, Jennifer Lautenschlager, Sean O'Brien, Steve Shellman, James Starz, and Michael Ward. 2015. CAMEO.CDB.09b5.pdf. In *ICEWS Coded Event Data*. Harvard Dataverse. https://doi.org/10.7910/DVN/28075/SCJPXX

[3] Borui Cai, Yong Xiang, Longxiang Gao, He Zhang, Yunfeng Li, and Jianxin Li. 2022. Temporal Knowledge Graph Completion: A Survey. *CoRR* abs/2201.08236 (2022).

[4] Tommaso Caselli and Piek Vossen. 2017. The Event StoryLine Corpus: A New Benchmark for Causal and Temporal Relation Extraction. In *NEWS@ACL*. Association for Computational Linguistics, 77–86.

[5] Tommaso Caselli and Piek Vossen. 2017. The event storyline corpus: A new benchmark for causal and temporal relation extraction. In *Proceedings of the Events and Stories in the News Workshop*. 77–86.

[6] Wei-Lin Chiang, Zhuohan Li, Zi Lin, Ying Sheng, Zhanghao Wu, Hao Zhang, Lianmin Zheng, Siyuan Zhuang, Yonghao Zhuang, Joseph E. Gonzalez, Ion Stoica, and Eric P. Xing. 2023. Vicuna: An Open-Source Chatbot Impressing GPT-4 with 90%* ChatGPT Quality. https://lmsys.org/blog/2023-03-30-vicuna/

[7] Kyunghyun Cho, Bart van Merrienboer, Çaglar Gülçehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. In *EMNLP*. ACL, 1724–1734.

[8] Songgaojun Deng, Huzefa Rangwala, and Yue Ning. 2020. Dynamic Knowledge Graph based Multi-Event Forecasting. In *KDD*. ACM, 1585–1595.

[9] Songgaojun Deng, Huzefa Rangwala, and Yue Ning. 2021. Understanding Event Predictions via Contextualized Multilevel Feature Learning. In *CIKM*. ACM, 342–351.

[10] Tim Dettmers, Pasquale Minervini, Pontus Stenetorp, and Sebastian Riedel. 2018. Convolutional 2D Knowledge Graph Embeddings. In *AAAI*. AAAI Press, 1811–1818.

[11] Zifeng Ding, Zhen Han, Yunpu Ma, and Volker Tresp. 2021. Temporal Knowledge Graph Forecasting with Neural ODE. abs/2101.05151 (2021). https://api.semanticscholar.org/CorpusID:231592393

[12] Xinya Du, Zixuan Zhang, Sha Li, Pengfei Yu, Hongwei Wang, Tuan Lai, Xudong Lin, Ziqi Wang, Iris Liu, Ben Zhou, et al. 2022. Resin-11: Schema-guided event prediction for 11 newsworthy scenarios. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: System Demonstrations*. 54–63.

[13] Tianyu Gao, Xingcheng Yao, and Danqi Chen. 2021. SimCSE: Simple Contrastive Learning of Sentence Embeddings. In *EMNLP (1)*. Association for Computational Linguistics, 6894–6910.

[14] Demian Gholipour Ghalandari, Chris Hokamp, Nghia The Pham, John Glover, and Georgiana Ifrim. 2020. A Large-Scale Multi-Document Summarization Dataset from the Wikipedia Current Events Portal. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel Tetreault (Eds.). Association for Computational Linguistics, Online, 1302–1308. https://doi.org/10.18653/v1/2020.acl-main.120

[15] Xavier Glorot and Yoshua Bengio. 2010. Understanding the difficulty of training deep feedforward neural networks. In *AISTATS (JMLR Proceedings, Vol. 9)*. JMLR.org, 249–256.

[16] Maarten Grootendorst. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *CoRR* abs/2203.05794 (2022).

[17] Yizhu Jiao, Ming Zhong, Jiaming Shen, Yunyi Zhang, Chao Zhang, and Jiawei Han. 2023. Unsupervised Event Chain Mining from Multiple Documents. In *WWW*. ACM, 1948–1959.

[18] Woojeong Jin, Rahul Khanna, Suji Kim, Dong-Ho Lee, Fred Morstatter, Aram Galstyan, and Xiang Ren. 2021. ForecastQA: A Question Answering Challenge for Event Forecasting with Temporal Text Data. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, Chengqing Zong, Fei Xia, Wenjie Li, and Roberto Navigli (Eds.). Association for Computational Linguistics, Online, 4636–4650. https://doi.org/10.18653/v1/2021.acl-long.357

[19] Woojeong Jin, Meng Qu, Xisen Jin, and Xiang Ren. 2020. Recurrent Event Network: Autoregressive Structure Inferenceover Temporal Knowledge Graphs. In *EMNLP (1)*. Association for Computational Linguistics, 6669–6683.

[20] Xiaomeng Jin, Manling Li, and Heng Ji. 2022. Event Schema Induction with Double Graph Autoencoders. In *NAACL-HLT*. Association for Computational Linguistics, 2013–2025.

[21] Diederik P. Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *ICLR (Poster)*.

[22] Kalev Leetaru and Philip A Schrodt. 2013. Gdelt: Global data on events, location, and tone, 1979–2012. In *ISA annual convention*, Vol. 2. Citeseer, 1–49.

[23] Manling Li, Sha Li, Zhenhailong Wang, Lifu Huang, Kyunghyun Cho, Heng Ji, Jiawei Han, and Clare R. Voss. 2021. The Future is not One-dimensional: Complex Event Schema Induction by Graph Modeling for Event Prediction. In *EMNLP (1)*. Association for Computational Linguistics, 5203–5215.

[24] Zixuan Li, Zhongni Hou, Saiping Guan, Xiaolong Jin, Weihua Peng, Long Bai, Yajuan Lyu, Wei Li, Jiafeng Guo, and Xueqi Cheng. 2022. HiSMatch: Historical Structure Matching based Temporal Knowledge Graph Reasoning. *CoRR* abs/2210.09708 (2022).

[25] Zixuan Li, Xiaolong Jin, Saiping Guan, Wei Li, Jiafeng Guo, Yuanzhuo Wang, and Xueqi Cheng. 2021. Search from History and Reason for Future: Two-stage Reasoning on Temporal Knowledge Graphs. In *ACL*. https://api.semanticscholar.org/CorpusID:235266233

[26] Zixuan Li, Xiaolong Jin, Wei Li, Saiping Guan, Jiafeng Guo, Huawei Shen, Yuanzhuo Wang, and Xueqi Cheng. 2021. Temporal Knowledge Graph Reasoning Based on Evolutional Representation Learning. In *SIGIR*. ACM, 408–417.

[27] Ying Lin, Heng Ji, Fei Huang, and Lingfei Wu. 2020. A Joint Neural Model for Information Extraction with Global Features. In *ACL*. Association for Computational Linguistics, 7999–8009.

[28] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *CoRR* abs/1907.11692 (2019).

[29] Leland McInnes and John Healy. 2017. Accelerated Hierarchical Density Based Clustering. In *ICDM Workshops*. IEEE Computer Society, 33–42.

[30] Leland McInnes and John Healy. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *CoRR* abs/1802.03426 (2018).

[31] Fred Morstatter. 2021. RCT-B. (2021). https://doi.org/10.7910/DVN/ROTHFT

[32] OpenAI. 2023. GPT-4 Technical Report. *CoRR* abs/2303.08774 (2023).

[33] Namyong Park, Fuchen Liu, Purvanshi Mehta, Dana Cristofor, Christos Faloutsos, and Yuxiao Dong. 2022. EvoKG: Jointly Modeling Event Time and Network Structure for Reasoning over Temporal Knowledge Graphs. In *WSDM*.

[34] Revanth Gangi Reddy, Yi R. Fung, Qi Zeng, Manling Li, Ziqi Wang, Paul Sullivan, and Heng Ji. 2023. SmartBook: AI-Assisted Situation Report Generation. (2023). arXiv:2303.14337 [cs.CL]

[35] Michael Sejr Schlichtkrull, Thomas N. Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, and Max Welling. 2018. Modeling Relational Data with Graph Convolutional Networks. In *ESWC (Lecture Notes in Computer Science, Vol. 10843)*. Springer, 593–607.

[36] Chao Shang, Yun Tang, Jing Huang, Jinbo Bi, Xiaodong He, and Bowen Zhou. 2019. End-to-End Structure-Aware Convolutional Networks for Knowledge Base Completion. In *AAAI*. AAAI Press, 3060–3067.

[37] Haohai Sun, Jialu Zhong, Yunpu Ma, Zhen Han, and Kun He. 2021. TimeTraveler: Reinforcement Learning for Temporal Knowledge Graph Forecasting. In *EMNLP*. https://api.semanticscholar.org/CorpusID:237454564

[38] Jialong Tang, Hongyu Lin, Zhuoqun Li, Yaojie Lu, Xianpei Han, and Le Sun. 2023. Harvesting Event Schemas from Large Language Models. *CoRR* abs/2305.07280 (2023).

[39] Meihan Tong, Bin Xu, Shuai Wang, Meihuan Han, Yixin Cao, Jiangqi Zhu, Siyu Chen, Lei Hou, and Juanzi Li. 2022. DocEE: A Large-Scale and Fine-grained Benchmark for Document-level Event Extraction. In *NAACL-HLT*. Association for Computational Linguistics, 3970–3982.

[40] Rakshit Trivedi, Hanjun Dai, Yichen Wang, and Le Song. 2017. Know-evolve: deep temporal reasoning for dynamic knowledge graphs. In *ICML*. 3462–3471.

[41] Christopher Walker, Stephanie Strassel, Julie Medero, and Kazuaki Maeda. 2006. ACE 2005 Multilingual Training Corpus. *Linguistic Data Consortium* 57 (2006).

[42] Hongwei Wang, Zixuan Zhang, Sha Li, Jiawei Han, Yizhou Sun, Hanghang Tong, Joseph P. Olive, and Heng Ji. 2022. Schema-Guided Event Graph Completion. *CoRR* abs/2206.02921 (2022).

[43] Xiang Wei, Xingyu Cui, Ning Cheng, Xiaobin Wang, Xin Zhang, Shen Huang, Pengjun Xie, Jinan Xu, Yufeng Chen, Meishan Zhang, Yong Jiang, and Wenjuan Han. 2023. Zero-Shot Information Extraction via Chatting with ChatGPT. *CoRR* abs/2302.10205 (2023).

[44] Nils B Weidmann and Michael D Ward. 2010. Predicting conflict in space and time. *Journal of Conflict Resolution* 54, 6 (2010), 883–901.

[45] Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. 2015. Embedding Entities and Relations for Learning and Inference in Knowledge Bases. In *ICLR (Poster)*.

[46] Susik Yoon, Yu Meng, Dongha Lee, and Jiawei Han. 2023. SCStory: Self-supervised and Continual Online Story Discovery. In *WWW*. ACM, 1853–1864.

[47] Liang Zhao. 2022. Event Prediction in the Big Data Era: A Systematic Survey. *ACM Comput. Surv.* 54, 5 (2022), 94:1–94:37.

[48] Cunchao Zhu, Muhao Chen, Changjun Fan, Guangquan Cheng, and Yan Zhan. 2020. Learning from History: Modeling Temporal Knowledge Graphs with Sequential Copy-Generation Networks. In *AAAI Conference on Artificial Intelligence*. https://api.semanticscholar.org/CorpusID:229180723

[49] Yuqi Zhu, Xiaohan Wang, Jing Chen, Shuofei Qiao, Yixin Ou, Yunzhi Yao, Shumin Deng, Huajun Chen, and Ningyu Zhang. 2023. LLMs for Knowledge Graph Construction and Reasoning: Recent Capabilities and Future Opportunities. *CoRR* abs/2305.13168 (2023).