Scalable temporal clique enumeration[*]

Kaijie Zhu[1,2]

[1]Eindhoven University of Technology, The Netherlands

`k.zhu@tue.nl`

[2]National Digital Switching System Engineering & Technological R&D Center, China

Abstract

We study the problem of enumeration of all k-sized subsets of temporal events that mutually overlap at some point in a query time window. This problem arises in many application domains, e.g., in social networks, life sciences, smart cities, telecommunications, and others. To our knowledge, the general temporal clique problem has not been identified and studied before. The special case when $k = 2$ has been well-studied in the literature as the *interval join* problem, e.g., [1, 2]. Although current competitive methods for interval joins can be easily adapted to clique enumeration, they still suffer from scalability issues such as unnecessary scanning of records when applied to large data sets occurring in practice.

We propose a start time index (STI) approach that overcomes the efficiency bottlenecks of current methods which are based on 2-way join algorithms to enumerate temporal k-cliques [3]. Additionally, we investigate how precomputed checkpoints can be used to further improve the efficiency of STI. Our experimental results demonstrate that STI outperforms the state of the art by a wide margin and that our checkpointing strategies are effective.

References

[1] Panagiotis Bouros and Nikos Mamoulis. A forward scan based plane sweep algorithm for parallel interval joins. *Proceedings of the VLDB Endowment*, 10(11):1346–1357, 2017.

[2] Danila Piatov, Sven Helmer, and Anton Dignös. An interval join optimized for modern hardware. In *IEEE 32nd International Conference on Data Engineering (ICDE)*, pages 1098–1109, 2016.

[3] Kaijie Zhu, George Fletcher, Nikolay Yakovets, Odysseas Papapetrou, and Yuqing Wu. Scalable temporal clique enumeration. In *Proceedings of the 16th International Symposium on Spatial and Temporal Databases (SSTD)*, pages 120–129, 2019.