Memento Filter: A Fast, Dynamic, and Robust Range Filter

Navid Eslami

Department of Computer Science
University of Toronto
navideslami@cs.toronto.edu

Range filters are probabilistic data structures that answer approximate range emptiness queries. They aid in avoiding processing empty range queries and have use cases in many application domains such as key-value stores [7, 16], social web analytics [4], statistics aggregation of time series [8], and SQL table accesses [10]. However, current range filter designs do not support dynamically changing and growing datasets [1, 19, 11, 18, 13, 9, 17, 2, 5]. Moreover, several of these designs also exhibit impractically high false positive rates under correlated workloads [1, 19, 18, 13, 9, 17, 2], which are common in practice [5]. These impediments restrict the applicability of range filters across a wide range of use cases.

We introduce Memento filter, the first range filter to offer dynamicity, fast operations, and a robust false positive rate guarantee for any workload. Memento filter partitions the key universe and clusters its keys according to this partitioning. For each cluster, it stores a fingerprint and a list of key suffixes contiguously. The encoding of these lists makes them amenable to existing dynamic filter structures. Due to the well-defined one-to-one mapping from keys to suffixes, Memento filter supports inserts and deletes and can even expand to accommodate a growing dataset.

We implement Memento filter on top of a Rank-and-Select Quotient filter [14] and InfiniFilter [6] and demonstrate that it achieves competitive false positive rates and performance with the state-of-the-art while also providing dynamicity. Due to its dynamicity, Memento filter is the first range filter applicable to B-Trees [3, 15]. We showcase this by integrating Memento filter into WiredTiger, a B-Tree-based key-value store [12]. Memento filter doubles WiredTiger's range query throughput when 50% of the queries are empty while keeping all other cost metrics unharmed.

References

[1] Karolina Alexiou, Donald Kossmann, and Per-Åke Larson. Adaptive range filters for cold data: avoiding trips to siberia. Proc. VLDB Endow., 6(14):17141725, sep 2013.

[2] Guanduo Chen, Zhenying He, Meng Li, and Siqiang Luo. Oasis: An optimal disjoint segmented learned range filter. Proc. VLDB Endow., 17(8):19111924, may 2024.

[3] Douglas Comer. Ubiquitous b-tree. ACM Comput. Surv., 11(2):121137, jun 1979.

[4] Brian F. Cooper, Adam Silberstein, Erwin Tam, Raghu Ramakrishnan, and Russell Sears. Benchmarking cloud serving systems with ycsb. In Proceedings of the 1st ACM Symposium on Cloud Computing, SoCC '10, page 143154, New York, NY, USA, 2010. Association for Computing Machinery.

[5] Marco Costa, Paolo Ferragina, and Giorgio Vinciguerra. Grafite: Taming adversarial queries with optimal range filters, 2023.

[6] Niv Dayan, Ioana Bercea, Pedro Reviriego, and Rasmus Pagh. Infinifilter: Expanding filters to infinity and beyond. Proc. ACM Manag. Data, 1(2), jun 2023.

[7] Siying Dong, Andrew Kryczka, Yanqin Jin, and Michael Stumm. Rocksdb: Evolution of development priorities in a key-value store serving large-scale applications. ACM Trans. Storage, 17(4), oct 2021.

[8] Tamer Kahveci and Ambuj K. Singh. Variable length queries for time series data. In Proceedings of the 17th International Conference on Data Engineering, page 273282, USA, 2001. IEEE Computer Society.

[9] Eric R. Knorr, Baptiste Lemaire, Andrew Lim, Siqiang Luo, Huanchen Zhang, Stratos Idreos,

and Michael Mitzenmacher. Proteus: A self-designing range filter. In Proceedings of the 2022 International Conference on Management of Data, SIGMOD '22, page 16701684, New York, NY, USA, 2022. Association for Computing Machinery.

[10] Cockroach Labs, 2015.

[11] Siqiang Luo, Subarna Chatterjee, Rafael Ketsetsidis, Niv Dayan, Wilson Qin, and Stratos Idreos. Rosetta: A robust space-time optimized range filter for key-value stores. In Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data, SIGMOD '20, page 20712086, New York, NY, USA, 2020. Association for Computing Machinery.

[12] MongoDB. Wiredtiger storage engine, 2024.

[13] Bernhard MöSSner, Christian Riegger, Arthur Bernhardt, and Ilia Petrov. bloomrf: On performing range-queries in bloom-filters with piecewise-monotone hash functions and prefix hashing, 2022.

[14] Prashant Pandey, Michael A. Bender, Rob Johnson, and Rob Patro. A general-purpose counting filter: Making every bit count. In Proceedings of the 2017 ACM International Conference on Management of Data, SIGMOD '17, page 775787, New York, NY, USA, 2017. Association for Computing Machinery.

[15] Raghu Ramakrishnan and Johannes Gehrke. Database Management Systems. McGraw-Hill, Inc., USA, 3 edition, 2002.

[16] Russell Sears, Mark Callaghan, and Eric Brewer. Rose: compressed, log-structured replication. Proc. VLDB Endow., 1(1):526537, aug 2008.

[17] Kapil Vaidya, Subarna Chatterjee, Eric Knorr, Michael Mitzenmacher, Stratos Idreos, and Tim Kraska. Snarf: A learning-enhanced range filter. Proc. VLDB Endow., 15(8):16321644, apr 2022.

[18] Ziwei Wang, Zheng Zhong, Jiarui Guo, Yuhan Wu, Haoyu Li, Tong Yang, Yaofeng Tu, Huanchen Zhang, and Bin Cui. Rencoder: A space-time efficient range filter with local encoder. In 2023 IEEE 39th International Conference on Data Engineering (ICDE), pages 2036–2049, 2023.

[19] Huanchen Zhang, Hyeontaek Lim, Viktor Leis, David G. Andersen, Michael Kaminsky, Kimberly Keeton, and Andrew Pavlo. Surf: Practical range query filtering with fast succinct tries. In Proceedings of the 2018 International Conference on Management of Data, SIGMOD '18, page 323336, New York, NY, USA, 2018. Association for Computing Machinery.