Team Kingfisher

Team members

Anand Sai Lattala
Nandan Chilukuri

GitHub Repository --> Link

Problem Statement

"Probabilistic Data Structure for Real-Time Event Tracking"

Design a probabilistic data structure that tracks the frequency of millions of real-time events with minimal memory and supports accurate percentile queries.

Constraints

- Sub-linear space complexity while maintaining accuracy.
- Distributed systems demand mergeable data structures.
- Real-time systems need fast updates and queries.

Solution Overview

- Used sketch-based probabilistic methods (e.g., Count-Min Sketch) for tracking frequencies.
- Implement quantile summaries for accurate percentile queries.
- Support merging capabilities to ensure compatibility in distributed systems.

Tech Stack

Language :-

- Python

Algorithm:-

- Count-Min Sketch
- Percentile Calculation

Libraries:-

- Numpy:-Provides efficient handling of numerical operations and array manipulations.
- Random:-Used to simulate random event streams for testing the Count-Min Sketch.
- Hashlib – Used for generating unique and deterministic hash values for events.

Implementation Plan

- Research existing probabilistic data structures (Bloom filters, Count-Min Sketch).
- Design and prototype the data structure.
- Test with simulated real-time data streams. Optimize for accuracy and memory.
- Implement merging features for distributed systems. Document the approach.

Expected Outcomes

- Achieve sub-linear space usage.
- Mergeable data structure for distributed systems.
- Error margin under 5% for frequency tracking.
- Enables scalable and efficient real-time event tracking.

Challenges

- Data accuracy trade-offs due to sub-linear space.
- Integration challenges in distributed environments.

Contingency Plan:

- Use error bounds to calibrate structure size.
- Ensure modular testing for easy debugging.

Conclusion

- Our solution balances memory efficiency, accuracy, and scalability for real-time systems.
- Designed with modern challenges in mind, it’s tailored for high-volume environments.

Thank You!