Efficient Indexing Data Structures for Flash-Based Sensor Devices

SONG LIN
University of California, Riverside
DEMETRIOS ZEINALIPOUR-YAZTI
University of Cyprus
and
VANA KALOGERAKI, DIMITRIOS GUNOPULOS, and WALID A. NAJJAR
University of California, Riverside

Presenter: Xiangnan Xu

Outline

- Introduction
- The Memory Hierarchy
- Problem Definition
- The Data Structures
- Indexing and Searching in MicroHash
- Indexing and Searching in MicroGF
- Experiment Evaluation
- Conclusion

Introduction

- Flash memory is the most prevalent storage medium found on modern wireless sensor devices (WSDs).
- This paper presents two efficient external memory index structures, MicroHash and MicroGF, for efficient retrieval of records stored on local flash memory of WSDs. They all organize data and index pages using a sorted by timestamp file organization.
- Key Idea: eliminate the expensive random access deletions

Introduction

- **MircoHash**: support equality queries in sensor nodes with limited processing capabilities and a low energy budget
- **MicroGF**: support spatial queries in sensor nodes equipped with GPS capabilities

The authors implement these two structures in nesC, the programming language of the TinyOS operating system. And the trace-driven experimentation with several real datasets reveals the excellent index and search capabilities at a small cost of constructing and maintaining the index.

The Memory Hierarchy

- System Architecture of Wireless Sensor
- Overview of Flash Memory
  - Types
  - Distinctive Constraints
    1. Delete Constraints
    2. Write Constraints
    3. Wear Constraints

Problem Definition

- Value-Based Equality Queries (MicroHash, MicroGF)
- Time-Based Range and Equality Queries (MicroHash, MicroGF)
- Spatial Queries (MicroGF)

The Data Structures

- In-Memory (SRAM) Data Structures

➢ Header(8B)

① 3 bits TYP

② 16 bits CRC

③ 7 bits SIZ

④ 23 bits PPA

⑤ 15 bits PWC

➢ Payload(504B)

Types: Root Page, Directory Page, Index Page, Data Page

Indexing and Searching in MicroHash

- Indexing in MicroHash
  - Initialization phase
  - Growing phase
  - Repartition phase
  - Deletion phase

- Searching in MicroHash
  - Searching by value
  - Searching by Timestamp

① LBSearch

② ScaleSearch

Indexing and Searching in MicroHash

- Search Optimization
  - Elf-like chaining (ELC)
  - Two-phase read
  - Lossless Compression by exploiting Temporal Locality

Indexing and Searching in MicroGF

- MicroGF Index Data Structures
  - Directory page data structure
  - Index page data structure
- Indexing in MicroGF
- Searching in MicroGF
- Comparison of methods to index spatial records
  - Grid Files
  - Quadtrees
  - MicroGF

Experimental Evaluation

- Evaluation parameters
- **Space overhead**

  overhead of maintaining the additional index pages, the overhead ratio is defined as follows:

$$\Phi = \frac{IndexPages}{DataPages + IndexPages}$$

- **Search performance**

  the average number of pages accessed for finding the required record
- **Energy consumption** for indexing the data records

Conclusion

- The experiment shows that the structures MicroHash and MicroGF can offer high-performance indexing and searching capabilities in the presence of a low-energy budget, they are both efficient and practical

Thank you!