bring2lite: a structural Concept and Tool for Forensic Data Analysis and Recovery of Deleted SQLite Records

Chr. Meng, H. Baier

Hochschule Darmstadt, CRISP, da/sec Security Group

2019-07-15
Location

Darmstadt
Agenda

Motivation

Background

Structural Analysis on SQLite Record Deletion

Concept to Extract Deleted Content

The Tool bring2lite

Evaluation

Conclusion and Future Work
Motivation

Background

Structural Analysis on SQLite Record Deletion

Concept to Extract Deleted Content

The Tool bring2lite

Evaluation

Conclusion and Future Work
Widespread use
Motivation

Goals of this paper

1. Structural approach:
   - Analyse deletion behaviour of SQLite depending on different database parameters, which affect the erasure of database data
   - Relevant pragmas: secure_delete, auto_vacuum, journal_mode

2. Develop a concept to parse and process deleted SQLite records

3. Provide a proof of concept implementation: bring2lite

4. Evaluation with respect to a common test corpus
Motivation

Background

Structural Analysis on SQLite Record Deletion

Concept to Extract Deleted Content

The Tool bring2lite

Evaluation

Conclusion and Future Work
Background

File format

1. SQLite file organised in pages:
   ▶ Page numbers start with 1
   ▶ Page 1 contains SQLite file header (100 bytes) and root page of sqlite_master table
   ▶ SQLite header defines page size (e.g., 4 KiB)

2. Five different page types:
   ▶ b-tree pages for tables / indexes (interior, leaf pages)
   ▶ Overflow pages
   ▶ Freelist pages
   ▶ Further two forensically irrelevant page types
Pragmas

1. A pragma is a configuration option.

2. Relevant pragmas in the scope of deletion are:

   2.1 secure_delete:
       ▶ Settings 0, 1, FAST
       ▶ secure_delete = 1: deleted content overwritten with zero bytes in database file

   2.2 auto_vacuum: if turned on, deletes unused pages (does not keep them in a freelist)

   2.3 journal_mode:
       ▶ WAL journal (Write-ahead log)
       ▶ Rollback journal file
Table b-tree leaf page:

- **Page Header**: 8 bytes
- **Cell Pointer Array**: N times 2 byte pointer
- **Unallocated Area**
  - Bytes of payload
  - Row ID
  - Serial types
  - Size of this varint and serial types
  - 1 to N varints
- **Freeblock**
  - Size of the deleted cell
  - Variable length
  - Payload
  - Overflow page

- **Cells**
  - Varint
  - Varint
  - Varint
  - 4 byte
Motivation

Background

Structural Analysis on SQLite Record Deletion

Concept to Extract Deleted Content

The Tool bring2lite

Evaluation

Conclusion and Future Work
Goal and scenarios

- **Main goal:** observe how SQLite removes records under different conditions, i.e. to learn the reality of SQLite deletion.

- **Six scenarios, e.g.,**
  - **Scenario S1** Insert 1 record, delete it.
  - **Scenario S2** Insert 3 records, delete 1 record.
  - **Scenario S5** Insert records until a second page will be created, delete all records on the first page.
  - **Scenario S6** Insert records until a second page will be created, delete all records.

- **12 pragma combinations for each scenario:**
  \[6 \cdot 12 = 72\] test cases.

- **Generation of test files on our own.**
Sample results of structural analysis

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>secure_delete=0</th>
<th>auto_vacuum=0</th>
<th>secure_delete=FAST</th>
<th>auto_vacuum=0</th>
<th>secure_delete=1</th>
<th>auto_vacuum=0</th>
<th>secure_delete=1</th>
<th>auto_vacuum=PERIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>S6</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>
Concept

Motivation

Background

Structural Analysis on SQLite Record Deletion

Concept to Extract Deleted Content

The Tool bring2lite

Evaluation

Conclusion and Future Work
Start with processing SQLite database header

- Parse *sqlite_master* table:
  - Its root page is page 1 (directly following the database header)
  - Extract all schemas and connect every database page to its schema

- Parse regular records (i.e. all active database entries)
Parse freeblocks (in an active page):
  ✔ Entry point is page header entry
  ✔ Linked list containing at least 4 bytes:
    ❑ 2 bytes: pointer to subsequent freeblock (if there is any)
    ❑ 2 bytes: length of freeblock
  ✔ Cell header and payload header information partly overwritten
Parse unallocated area (in an active page):
- Start after cell pointer area
- Stop at cell area
- Parse for bytes different from zero bytes
Parse freelist pages:

- Entry point is from SQLite database header entry
- Freelist structure is easily processed
- If available, parse journal file (WAL, rollback journal)
The Tool bring2lite

Motivation

Background

Structural Analysis on SQLite Record Deletion

Concept to Extract Deleted Content

The Tool bring2lite

Evaluation

Conclusion and Future Work
The Tool bring2lite

Overview

- Proof of concept implementation in Python
- Class diagramme:

- Command line oriented
- Interprets SQLite data types (e.g., varints as integers)
- Source freely available via https://github.com/bring2lite/bring2lite
The Tool bring2lite

Sample usage

```bash
~/bring2lite$ python3.6 ./bring2lite/main.py \
   --filename ./db/database.sqlite --out ./results
```

```bash
~/bring2lite$ ls -l results
drwxr-xr-x cm cm 16384 Mr 27 20:34 freeblocks
drwxr-xr-x cm cm 16384 Mr 27 20:34 freelists
drwxr-xr-x cm cm 16384 Mr 27 20:34 regular-page-parsing
drwxr-xr-x cm cm 16384 Mr 27 20:34 schemas
drwxr-xr-x cm cm 16384 Mr 27 20:34 unalloc-parsing
```
Evaluation

Motivation

Background

Structural Analysis on SQLite Record Deletion

Concept to Extract Deleted Content

The Tool bring2lite

Evaluation

Conclusion and Future Work
Evaluation

Based on corpus by Nemetz, Schmitt, Freiling

<table>
<thead>
<tr>
<th>Code</th>
<th>SQL Commands</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-01</td>
<td>create 1, insert 20, drop</td>
</tr>
<tr>
<td>A-02</td>
<td>create 1, insert 20, delete 20, drop</td>
</tr>
<tr>
<td>A-03</td>
<td>create 2, insert 10/each, drop each</td>
</tr>
<tr>
<td>A-04</td>
<td>create 2, insert 10/each, delete 10/1, drop each</td>
</tr>
<tr>
<td>A-05</td>
<td>create 2, insert 10/each, delete 10/each, drop each</td>
</tr>
<tr>
<td>B-01</td>
<td>create 1, insert 10, drop, create 1, insert 5</td>
</tr>
<tr>
<td>B-02</td>
<td>create 3, insert 10/each, drop 1, create 1, insert 5</td>
</tr>
<tr>
<td>C-01</td>
<td>create 1 (int cols), insert 20, delete 7</td>
</tr>
<tr>
<td>C-02</td>
<td>create 2 (int cols), insert 20/each, delete 5/each</td>
</tr>
<tr>
<td>C-03</td>
<td>create 1 (text cols), insert 20, delete 7</td>
</tr>
<tr>
<td>C-04</td>
<td>create 2 (text cols), insert 20/each, delete 5/each</td>
</tr>
<tr>
<td>C-05</td>
<td>create 2 (int/text), insert 20/each, delete 5/each</td>
</tr>
<tr>
<td>C-06</td>
<td>create 1 (float cols), insert 20, delete 7</td>
</tr>
<tr>
<td>C-07</td>
<td>create 2 (float cols), insert 20/each, delete 5/each</td>
</tr>
<tr>
<td>C-08</td>
<td>create 2 (float/text), insert 20/each, delete 5/each</td>
</tr>
<tr>
<td>C-09</td>
<td>create 1 (float/text), insert 10, delete 10</td>
</tr>
<tr>
<td>C-10</td>
<td>create 2 (float/text), insert 10/each, delete 10/each</td>
</tr>
<tr>
<td>D-01</td>
<td>create, insert 10, delete 5, insert 3</td>
</tr>
<tr>
<td>D-02</td>
<td>create, insert 10, delete 5, insert 5</td>
</tr>
<tr>
<td>D-03</td>
<td>create, insert 10, delete 5, insert 10: match 1</td>
</tr>
<tr>
<td>D-04</td>
<td>create, insert 10, delete 5, insert 3: match all</td>
</tr>
<tr>
<td>D-05</td>
<td>create, insert 10, delete 5, insert 5: match all</td>
</tr>
<tr>
<td>D-06</td>
<td>create, insert 10, delete 10, insert 5: match all</td>
</tr>
<tr>
<td>D-07</td>
<td>create 2, insert 10/each, delete 5/1, insert 5/2</td>
</tr>
<tr>
<td>D-08</td>
<td>create 2, insert 10/each (alt), delete 5/1, insert 5/2</td>
</tr>
<tr>
<td>E-01</td>
<td>records overflow like in 07-01, delete 7</td>
</tr>
<tr>
<td>E-02</td>
<td>records overflow like in 07-02, delete 5</td>
</tr>
</tbody>
</table>

Source: S. Nemetz, S. Schmitt, F. Freiling; A standardized corpus for SQLite database forensics; DFRWS EU 2018
## Evaluation

### Evaluation result

<table>
<thead>
<tr>
<th>Case</th>
<th>Undark</th>
<th>SQLite Deleted Records Parser</th>
<th>SQLite Doctor</th>
<th>Stellar Phoenix Repair for SQLite</th>
<th>SysTools SQLite Database Recovery</th>
<th>Sanderson Forensic Browser for SQLite</th>
<th>Sqlite Forensic Explorer</th>
<th>Autopsy SQLite Deleted Records Plugin</th>
<th>bring2lite</th>
</tr>
</thead>
<tbody>
<tr>
<td>0A-01</td>
<td>20/20*</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>20/20</td>
</tr>
<tr>
<td>0A-02</td>
<td>9/20*</td>
<td>20/20*</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>1/20</td>
</tr>
<tr>
<td>0A-03</td>
<td>20/20*</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>20/20</td>
</tr>
<tr>
<td>0A-04</td>
<td>15/20*</td>
<td>10/20*</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>13/20</td>
</tr>
<tr>
<td>0A-05</td>
<td>11/20*</td>
<td>20/20*</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>3/20</td>
</tr>
<tr>
<td>0B-01</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>4/10</td>
<td>4/10</td>
<td>4/10</td>
</tr>
<tr>
<td>0B-02</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>4/10</td>
</tr>
<tr>
<td>0C-01</td>
<td>0/7</td>
<td>0/7</td>
<td>0/7</td>
<td>0/7</td>
<td>7/7</td>
<td>7/7</td>
<td>7/7</td>
<td>7/7</td>
<td>6/7*</td>
</tr>
<tr>
<td>0C-02</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>10/10*</td>
<td>9/10</td>
<td>0/10</td>
<td>8/10*</td>
</tr>
<tr>
<td>0C-03</td>
<td>0/7</td>
<td>0/7</td>
<td>0/7</td>
<td>0/7</td>
<td>2/7</td>
<td>8/10*</td>
<td>4/7</td>
<td>0/7</td>
<td>6/7*</td>
</tr>
<tr>
<td>0C-04</td>
<td>0/10</td>
<td>0/10*</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>1/10*</td>
<td>8/10*</td>
<td>0/10</td>
<td>8/10*</td>
</tr>
<tr>
<td>0C-05</td>
<td>0/10</td>
<td>0/10*</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>10/10*</td>
<td>9/10</td>
<td>0/10</td>
<td>10/10</td>
</tr>
<tr>
<td>0C-06</td>
<td>0/7</td>
<td>0/7</td>
<td>0/7</td>
<td>0/7</td>
<td>0/7</td>
<td>5/7</td>
<td>6/7*</td>
<td>0/7</td>
<td>6/7*</td>
</tr>
<tr>
<td>0C-07</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>10/10*</td>
<td>10/10</td>
<td>0/10</td>
<td>9/10*</td>
</tr>
<tr>
<td>0C-08</td>
<td>0/10</td>
<td>0/10*</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>6/10</td>
<td>0/10</td>
<td>7/10*</td>
</tr>
<tr>
<td>0C-09</td>
<td>0/10</td>
<td>5/10*</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>2/10</td>
<td>0/10</td>
<td>9/10</td>
</tr>
<tr>
<td>0C-10</td>
<td>5/10*</td>
<td>20/20*</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>0/20</td>
<td>2/20</td>
<td>0/20</td>
<td>5/20</td>
</tr>
<tr>
<td>0D-01</td>
<td>0/5</td>
<td>2/5*</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>1/5</td>
<td>0/5</td>
<td>1/5</td>
</tr>
<tr>
<td>0D-02</td>
<td>0/5</td>
<td>1/5*</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>1/5</td>
<td>0/5</td>
<td>1/5</td>
</tr>
<tr>
<td>0D-03</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>1/5</td>
<td>0/5</td>
<td>0/5</td>
</tr>
<tr>
<td>0D-04</td>
<td>0/5</td>
<td>2/5*</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
</tr>
<tr>
<td>0D-05</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
</tr>
<tr>
<td>0D-06</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
</tr>
<tr>
<td>0D-07</td>
<td>0/5</td>
<td>5/5*</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>5/5</td>
<td>0/5</td>
<td>3/5*</td>
</tr>
<tr>
<td>0D-08</td>
<td>0/5</td>
<td>5/5*</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>3/5</td>
<td>0/5</td>
<td>3/5*</td>
</tr>
<tr>
<td>0E-01</td>
<td>3/7</td>
<td>0/7</td>
<td>0/7</td>
<td>0/7</td>
<td>3/7</td>
<td>0/7</td>
<td>0/7</td>
<td>0/7</td>
<td>5/7*</td>
</tr>
<tr>
<td>0E-02</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>0/5</td>
<td>3/5*</td>
</tr>
</tbody>
</table>

| Sum   | 95/278 | 139/278                      | 0/278         | 0/278                            | 0/278                            | 0/278                                 | 73/278                  | 0/278                          | 147/278   |
Evaluation

Sample database: 0B-02 (1/3)

Setting: create 3, insert 10/each, drop 1, create 1, insert 5

$ less 0B-02.sql

PRAGMA page_size=4096;
[REMOVED]

CREATE TABLE users ('id' INT UNSIGNED NOT NULL,
    'name' TEXT NOT NULL, 'surname' TEXT NULL,
    'codeA' INT NULL, 'codeB' FLOAT NULL
);

CREATE TABLE customers ('cid' INT UNSIGNED NOT NULL,
    'cname' TEXT NOT NULL, 'csurname' TEXT NULL,
    'ccodeA' INT NULL, 'ccodeB' FLOAT NULL
);

[REMOVED]
Sample database: 0B-02 (2/3)

- Files 0B-02.sql and 0B-02.db show:
  - Table users written to page 2 (10 entries)
  - Table customers written to page 3 (10 entries)
  - Table supplier written to page 4 (10 entries)
- Table customers dropped, i.e. page 3 released
- Table members written to page 3 (5 entries)
- Deleted content from table customers is in unallocated area of page 3 – if there is any:
  - Cell area in page 3 starts at offset 0x0f4f of page 3
  - Deleted content is starting at offset 0x0ead of page 3
Sample database: 0B-02 (3/3)

bring2lite retrieves these cells from dropped table customers

bring2lite/output$ ls
regular-page-parsing schemas unalloc-parsing

bring2lite/output$ ls unalloc-parsing
3-page.log

bring2lite/output$ less unalloc-parsing/3-page.log
INT,TEXT,TEXT,INT,REAL,
20010,Luisa,Kuhn,-1407291853,4892744407.93914,
20009,Christian,Schulze,527030628,4362154905.38727,
20008,Zoe,Schubert,-603005252,4007666590.16147,
20007,Luca,Scholz,1643805150,1166617011.72898,
+++++++++++++++++++++++++++++

Chr. Meng, H. Baier
bring2lite / 2019-07-15
Conclusion and Future Work

Motivation

Background

Structural Analysis on SQLite Record Deletion

Concept to Extract Deleted Content

The Tool bring2lite

Evaluation

Conclusion and Future Work
Conclusion and Future Work

Scope: recovering deleted SQLite records
Assess commercial and open source tools
bring2lite performs best
Future work:
  - Focus on test cases, where bring2lite performs poorly
  - Improve recovery performance of freeblocks and overflow pages
  - Sharpen the property 'meaningfulness' of recovered data
  - Consider anti-forensic measures
Contact

▶ harald.baier@h-da.de / meng.chr@googlemail.com

▶ Interested in internship at da/sec?