Dynamic Correlation of Digital Forensics Reports

Christoph Beckmeyer
Aachen University of Applied Sciences
christoph.beckmeyer@alumni.fh-aachen.de
Presentation Overview

Problem description

Proposed solution

Demo

Limitations & Future Work
Project started 2012

by Martin Pfeiffer as his Bachelor thesis.

Vision of a Research Tool.

Prototypical Database Design for Correlation.

Importer for Physical Analyzer 2.
Correlation in Digital Forensics

- Mobile Phone(s)
- Computer(s)
- Computer Forensic Tools (Sleuthkit, etc)
- Physical Analyser
- XRY
- Correlation
- Analysis Software
- Governmental Databases (e.g., rsCase)
1st Level Correlation:

Joining of sets of artifacts with same semantic from different sources.

2nd Level Correlation:

Establishing semantic relationships between individual artifacts.
Artifacts

The diagram illustrates the relationship between artifacts and their attributes. Each artifact is connected to multiple attributes, and each attribute can have several values. The diagram shows how the values of the attributes are linked to each other, indicating a dynamic correlation in digital forensics.
Use cases for correlation:

Case evidence spread across devices.

Extraction tools have different capabilities.

Organizations use different tools.
Correlation as a Problem

Many existing file formats

New file formats emerge

Existing formats change (CDR)

Across formats: different syntax for same semantics
Practitioners’ current Solutions

Excel

Often manual labor.

Custom Development

Longer development cycle.

Commercial Analysis Tools

Expensive.
### Reviewed tools:

<table>
<thead>
<tr>
<th>Tools</th>
<th>1st Level Correlation</th>
<th>2nd Level Correlation</th>
<th>Runtime Extensibility</th>
<th>Cross Device Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>XIRAF (Alink et al.)</td>
<td>Yes, wrapped with XML</td>
<td>Yes, predefined</td>
<td>Yes (Common Lisp)</td>
<td>Yes</td>
</tr>
<tr>
<td>FACE (Case et al.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zeitline (Buchholz)</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIC (Osborne et al.)</td>
<td>Yes</td>
<td></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ECF (Chen et al.)</td>
<td>Yes</td>
<td></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Rich Event Representation</td>
<td>Yes</td>
<td></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>(Schatz et al.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excel</td>
<td>Yes</td>
<td></td>
<td>Yes (VBScript)</td>
<td>Yes</td>
</tr>
<tr>
<td>Analysts Notebook</td>
<td>Yes</td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Physical Analyzer</td>
<td>No</td>
<td>Yes (but failed to use)</td>
<td></td>
<td>Yes</td>
</tr>
</tbody>
</table>
Previous Lessons from DIRECT:

Embedding into existing tool failed.

Modeling domain into RDBMS was inflexible.
Conclusions from current solutions:

Practitioners are in need for flexible tool to:

Quickly correlate by themselves. (Cut development cycle)

Current tools often correlate into static model.

Hinders fast adaptation to change in inputs.

Flexible correlation is possible with commercial tools.

But they might not support your use case. (e.g., comparison of inputs)
Proposed Solution

Introduction of Abstractions

1. Normalizing the file format.

2. Correlation using Drag’n Drop.

3. Normalization by attaching functions.
Introducing Abstractions

- 2nd Level Correlation
- Queries
- Value Normalization using Functions
- 1st Level Correlation using Drag'n'Drop
- Normalizing the File Format

Current scope
1. Normalizing the File Format

```xml
<model type="Call" id="82a...."
    deleted_state="Intact">
    <field name="Name" type="String">
        <empty />
    </field>
    <field name="Type"
        type="CallType">
        <value type="CallType"> <![CDATA[Outgoing]]> </value>
    </field>
    <field name="TimeStamp" type="TimeStamp">
        <value type="TimeStamp">2005-09-13T09:11:34+02:00</value>
    </field>
</model>
```

```
<view name="Anrufe">
    <item>
        <field name="Art">
            value="Entgegengenommen"
            class="STATUS" />
        <field name="Zeit">
            value="13.10.2013 11:37:30 (Gerät)"
            class="TIME" />
    </item>
</view>
```
Merging artifacts

2. Correlation using Drag’n Drop
2. Correlation using Drag’n Drop

Modeling the domain

Fusion of XRY and Physical Analyzer

Model conversations
3. Value Normalization

Normalization of values

Call

PA2

XRY

Case

Import

Artifact

Attribute

Value

Normalization of values

Wednesday, May 14, 14
3. Value Normalization

Attachable to Attributes

Examples:

- Date formats
- Phone numbers
- Matching against known files
Demo
Limitations

Better Graphical User Interface.

More functions.

Tracing to the original file.

Queries on the correlation result.
Future Work

Automatic Matching

On artifact / attribute names

On attributes values

2nd Level Correlation

Semantic Network

Deduction of relations
Thank you

Dynamic Correlation with DIRECT

1. Normalizing the file format.
2. Drag’n Drop correlation for artifacts & attributes.
3. Function library for value normalization.

Christoph Beckmeyer
chrisb@alumni.fh-aachen.de
www.it-forensik.fh-aachen.de/projekte/direct