Malware Narratives

Introduction

Version 1.0
Prerequisites

Interest in software diagnostics and malware analysis
Why?

- Communication language
- Malware diagnostics as software diagnostics
- Big DA+TA (Dump Artifacts + Trace Artifacts)
Software Diagnostics

A discipline studying abnormal software structure and behavior in software execution artifacts (such as memory dumps, software and network traces and logs) using pattern-driven, systemic and pattern-based analysis methodologies.
Diagnostics Pattern

A common recurrent identifiable problem together with a set of recommendations and possible solutions to apply in a specific context.
Pattern Orientation

Pattern-driven

- Finding patterns in software artefacts
- Using checklists and pattern catalogs

Pattern-based

- Pattern catalog evolution
- Catalog packaging and delivery
Catalog Classification

- By abstraction
  Meta-patterns

- By artifact type
  Software Log*  Memory Dump  Network Trace*

- By story type
  Problem Description  Software Disruption  UI Problem

- By intention
  Malware
Malware

Software that uses planned alteration of structure and behavior of software to serve malicious purposes.
Memory Analysis Patterns

Software Diagnostics

Memory Dump Analysis Patterns

Malware Analysis Patterns
Traces and Logs

- Windows logs and patterns
- Linux logs and patterns
- Event logs and patterns
- Mac OS X logs and patterns
- Web Server logs and patterns
- Network traces and patterns
- GUI logs and patterns
Trace and Log Patterns

Software Trace and Log Analysis Patterns

- Windows logs
- Event logs
- Linux logs
- Mac OS X logs
- Web Server logs
- Network traces
- GUI logs

© 2013 Software Diagnostics Services
Software Narrative

A temporal sequence of events related to software execution.
Narrative Taxonomy

- Incident stories
- **Software traces and logs**
- Malware analysis stories
Malware Narrative Patterns
Software Log

- A sequence of formatted messages
- Arranged by time
- A narrative story
Minimal Log Graphs

<table>
<thead>
<tr>
<th>No</th>
<th>Module</th>
<th>PID</th>
<th>TID</th>
<th>Date</th>
<th>Time</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ModuleA</td>
<td>4280</td>
<td>1736</td>
<td>5/28/2012</td>
<td>08:53:50.496</td>
<td>Trace message 1</td>
</tr>
<tr>
<td>2</td>
<td>ModuleB</td>
<td>6212</td>
<td>6216</td>
<td>5/28/2012</td>
<td>08:53:52.876</td>
<td>Trace message 2</td>
</tr>
</tbody>
</table>

[...]
Pattern-Based Analysis

Usage → Software Trace → Discovery

Pattern Catalog + New Pattern
Pattern Classification

- Vocabulary
- Error
- Trace as a Whole
- Large Scale
- Activity
- Message
- Block
- Trace Set
Reference and Course

- Free catalog
  Software Log Analysis Patterns

- Free reference graphical slides
  Accelerated-Windows-Software-Trace-Analysis-Public.pdf

- Training course*
  Accelerated Windows Software Trace Analysis

* Available as a full color paperback book, PDF book, on SkillsSoft Books 24x7. Recording is available for all book formats
Vocabulary Patterns

- Basic Facts*
- Vocabulary Index

* patterns marked with yellow color are most likely to be useful for malware detection and analysis
Error Patterns

- Error Message
- Exception Stack Trace
- False Positive Error
- Periodic Error
- Error Distribution
Trace as a Whole

- Partition
- Circular Trace
- Message Density
- Message Current
- Trace Acceleration
- No Trace Metafile
- Empty Trace
- Missing Module
- Guest Module

- Truncated Trace
- Visibility Limit
- Sparse Trace
Large Scale Patterns

- Characteristic Block
- Background Modules
- Foreground Modules
- Layered Periodization
- Focus of Tracing
- Event Sequence Order
- Trace Frames
## Characteristic Block

<table>
<thead>
<tr>
<th>Time</th>
<th>PID TID Time</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Foreground Modules
Focus of Tracing

Activity regions: $J_m^1$, $J_m^2$, $J_m^3$
Activity Patterns

- Thread of Activity
- Adjoint Thread of Activity
- No Activity
- Activity Region
- Discontinuity
- Time Delta
- Glued Activity
- Break-in Activity
- Resume Activity
- Data Flow
Thread of Activity
Adjoint Thread of Activity

Time

# PID TID Time Func Message

Time

# PID TID Time Func Message
Activity Region

Message current: $J_{m2} > \max (J_{m1}, J_{m3})$
Glued Activity

ATID: Adjoint Thread ID

ImageA
ATID 2

ImageB
ATID 3

Trace Session 1

Trace Session 2
Break-in Activity

Discontinuity
Message Patterns

- Significant Event
- Defamiliarizing Effect
- Anchor Messages
- Diegetic Messages
- Message Change
- Message Invariant
- UI Message
- Original Message
- Implementation Discourse
- Opposition Messages

- Linked Messages
- Gossip
- Counter Value
- Abnormal Value*
- Message Context
- Marked Messages
- Incomplete History
- Message Interleave
- Fiber Bundle

* added recently
Defamiliarizing Effect
Abnormal Value
Marked Messages

Annotated messages:

network activity [+]
process A launched [+]
process B launched [-]
process A exited [-]

[+] activity is present in a trace
[-] activity is undetected or not present
Fiber Bundle

Trace messages

I/O stack

Thread stack trace
Block Patterns

- Macrofunction
- Periodic Message Block
- Intra-Correlation
Periodic Message Block
Trace Set Patterns

- Master Trace
- Bifurcation Point
- Inter-Correlation
- Relative Density
- News Value
- Impossible Trace
- Split Trace
Master Trace
Inter-Correlation

Logging Tool

System

Logging Tool

Log File

Log File
# Impossible Trace

<table>
<thead>
<tr>
<th>#</th>
<th>Module</th>
<th>PID</th>
<th>TID</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>1001</td>
<td>ModuleA</td>
<td>202</td>
<td>404</td>
<td>foo: start</td>
</tr>
<tr>
<td>1002</td>
<td>ModuleA</td>
<td>202</td>
<td>404</td>
<td>foo: end</td>
</tr>
</tbody>
</table>

```c
void foo()
{
    TRACE("foo: start");
    bar();
    TRACE("foo: end");
}

void bar()
{
    TRACE("bar: start");
    // some code ...
    TRACE("bar: end");
}
```
Grand Unification

- Narrative and Trace
  \[ N : T \rightarrow M \]

- Generalized Narrative and Trace
  \[ GN : A \rightarrow M \]
  \[ GN_3 \circ GN_2 \circ GN_1 : M \rightarrow M \rightarrow M \]
Further Reading

- Software Diagnostics Institute
- Memory Dump Analysis Anthology: Volumes 3, 4, 5, 6, ...
  Volume 7 is in preparation (April, 2013)
  Volume 8 is planned for November, 2013
- Introduction to Software Narratology
- Accelerated Windows Software Trace Analysis
What’s Next?

Pattern-Oriented Network Trace Analysis
Q&A

Please send your feedback using the contact form on DumpAnalysis.com
Thank you for attendance!