Security & Assured Autonomy

Howie Shrobe
MIT CSAIL
Assured Autonomy

The expectation that a system entrusted to make decisions on its own, will with high probability:

• Make good enough decisions to achieve (most of) its goals without harm
• Be able to justify those decisions

• A secure system is one that will behave as designed and implemented even when under attack
  • Normally thought of as Confidentiality, Integrity and Availability
  • For Cyber-Physical systems must also include predictable timing

• Security is necessary condition for Assured Autonomy
Cars are (unsafe) rolling computers
Full Stack Security is critical

• Vulnerabilities are exploited at many levels:
  • Processor
  • Software
  • Interconnect

• Interconnecting components can amplify the problem:
  • The CANBUS is insecure
  • It was connected to the entertainment/control panel systems for convenience
  • Vulnerabilities in the entertainment system now are vulnerabilities in the control system
An Autonomous System Stack

- Application Specific Models and Code
  - Tensorflow, SOAR, ...
- Planning & Control
- AI Libraries & Frameworks
  - OpenCV, JPEG, ...
- Supporting Libraries
  - LibC, LibPython, ...
- Language Runtime Support
- Compiler & Development Env
  - GCC, LLVM, ...
- Network Hardware and Software Stack
  - 802..., TCP/IP, MODBUS, CANBUS, ...
- OS
- Rest of Hardware
  - Linux, FreeRTOS, ...
- Processor
  - Memory, Interconnect, GPU, Peripherals,
    Intel, AMD, ARM, ...
# Tensorflow vulnerability list

<table>
<thead>
<tr>
<th>Advisory Number</th>
<th>Type</th>
<th>Versions affected</th>
<th>Reported by</th>
<th>Additional Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFSA-2019-001</td>
<td>Null Pointer Dereference Error in Decoding GIF Files</td>
<td>&lt;= 1.12</td>
<td>Baidu Security Lab</td>
<td></td>
</tr>
<tr>
<td>TFSA-2018-006</td>
<td>Crafted Configuration File results in Invalid Memory Access</td>
<td>&lt;= 1.7</td>
<td>Blade Team of Tencent</td>
<td></td>
</tr>
<tr>
<td>TFSA-2018-005</td>
<td>Old Snappy Library Usage Resulting in Memcpy Parameter Overlap</td>
<td>&lt;= 1.7</td>
<td>Blade Team of Tencent</td>
<td></td>
</tr>
<tr>
<td>TFSA-2018-004</td>
<td>Checkpoint Meta File Out-of-Bounds Read</td>
<td>&lt;= 1.7</td>
<td>Blade Team of Tencent</td>
<td></td>
</tr>
<tr>
<td>TFSA-2018-003</td>
<td>TensorFlow Lite TOCO FlatBuffer Parsing Vulnerability</td>
<td>&lt;= 1.7</td>
<td>Blade Team of Tencent</td>
<td></td>
</tr>
<tr>
<td>TFSA-2018-002</td>
<td>GIF File Parsing Null Pointer Dereference Error</td>
<td>&lt;= 1.5</td>
<td>Blade Team of Tencent</td>
<td></td>
</tr>
<tr>
<td>TFSA-2018-001</td>
<td>BMP File Parser Out-of-bounds Read</td>
<td>&lt;= 1.6</td>
<td>Blade Team of Tencent</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>Out Of Bounds Read</td>
<td>&lt;= 1.4</td>
<td>Blade Team of Tencent</td>
<td>issue report</td>
</tr>
</tbody>
</table>
Full Stack Security (CSAIL, Draper, MIT-LL)

- Memory and Type Safe: E.g. RUST, ...
- Hardware enforced compartmentalization into minimal sized components
The Underlying Problem:

- Computer hardware and many language runtimes don’t represent or properly manage important semantic distinctions
- At runtime, there’s nothing but “Raw Seething Bits”
The solution

• Tag all data with meta-data
• State policies in the form of rules referencing the meta-data
• Enforce the policies at most efficient & lowest level possible to guarantee complete mediation

• Example 1: Memory Safety
  • All pointers includes bounds information
  • Attempts to reference outside the bounds are trapped

• Example 2: Type safety
  • All objects are typed
  • Operations can only be performed on relevant object types (e.g. you can’t execute a string)
Summary

• If we want to have any confidence in an autonomous system then its necessary (but not sufficient) to guarantee that vulnerabilities cannot be exploited to change the reasoning of the autonomy software
  • All data must carry meta-data
  • Policies related to the meta-data must be systematically enforced
  • Enforcement must be done across all levels
  • A “belt and suspenders” approach is necessary
  • Prevention + Containment