Security/Privacy & Assured Autonomy

Howie Shrobe
MIT CSAIL
Panelists

• Tony Dahbura – Johns Hopkins
• Jeremy Daily -- Colorado State
• Greg Falco – MIT & Stanford
• Ryan Gerdes – Virginia Tech
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 causing (much) harm
- Be able to explain & justify those decision

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, etc.

Security is necessary condition for Assured Autonomy
Cars are (unsafe) rolling computers
An Autonomous System Stack

Autonomy Specific Models and Code
- Tensorflow, SOAR 

AI Libraries & Frameworks
- Tensorflow, SOAR 

Supporting Libraries
- OpenCV, JPEG, 
- LibC, LibPython,
- GCC, LLVM,

Language Runtime Support
- 802..., TCP/IP, MODBUS, CANBUS...

Compiler & Development Env
- Linux, FreeRTOS,

Network Hardware and Software Stack
- Memory, Interconnect, GPU,

OS
- Intel, AMD, ARM...

Rest of Hardware

Processor
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>
</tbody>
</table>
Summary

• If we want to have any confidence in an autonomous system then it's 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