Scission: Signal Characteristic-Based Sender Identification and Intrusion Detection in Automotive Networks

1st ITG Workshop on IT Security (ITSec)
University of Tübingen
April 2, 2020

Marcel Kneib(1), Christopher Huth(2), Paul Duplys(2)

(1) Bosch Engineering GmbH,
(2) Robert Bosch GmbH
Introduction

- Attacks on vehicles...
  - on the rise due to increased connectivity features
  - may be highly scalable
  - result in threats for humans and the environment
- Demonstrated by Miller and Valasek [31]

Introduction

- Controller Area Network widely used for in-vehicle communication
  - 500 kb/s bandwidth
  - 64 bit payload
  - No sender authenticity

- Message Authentication Codes hard to apply

- Intrusion Detection Systems
  - Signatures
  - Anomalies
  - Physical properties
    - Clock drifts [4]
    - Variations in the analog signal [33, 6]


Scission Overview

Sampling → Preprocessing → Feature Extraction → Classification → Detection

- **Sampling**
  - Raw data

- **Preprocessing**
  - Filters, normalization

- **Feature Extraction**
  - Mean (mean(x))
  - Skewness (skew(x))
  - Kurtosis (kurt(x))
  - Variance (var(x))

- **Classification**
  - Model

- **Detection**
  - Alarm
  - OK
Sampling and Preprocessing

- Sampling differential signal (20 MS/s)
- Cluster symbols based on its signal shape
Feature Selection

► Statistical features (time, frequency) individual for each group
  ► Mean, Standard Deviation, Variance, Skewness, ...

![Graph showing voltage over time for two ECU groups, ECU 0 and ECU 1.]

<table>
<thead>
<tr>
<th></th>
<th>Concatenated</th>
<th>Rising</th>
<th>Falling</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ECU 0</strong></td>
<td>1.286 V</td>
<td>1.623 V</td>
<td>0.289 V</td>
<td>1.947 V</td>
</tr>
<tr>
<td><strong>ECU 1</strong></td>
<td>1.285 V</td>
<td>1.691 V</td>
<td>0.275 V</td>
<td>1.890 V</td>
</tr>
<tr>
<td><strong>Difference</strong></td>
<td>0.001 V</td>
<td>0.068 V</td>
<td>0.014 V</td>
<td>0.057 V</td>
</tr>
</tbody>
</table>
Model Generation and Classification

- Logistic Regression

- Supervised learning with 200 frames per ECU

- Initial training in safe environment
  - Initiated by secure diagnostic access
  - Key between ECUs and Scission assigned

- Performance Monitoring (aging, corrosion, ...)
  - Probabilities of each ECU
  - Online adaption of the classifiers
  - MAC supported adaption/learning
    - AUTOSAR Secure Onboard Communication (SecOC)

<table>
<thead>
<tr>
<th>ECU 0</th>
<th>ECU 1</th>
<th>ECU 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>95 %</td>
<td>3 %</td>
<td>2 %</td>
</tr>
</tbody>
</table>
Intrusion Detection

- Sender identification based on the highest probability

- In-vehicle communication is static
  - Each identifier is used by only one ECU
  - Alarm if an identifier is used by a invalid ECU

- False positives
  - Due to interferences (start of a strong consumer)
  - Alarm if probability of invalid ECU exceeds threshold $t_{\text{max}}$ (e.g. 70 %)
  - Leads to a higher false negative rate

<table>
<thead>
<tr>
<th></th>
<th>ECU 0</th>
<th>ECU 1</th>
<th>ECU 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>95 %</td>
<td>3 %</td>
<td>2 %</td>
<td></td>
</tr>
<tr>
<td>2 %</td>
<td>98 %</td>
<td>0 %</td>
<td></td>
</tr>
<tr>
<td>49.9 %</td>
<td>50.1 %</td>
<td>0 %</td>
<td></td>
</tr>
</tbody>
</table>
Evaluation

<table>
<thead>
<tr>
<th></th>
<th>ECU</th>
<th>Frames</th>
<th>Avg. accuracy</th>
<th>Min. accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prototype</td>
<td>10</td>
<td>56,560</td>
<td>99.9 %</td>
<td>99.58 %</td>
</tr>
<tr>
<td>Fiat</td>
<td>6+2</td>
<td>25,979</td>
<td>99.6 %</td>
<td>98.56 %</td>
</tr>
<tr>
<td>Porsche</td>
<td>6+2</td>
<td>6,389</td>
<td>99.88 %</td>
<td>99.58 %</td>
</tr>
</tbody>
</table>

99.85% Identification rate → FP after 666 frames → threshold $t_{max}$

<table>
<thead>
<tr>
<th>Predicted</th>
<th>No attack</th>
<th>Attack</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prototype</td>
<td>100 %</td>
<td>0 %</td>
</tr>
<tr>
<td>Attack</td>
<td>1.5 %</td>
<td>98.5 %</td>
</tr>
<tr>
<td>Fiat</td>
<td>100 %</td>
<td>0 %</td>
</tr>
<tr>
<td>Attack</td>
<td>0 %</td>
<td>100 %</td>
</tr>
<tr>
<td>Porsche</td>
<td>100 %</td>
<td>0 %</td>
</tr>
<tr>
<td>Attack</td>
<td>3.18 %</td>
<td>96.82 %</td>
</tr>
</tbody>
</table>
Conclusion

- Sender identification based on physical properties of CAN signals
- Reduction in the necessary hardware requirements
- Evaluated on series production vehicles
  - High identification rate
  - No false positives
- Scission can improve the security of modern vehicles
  - IDS extension
  - Additional security functionality for gateways
  - Standalone system
- Outlook
  - Further reduction of hardware/performance requirements
  - Implementation on an embedded platform
THANK YOU!

Marcel.Kneib@de.bosch.com

Christopher.huth@de.bosch.com
Stability

- Characteristics remain unchanged over several months [33]

- Fiat under changing conditions
  1. Measurement (includes training)
     - Engine off | 25°C (77°F) | 3369 frames | 100% identification
  2. Measurement
     - Driving 30 min. | 32°C (89.6°F) | 6672 frames | 100% identification
  3. Measurement (3 hours of cooling at 23°C (73.4°F))
     - Driving 20 min. | 36°C (96.6°F) | 4863 frames | 100% identification

- Biggest change in the voltage level between 0.012V and 0.026V

Reaction on intrusion

▶ Warn the driver

▶ Log the attack

▶ Prevent the attack
  ▶ Invalidation of the CRC
  ▶ Error Frame

▶ Send the detected attack to Cloud-IDS
  1. Analyze the attack
  2. Update the in-vehicle Signature-based IDS
  3. Find the vulnerability
  4. Update the vulnerable ECU
Additional / Unknown ECU

- Lower threshold $t_{\text{min}}$ (e.g. 30%)
- Counter for each ECU
  - Increment if an unexpected ECUs probability $>$ 30% but $<$ 70%
  - Decrement if expected ECU $>$ 30%
- Additional ECU (connected to the bus after training)
  - Counter of several ECUs will rise (no frames are necessary)
- Unknown ECU (connected but not considered during learning)
  - Detection like normal attack or
  - Counter of the faked ECU will rise
Scission-aware Attacker

- Influencing all ECUs (draining battery)
  - Quick and significantly → System maybe inactive during model adaption
  - Slow → System adapts model continuously

- Influencing its own signal (heating up / cooling down) to impersonate another ECU
  - No information about its own or the signal of the other ECU
  - Several signal characteristics must be similar
  - Precise adaption must be possible