SCISSION

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

Marcel Kneib and Christopher Huth

CCS 2018

Presented by

Alokparna Bandyopadhyay

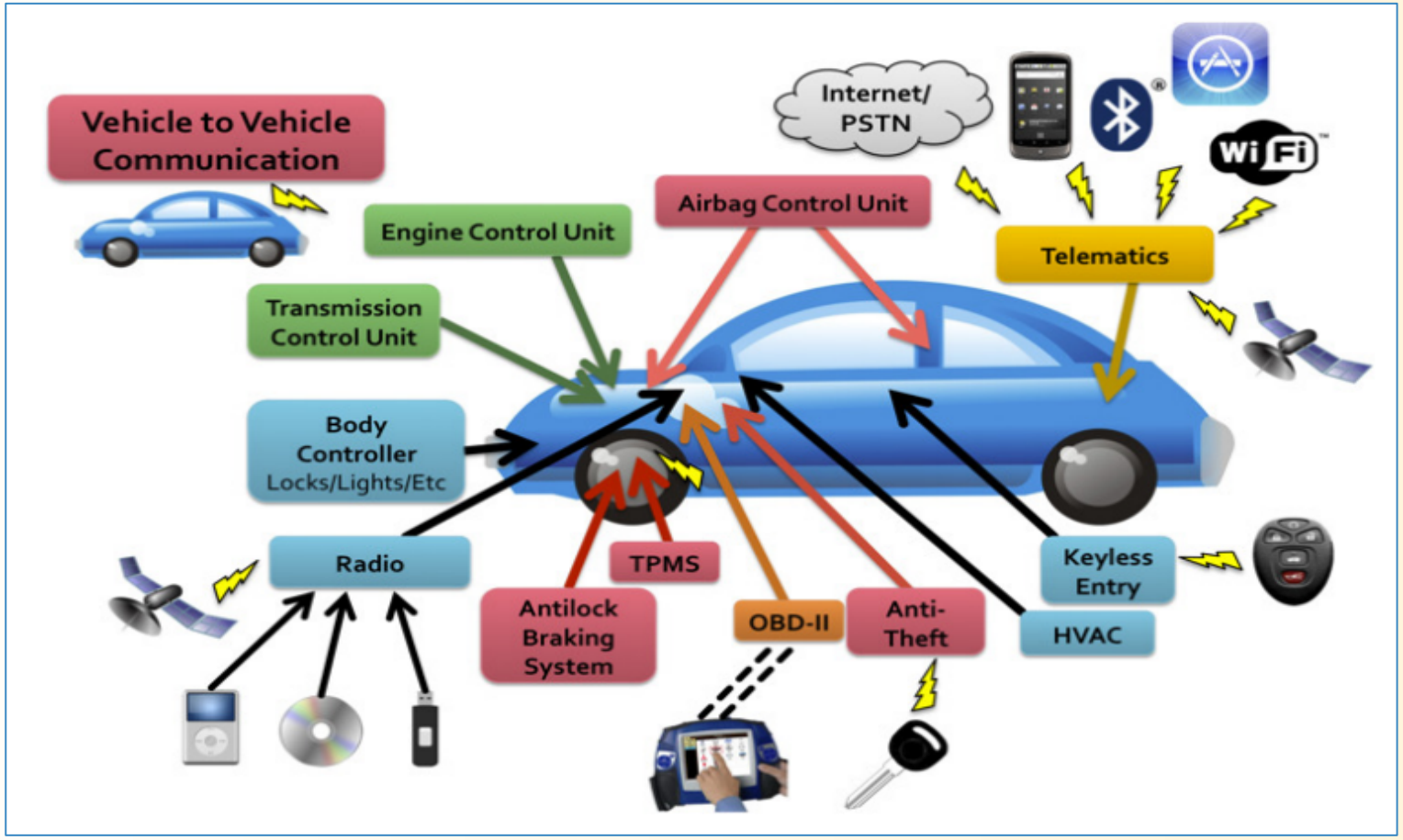
Fall 2018, Wayne State University

Overview

- Introduction
- Control Area Network (CAN)
- System and Threat Model
- SCISSION
- Evaluation
- Discussion & Conclusion

Introduction

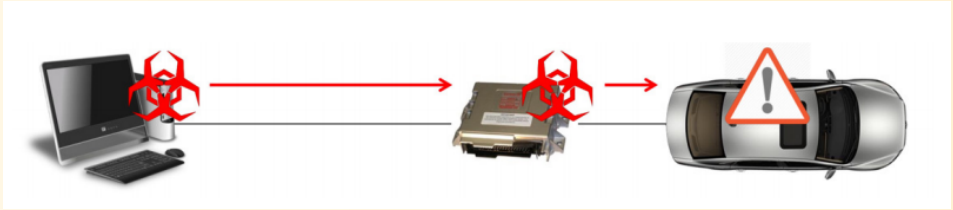
Automotive Components of a Modern Car

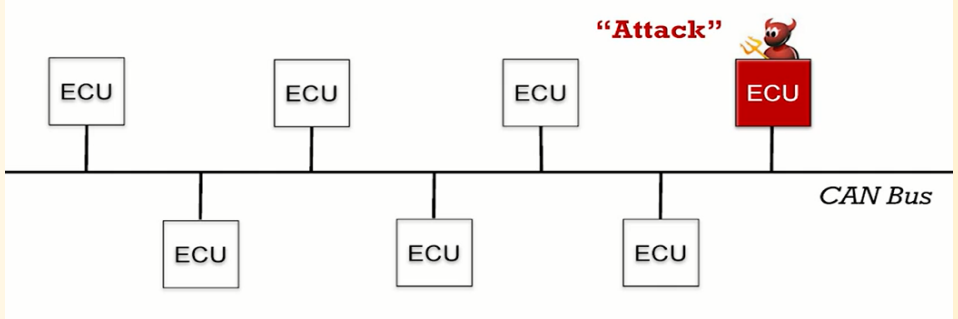


Increased connectivity in connected vehicles

Security Concerns

- Modern cars with remote and/or driverless control has various remote connections (e.g. Bluetooth, Cellular Radio, WiFi, etc.)
  - Attackers exploit remote access points to compromise ECUs in the network
  - Remotely control or even shut down a vehicle
  - No security features in most in-vehicle networks (e.g. CAN Bus)
  - Attacker identification and authentication not possible



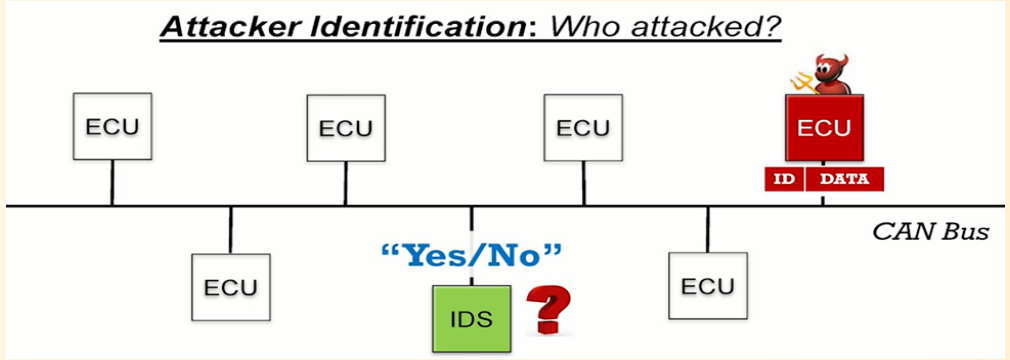


Defense against Attacks

- Efficient Intrusion Detection Systems (IDS) are proposed in the past to identify presence of an attack
  - *Signature Based*: Detects known attack based on their message pattern and content
    - Problem: Difficult to deploy due to lack of data
  - *Anomaly Based*: Expected characteristics are explicitly specified to detect unknown attacks
    - Problem: False Positives

Motivation for Scission

- Attacker Identification is essential
  - Forensic isolation of attacker
  - Vulnerability removal
  - Faster compared to software updates
  - Economic compared to manufacturer recall
- Difference in CAN signals can be used as fingerprints
- Can be used for smart sensors with low computational capacity
- Difficult for remote attackers to circumvent such systems



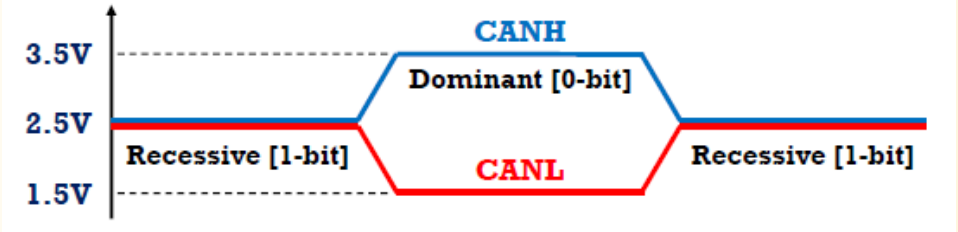
Contribution of Scission

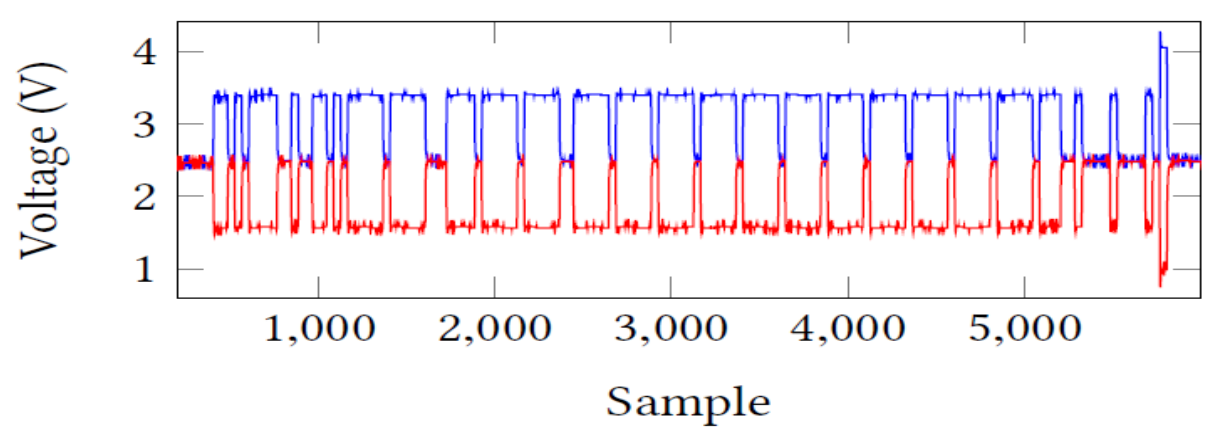
- Uses immutable physical properties of CAN signals as *fingerprints* to identify the sender of CAN messages
- Detect unauthorized messages from compromised, unknown or additional ECUs
- High detection rate with minimal false positives
- No additional computation required
- Does not reduce bandwidth and requires low resources
- Cost effective feasibility

Control Area Network (CAN)

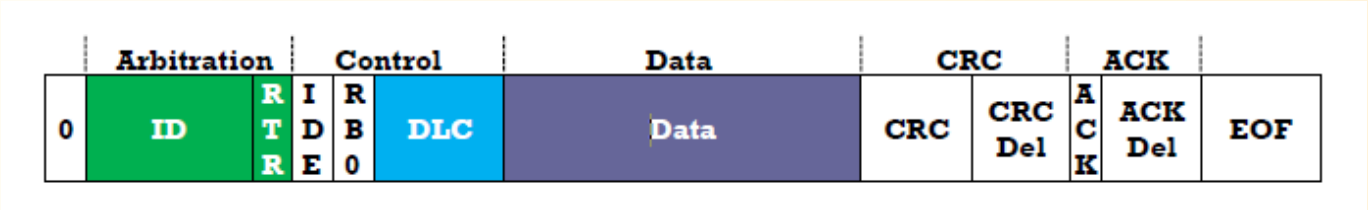
CAN Signal

CAN transceivers have two dedicated CAN wires: CAN High (blue) and CAN Low (red)





CAN Data Frame



Format of a standard CAN data frame

- Data transmitted – 8 bytes of payload
- Frames contain unique ID based on priority and meaning of data
- Node address is not present
- Several bus participants try to access the broadcast bus simultaneously
- Only one ECU can broadcast at a time based on the priority of its identifier

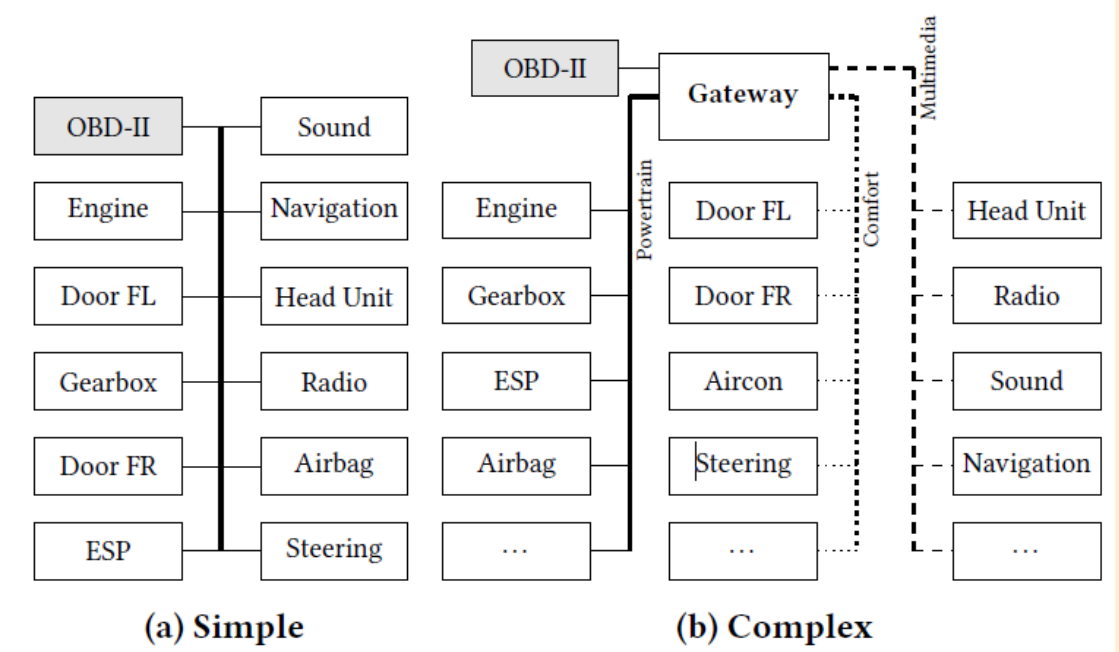
Signal Characteristics

- Sources of signal characteristics for extraction of CAN fingerprints:
  - Variations in supply voltages
  - Variations in grounding
  - Variations in resistors, termination and cables
  - Imperfections in bus topology causing reflections

System and Threat Model

System Model

- In-vehicle protocol used: *CAN Bus*
- Network of several separate CAN Buses with several ECUs connected to each
- In-vehicle network architecture
  - *Simple*: Fewer buses, less secure
  - *Complex*: ECUs separated according to functionality, individual buses connected through gateways with additional security mechanisms



System Model cont.

- Scission is physically integrated into the network via additional ECU
- Scission ECU is secured and trustworthy
- System cannot be bypassed by an attacker
- Gateways can be used to determine whether received messages have been sent from valid ECUs

Threat model

- Compromised ECU
  - Attackers access the monitored CAN through an exploited vulnerability of an existing ECU
  - Remotely and stealthily send a variety of CAN frames using all possible identifiers and any message content
- Unmonitored ECU
  - Malicious usage of a passive or unmonitored device
  - Exploit ECU update mechanism
  - Insert malicious code and turn a passive, listening-only device into a message sending device

Threat model cont.

- Additional ECU
  - Attach an additional bus participant directly to the guarded network or use the easy-to-reach On-board diagnostics (OBD)-II port of the vehicle
  - Physical access to the vehicle to control the vehicle maneuver
- Scission-aware Attacker
  - Remote attacker attempts to mislead the IDS by influencing its signal characteristics
  - Affects the absolute voltage level of the signals

Security Goal

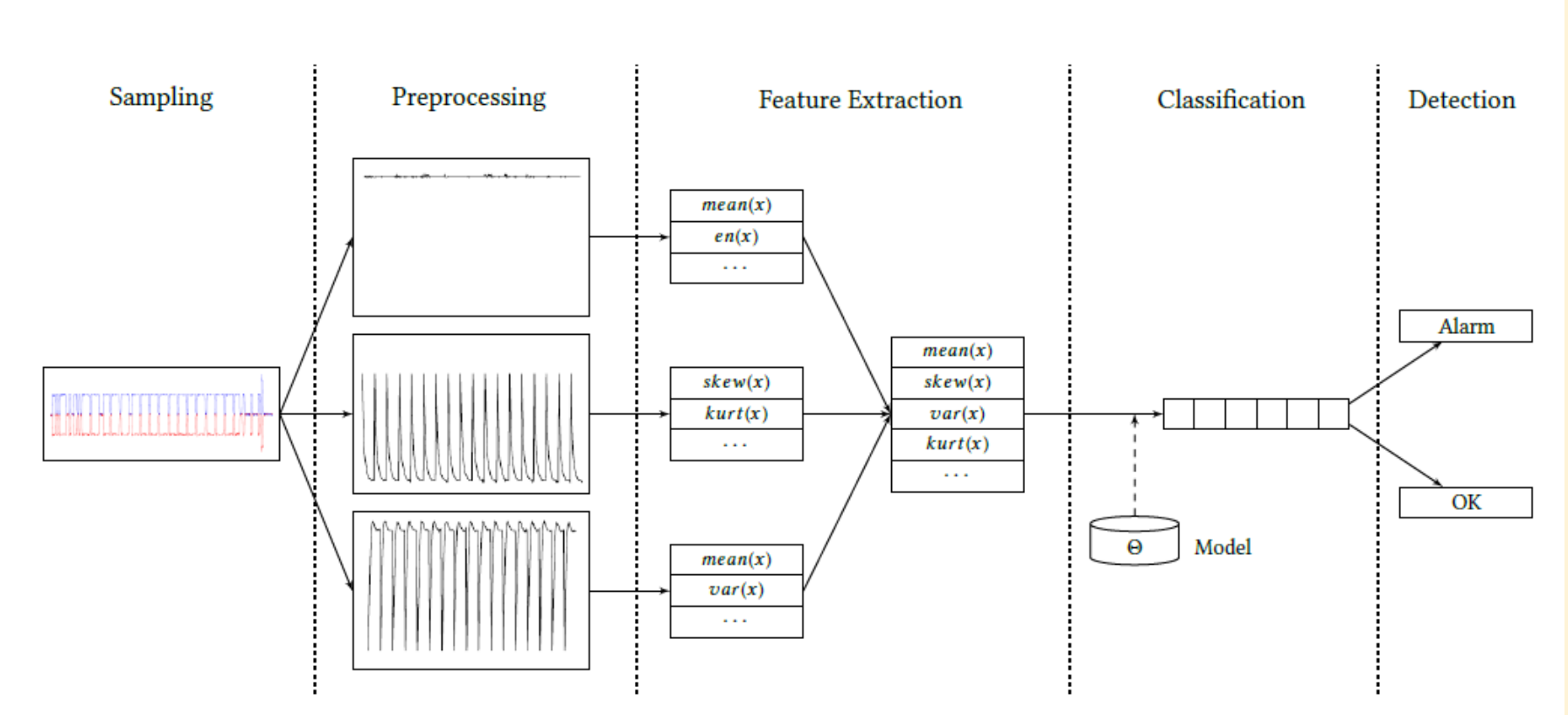
- CAN provides no security mechanism to identify an attacker
- Scission determines signal characteristics to create fingerprints for source ECUs
- System monitors network traffic to detect unauthorized messages from compromised, unknown or additional ECUs
- System detects
  - Counterfeit CAN frames from compromised and unknown ECUs
  - Remotely compromised ECUs

SCISSION

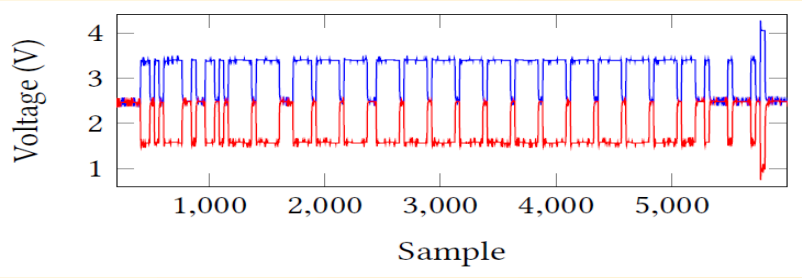
Signal Characteristic-Based Sender Identification

Overview of Scission

Scission fingerprints ECUs and achieves attacker identification in five phases



Phase 1: Sampling



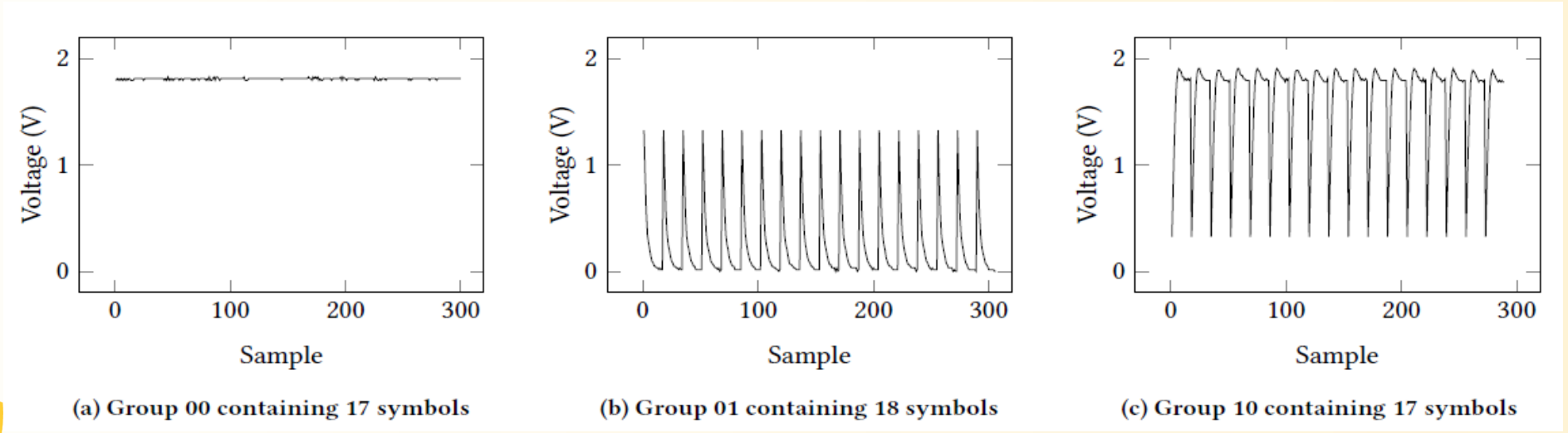
- Analog signals of the received frames are recorded
- Differential signal is used directly
  - Requires an additional circuit
  - System requires fewer resources because less data is stored temporarily
  - Signal noise can be compensated
  - Number of measured values per bit depends on the sampling and baud rate
- Separate signals are used
  - Can be influenced by electromagnetic interference or other variations
  - Incorrect predictions due to signal noise

Phase 2: Preprocessing

- Signal of each bit of the message recorded in sampling stage is processed individually
- Sets containing several analog values are subsequently divided into 3 groups
  - Group $G\downarrow 10$ – Set representing a dominant bit (0), contains a rising edge
  - Group $G\downarrow 00$ – Set representing a dominant bit (0), does not contain a rising edge
  - Group $G\downarrow 01$ – Set representing a recessive bit (1), containing a falling edge
- Dominant bits, whose previous bits were also dominant, are discarded since these bits are unsuitable for classification

Phase 2: Preprocessing cont.

- Separate groups makes the system robust and accurate
  - Possible to use all bits after sampling for identification, independent of the transmitted data
  - Distinguishable characteristics of the different groups does not counterbalance each other
  - Makes the important characteristics more observable



(a) Group 00 containing 17 symbols

(b) Group 01 containing 18 symbols

(c) Group 10 containing 17 symbols

Phase 3: Feature Extraction

- System extracts and evaluates different statistical features for each of the previous prepared groups
- Time domain and magnitude of frequency domain are considered
- Relief-F algorithm from the Weka 3 Toolkit is used for selection of most significant features
- Best features of the test setups are combined to get a general feature set
- Most important characteristics are found in $G\downarrow 10$, which contain the rising edges
- Feature vector F(V) represents the fingerprint extracted from the received CAN signal

Feature	Description
Mean	$\mu = \frac{1}{N}\sum_{i=1}^{N} x(i)$
Standard Deviation	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x(i)-\mu)^2}$
Variance	$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x(i)-\mu)^2$
Skewness	$skew = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x(i)-\mu}{\sigma}\right)^3$
Kurtosis	$kurt = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x(i)-\mu}{\sigma}\right)^4$
Root Mean Square	$rms = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x(i)^2}$
Maximum	$max = max(x(i))_{i=1\ldots N}$
Energy	$en = \frac{1}{N}\sum_{i=1}^{N} x(i)^2$

Features considered in the selection, where x are the measured values in the *time domain* respectively the magnitude values in the *frequency domain* and N is the number of elements

1 $max(G_{10})$	2 $en(G_{10}^{FFT})$	3 $en(G_{00})$	4 $max(G_{00})$	5 $\mu(G_{10})$	6 $\mu(G_{00})$
7 $max(G_{10}^{FFT})$	8 $\mu(G_{10}^{FFT})$	9 $skew(G_{10})$	10 $kurt(G_{10}^{FFT})$	11 $kurt(G_{10})$	12 $\sigma^2(G_{10}^{FFT})$
13 $skew(G_{10}^{FFT})$	14 $skew(G_{01})$	15 $kurt(G_{01})$	16 $skew(G_{01}^{FFT})$	17 $kurt(G_{01}^{FFT})$	18 $\sigma^2(G_{10})$

Selected features for classification ordered by their rank

Phase 4 & 5: Classification & Detection

- Finding the sender ECU of a received frame is a classification problem
- Several machine learning techniques are used to identify the class of the new observation
  - Logistic Regression is used for training and prediction
- *Training Phase:*
  - Generate Fingerprints of multiple CAN frames for each of the different ECUs
  - Train the Supervised Learning model
- *Detection Phase:*
  - Compare the features of the newly received frames with the features collected for model generation
  - Predict the sender ECU

Deployment & Lifecycle

- Vehicle is considered to be in a safe environment during initial deployment phase
  - A key is assigned to each ECU to enable secure communication with the IDS
  - A safe training phase is carried out to avoid forged frames
- Performance monitor evaluates the quality of the classifiers
  - Model constantly adapts to changes ensuring high accuracy
  - Stochastic algorithms and online machine learning methods are used to update the existing model
- Influence of potential malicious data during the training phase is avoided by countermeasures of poisoning attacks
- Requires less bandwidth, can be implemented in ECUs with less resources and no additional hardware accelerators

Security of Scission

- *Detecting Compromised ECUs*
  - System calculates the probability of the ECU being allowed to send frames with the specified identifier
  - If the estimated probability is below the threshold $t\downarrow min$, the frame is marked as suspicious
  - The frame marked as suspicious is classified as malicious if the probability of the suspect device exceeds the threshold $t\downarrow max$ and trigger an alarm
  - If the probability does not exceed $t\downarrow max$, the frame is considered trustworthy to reduce false positives
- *Detecting Unmonitored and Additional ECUs*
  - Fingerprint of the unmonitored/additional ECU matches that of another ECU which is not allowed to use the received identifier → *Attack is detected*
  - Unmonitored/additional ECU has very similar characteristics to a trustworthy ECU which the attacker imitates → *Attack cannot be detected*
  - No ECU could be assigned → Frame is marked as suspicious

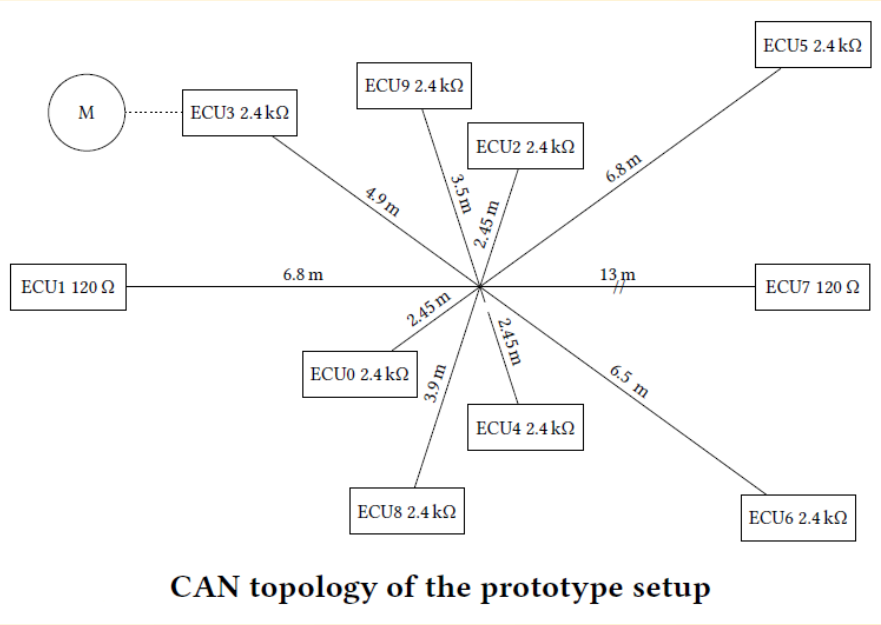
Security of Scission cont.

- *Detecting Scission-aware Attacker*
  - To impersonate a specific ECU, an attacker may influence its own voltage level by heating or cooling up the compromised ECU
  - Scission is able to continuously adapt to the slightly changing conditions
  - Scission uses several signal characteristics, it is unlikely for an attacker to impersonate a specific ECU
  - Attacker is not able to precisely adapt its signal due to the absence of general information about the characteristics
  - Cannot evade Scission

Evaluation

Evaluation Setup & Goal

- Prototype setup has 9 ECUs interconnected with each other
- Two real life cars used – Fiat 500 & Porsche Panamera S E-Hybrid
- Digital storage oscilloscope *PicoScope 5204* with a sampling rate of *500 MS/s* and a resolution of *8 bits* is used to record signals
- Two measurement series were created per frame, one for CAN low and one for CAN high, which were then combined to obtain the differential signal
- Evaluation Goal
  - Fingerprinting approach is able to identify the senders of received CAN frames with a high probability
  - Evaluate the ability of Scission to identify compromised, unmonitored and additional ECUs based on fingerprints



CAN topology of the prototype setup

Performance Evaluation

	ECU 0	ECU 1	ECU 2	ECU 3	ECU 4	ECU 5	ECU 6	ECU 7	ECU 8	ECU 9
ECU 0	**100**	0	0	0	0	0	0	0	0	0.42
ECU 1	0	**100**	0	0.29	0	0	0	0	0	0
ECU 2	0	0	**100**	0	0	0	0	0	0	0
ECU 3	0	0	0	**99.71**	0	0	0	0	0	0
ECU 4	0	0	0	0	**100**	0.18	0	0	0	0
ECU 5	0	0	0	0	0	**99.82**	0	0	0	0
ECU 6	0	0	0	0	0	0	**100**	0	0	0
ECU 7	0	0	0	0	0	0	0	**100**	0	0
ECU 8	0	0	0	0	0	0	0	0	**100**	0
ECU 9	0	0	0	0	0	0	0	0	0	**99.58**

Prototype Setup

	ECU 0	ECU 1	ECU 2	ECU 3	ECU 4	ECU 5	ECU 6	ECU 7
ECU 0	**100**	0	0	0	0	0	0	0.42
ECU 1	0.00	**100**	0	0.29	0	0	0	0
ECU 2	0.00	0	**100**	0	0	0	0	0
ECU 3	0.00	0	0	**99.71**	0	0	0	0
ECU 4	0.00	0	0	0	**100**	0.18	0	0
ECU 5	0.00	0	0	0	0	**99.82**	0	0
ECU 6	0.00	0	0	0	0	0	**100**	0
ECU 7	0.00	0	0	0	0	0	0	**99.58**

Porsche Panamera S E-Hybrid

	ECU 0	ECU 1	ECU 2	ECU 3	ECU 4	ECU 5	ECU 6	ECU 7
ECU 0	**99.90**	0	0.10	0	0	0	0	0
ECU 1	0	**99.89**	0	0.04	0	0.97	0	1.44
ECU 2	0.10	0	**99.72**	0	0	0.03	0	0
ECU 3	0	0	0	**99.96**	0	0	0	0
ECU 4	0	0	0	0	**100**	0.21	0	0
ECU 5	0	0	0.18	0	0	**98.75**	0	0
ECU 6	0	0	0	0	0	0	**100**	0
ECU 7	0	0.11	0	0	0	0.03	0	**98.56**

Fiat 500

Confusion matrix for the identification of ECUs

Performance Evaluation cont.

		Predicted		Suspicious Frames
		No attack	Attack	
Prototype	No attack	100	0	0
	Attack	1.5	98.5	0.2
Fiat	No attack	100	0	0.01
	Attack	0	100	0
Porsche	No attack	100	0	0.01
	Attack	3.18	96.82	3.18

Confusion Matrix of Scission

Samplerate (MS/s)	1	2	5	10	15	20
Identification rate	88.23	99.57	99.71	99.72	99.85	100
False positive rate	2.94	0.35	0.26	0.17	0.14	0

Performance for different sampling rates.

Discussion & Conclusion

Limitations

- If an attacker works with the identifiers that the ECU is allowed to use under normal conditions, Scission cannot detect them
- In case of additional ECUs, if the bus is modified without influencing the characteristics, the system will not longer be able to reliably recognize the change

Conclusion

- Usage of Scisson IDS in in-vehicle networks is a promising technology for improving their security
- Scission extracts fingerprints from the CAN signals for attacker identification with zero false positives
- Able to identify the correct sender with a probability of 99.85 %
- No impact on the available bandwidth – can be implemented in smart sensors
- Fingerprinting technology can enhance classical IDS approaches
- Can be used as a basis for stand-alone system or improve the security of gateways connecting different buses

THANK YOU