Towards Automatically Reverse Engineering Vehicle Diagnostic Protocols

Le Yu[1] , Yangyang Liu[1] , Pengfei Jing[1] , Xiapu Luo[1*] , Lei Xue[1]

Kaifa Zhao[1] , Yajin Zhou[2] , Ting Wang[3] , Guofei Gu[4] , Sen Nie[5] , Shi Wu[5]

[1]The Hong Kong Polytechnic University *[2]Zhejiang University*

[3]The Pennsylvania State University *[4]Texas A&M University*

[5]Tencent Keen Security Lab

31st USENIX Security Symposium

AUGUST 10–12, 2022

BOSTON, MA, USA

Motivation

- In-vehicle protocols: Very important to the security assessment and protection of modern vehicles.
  - Used in the **communication** between ECUs (e.g., CAN protocol).
  - Used in **accessing** and **manipulating** ECUs (e.g., diagnostic protocols).
- Vehicle diagnostic protocols
  - Reading sensor values or controlling ECUs through the **OBD port**.
  - *Example*: Keyword Protocol 2000 (**KWP 2000**) and Unified Diagnostic Services (**UDS**).
    - The standards define the low-level message formats.
    - The ***vehicle manufacturers*** define the *syntactic information, semantic meaning of messages*, and *formulas for encoding the return values*. ➔ **Proprietary *without*** publicly available documents.
  - Widely exploited to launch various attacks on vehicles.
    - Example: Miller et al. send **diagnostic messages** through OBD port to *kill engine* or *control fuel gauge* of the Ford and Toyota [1].

[1] C. Miller and C. Valasek. *Adventures in automotive networks and control units*. Def Con, 2013.

State-of-the-art

- Traffic analysis based methods [2, 3] analyze the formats of CAN messages transmitted between ECUs:
  - Do ***not*** consider the **transmission layer protocol**.
  - Do ***not*** recover the **proprietary formats** and **formulas**.
- App analysis based method [4] only identifies the possible request messages.
  - Do ***not*** recover the **proprietary formats** and **formulas**.

[2] D. Frassinelli, S. Park, and S. Nürnberger. *<<i know where you parked last summer>> automated reverse engineering and privacy analysis of modern cars*. In Proc. S&P, 2020.
[3] M. Pesé, T. Stacer, C. Campos, E. Newberry, D. Chen, and K. Shin. Librecan: *Automated can message translator*. In Proc. CCS, 2019.
[4] H. Wen, Q. Zhao, Q. Chen, and Z. Lin. *Automated crossplatform reverse engineering of can bus commands from mobile apps*. In Proc. NDSS, 2020.

Threat Model

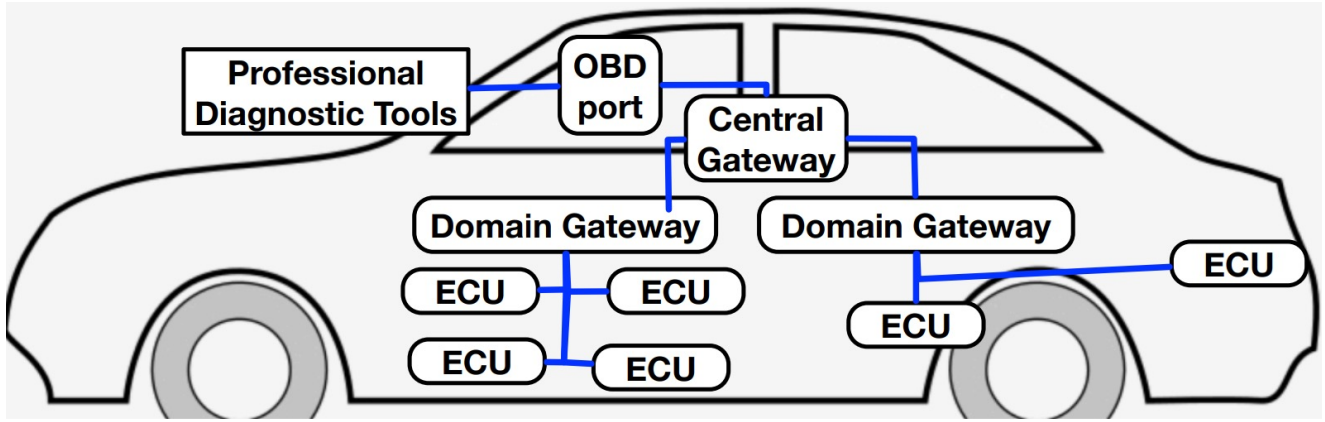

Fig 1. Communication system in the vehicle.

29 03 22 DB ... E5	Read brake pressure
12 03 22 DE ... 9C	Read accelerator position
43 05 31 01 ... 03	Control high beam (FLEL)
43 05 31 01 ... 01	Control low beam (FLEL)
60 05 31 01 ... 13	Control turn light (KOMBI)
01 02 ... 01	Reset collision safety module
60 02 ... 01	Reset combination instrument
Diagnostic Message(Lexus)	Functions
03 22 ... 7B	Read engine speed (Engine)
03 22 ... 59	Read throttle position (Engine)
04 30 01 ... 10	Control displayed speed (KOMBI)
04 30 02 ... 08	Control engine speed (KOMBI)
Diagnostic Message(Toyota)	Functions
40 05 30 11 00 ... 00	Unlock all doors
40 05 30 1C 00 ... 00	Turn on the wiper
40 05 30 11 00 ... 00	Unlock the trunk
Diagnostic Message(Kia)	Functions
04 2F B0 ... 03	Unlock central lock
04 2F B0 ... 03	Turn on all light on dashboard

Fig 2. Using reverse engineered diagnostic messages to attack BMW i3, Lexus NX300, Toyota Corolla, and Kia.

The ECUs of the vehicle are connected through bus systems.

- They communicate with each other by using CAN frames.
- They are connected to the gateway.

The OBD port was designed for On-board diagnostics (OBD).

The professional diagnostic tools connected to the OBD port can

- send diagnostic messages to ECUs and receive response messages
- read their values or even manipulate them.

Assumption of reverse engineering: A target vehicle + A diagnostic tool that works for the vehicle.

Tab 1. OSI model of the most popular diagnostic protocols (KWP 2000, UDS, and OBD-II)

Application Session	KWP 2000: ISO 14230-3 [5] or 15765-3 [6]	UDS: ISO 14229-2 [28]	OBD-II: ISO-15031 [34]
Transport	ISO 15765-2 [11] VW TP 2.0 [29]	ISO 15765-2 [11]	ISO 15765-2 [11]
Network			
Data Link Physical	K-Line: ISO 14230-1(2) [3,4], CAN: ISO 11898 [18]		

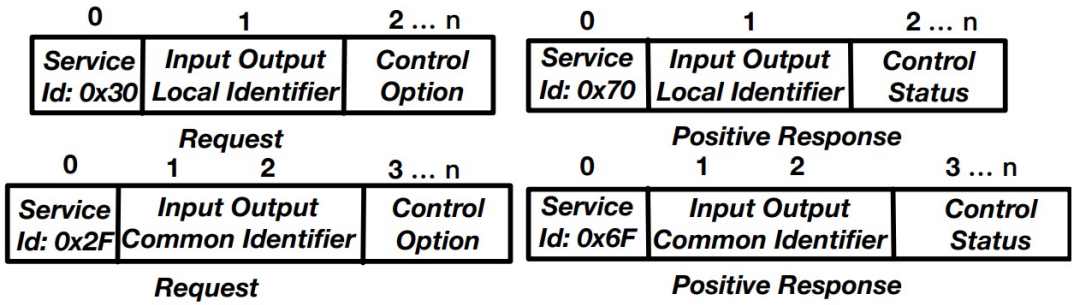

Fig 2. Request and positive response messages of the input output control by local identifier service and input output control by common identifier service of KWP 2000

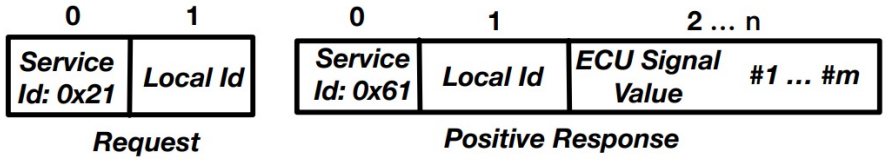

Fig 3. Request and positive response messages of the read data by local identifier service of KWP 2000

Example.

(1) Turn on/off the light

The diagnostic tool sends the request messages "**30 15 00 40 00**" and "**30 15 00 00 00**" to the Main Body Control ECU.

(2) Obtain the engine RPM

The diagnostic tool sends the request message "**21 07**" to the engine.

It receives a response message containing the ESV "**01 F1 10**" .

The formula type is **0x01**.

The corresponding formula is **X0 *X1/5**.

The value of X0 is **0xF1** (i.e., 241) and the value of X1 is **0x10** (i.e., 16) ➔ The actual ESV is **771.2 /min (i.e., 242*16/5).**

System Design: Overview

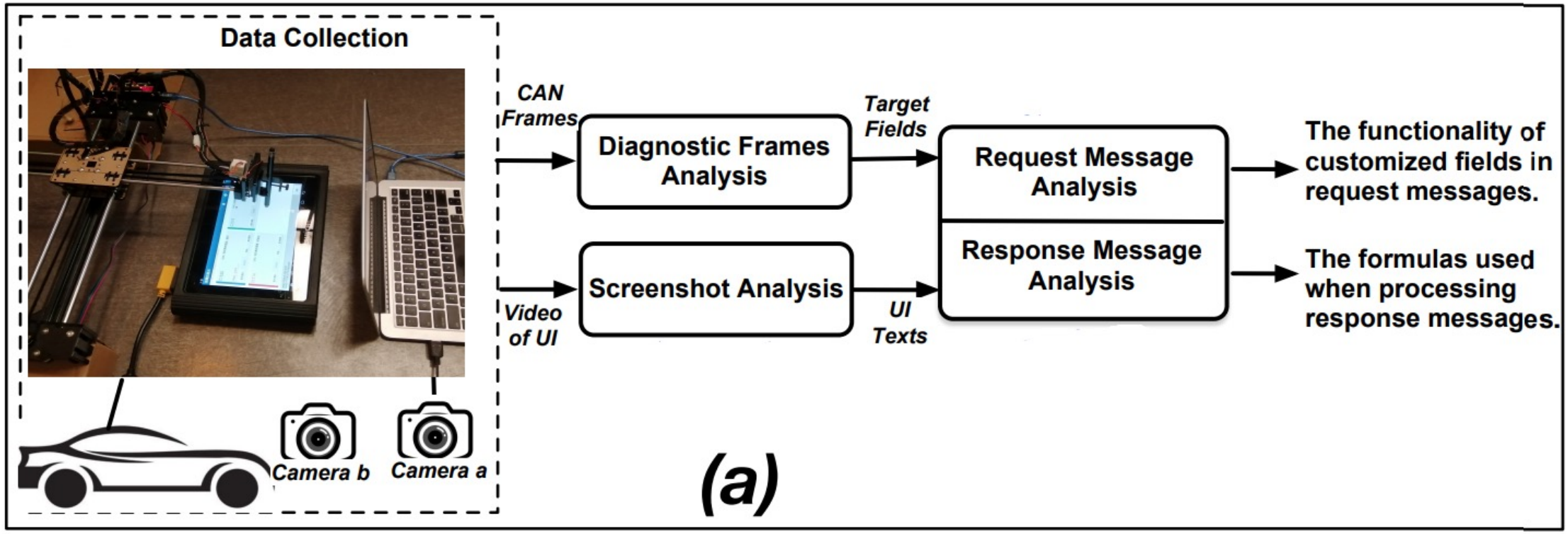

Fig 4. System overview

System Design: Data Collection

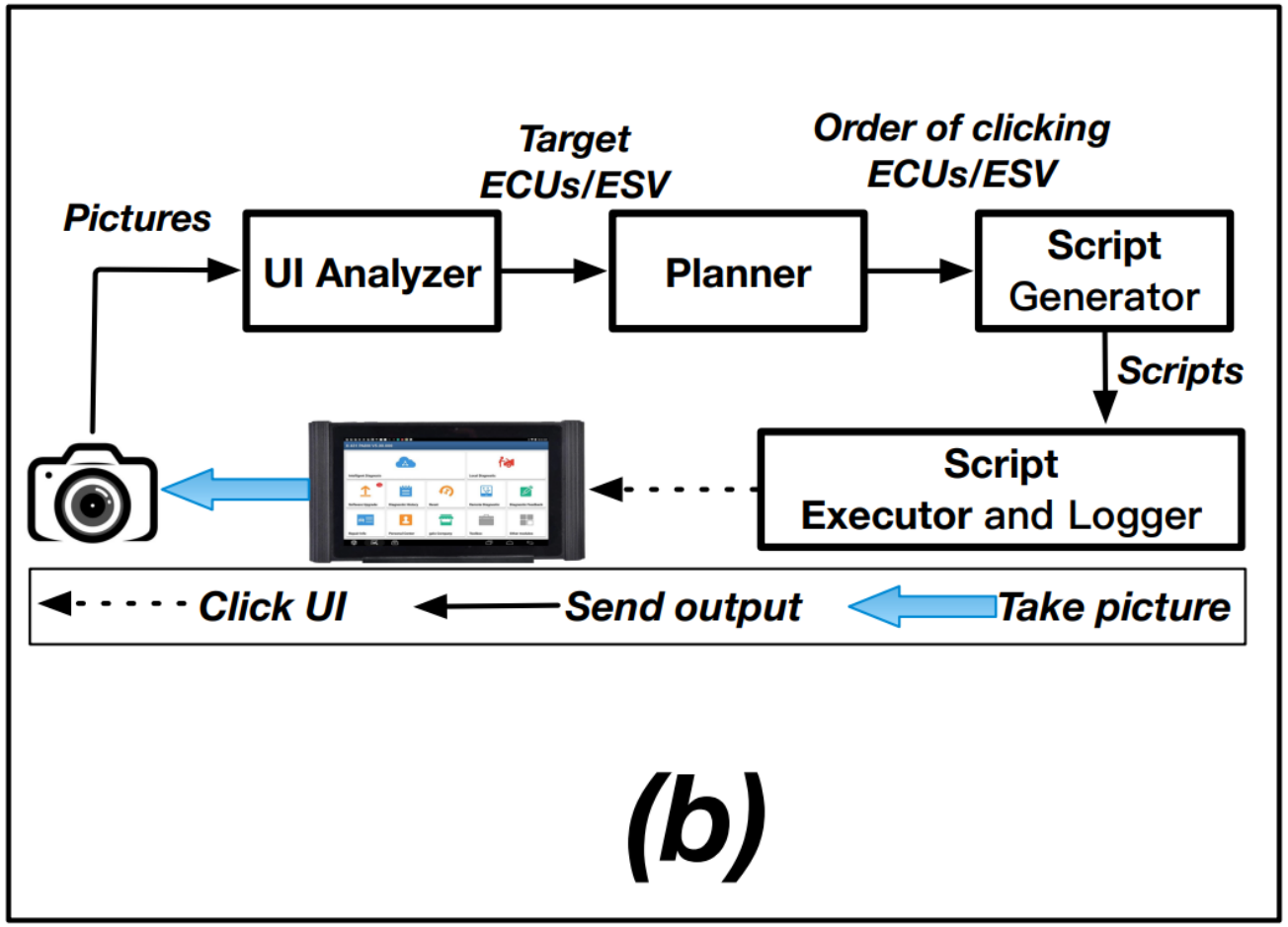

Fig 5. Steps of data collection

Diagnostic Frames Analysis

- **Step 1**: Screening Frames.
- **Step 2**: Assembling Payload.
- **Step 3**: Fields Extraction.
  - Extract the **local id**, **DID**, **ESV** and **ECR** contained in diagnostic messages.
  - For each **ECR**, we also extract the **IO control parameter** and **control state**.

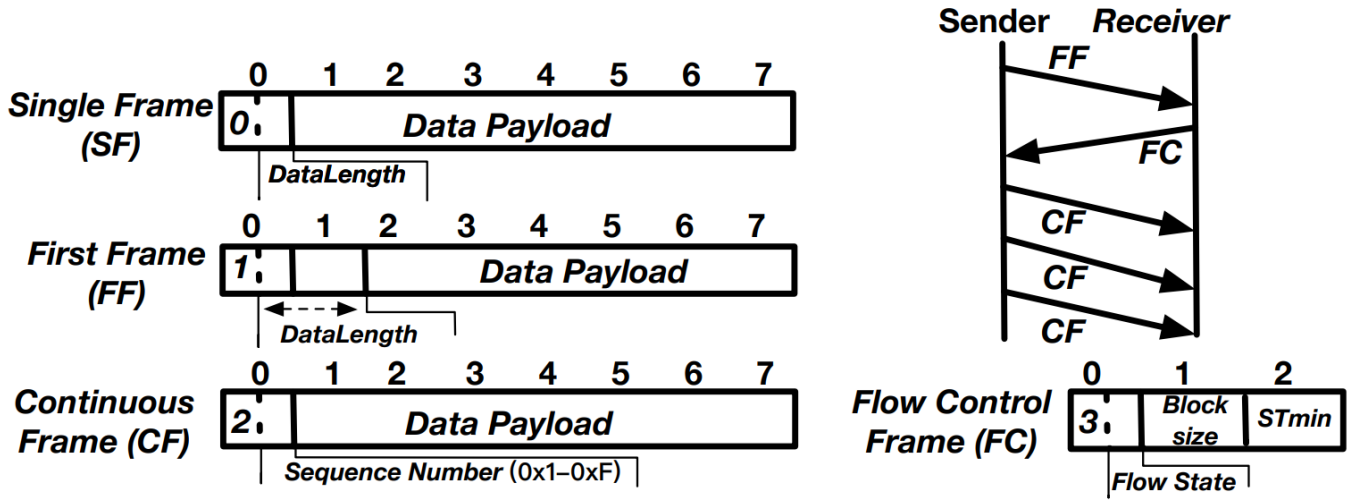

Fig 6: ISO 15765-2: Structures of single frame (SF), first frame (FF), continuous frame (CF), flow control frame (FC), and flow control mechanism.

Screenshot Analysis

UI Text Extraction.

- Use *MPlayer* to transform the video into a series of images
- Apply *Tesseract* to these images for extracting the text

Incorrect ESV Value Filtering.

- Since sometimes the OCR engine might miss some decimal points (e.g., “25.00” is incorrectly identified as “2500”), we remove incorrect ESVs by using pre-defined range of each ESV.

Request Message Analysis

- When reading the ESV through UDS or KWP 2000, the request message includes the **DID** or **local identifier**, which are customized by the manufacturers.

- When controlling vehicle components with UDS, the request message also includes **DID** of the target component.

Response Message Analysis

An improved genetic programming algorithm to infer the formulas.

- **Step 1**: Constructing the mapping between ESV in diagnostic messages and ESV displayed on UI.
- **Step 2**: Inferring the formula through genetic programming.
- **Step 3**: Pre-processing of the data set and post-processing of the formula.

	ESV in response messages (X)		**ESV** displayed on UI (Y)	
Range	Pre-process	Formula Post-processing	Pre-process	Formula Post-processing
$> 10^4$	$X' = X/10^4$	Replace(X', $X/10^4$)	$Y' = Y/10^4$	Replace(Y', $Y/10^4$)
10^3-10^4	$X' = X/10^3$	Replace(X', $X/10^3$)	$Y' = Y/10^3$	Replace(Y', $Y/10^3$)
10^2-10^3	$X' = X/100$	Replace(X', $X/100$)	$Y' = Y/100$	Replace(Y', $Y/100$)
10-10^2	$X' = X/10$	Replace(X', $X/10$)	$Y' = Y/10$	Replace(Y', $Y/10$)
0.1-1.0	-	-	$Y' = Y * 10$	Replace(Y', $Y * 10$)
10^{-2}-10^{-1}	-	-	$Y' = Y * 100$	Replace(Y', $Y * 100$)
10^{-3}-10^{-2}	-	-	$Y' = Y * 10^3$	Replace(Y', $Y * 10^3$)
$< 10^{-3}$	-	-	$Y' = Y * 10^4$	Replace(Y', $Y * 10^4$)

Tab 3. Vehicles and diagnostic tools used in experiments

Evaluation

Car	Vehicle Model	Protocol	Diagnostic Tools
Car A	Skoda Octavia	UDS	LAUNCH X431
Car B	Volkswagen Magotan	KWP 2000	VCDS
Car C	Volkswagen Lavida	KWP 2000	LAUNCH X431
Car D	Lexus NX300	UDS	Techstream
Car E	Mini Copper R56	UDS	AUTEL 919
Car F	Mini Copper R59	UDS	AUTEL 919
Car G	BMW i3	UDS	AUTEL 919
Car H	RongWei MARVEL X	UDS	AUTEL 919
Car I	Changan Eado	UDS	AUTEL 919
Car J	BMW 532Li	UDS	AUTEL 919
Car K	Volkswagen Passat	KWP 2000	AUTEL 919
Car L	Toyota Corolla	UDS	AUTEL 919
Car M	Peugeot 308	UDS	AUTEL 919
Car N	Kia k2 (UC)	UDS	AUTEL 919
Car O	Ford Kuga	UDS	AUTEL 919
Car P	Honda Accord	UDS	AUTEL 919
Car Q	Nissan Teana	UDS	AUTEL 919
Car R	Audi A4L	UDS	AUTEL 919

Tab 3. Vehicles and diagnostic tools used in experiments

Evaluation

Precision of the OCR engine

Diagnostic Tool	#Total Pics	#Correct Pics	Precision
AUTEL 919	500	488	97.6%
LAUNCH X431	500	425	85.0%

Tab 4. Performance of OCR engine

Result of OBD-II Frames

ESV	Request Message	Formula Ground Truth	Formula (GP) System Output
Absolute Throttle Position	01 11	$Y = \frac{X}{2.55}$	$\frac{Y}{10} = \frac{(X/100)}{0.255}$
Calculated Engine Load	01 04	$Y = \frac{X}{2.55}$	$\frac{Y}{10} = \frac{X/100}{0.255}$
Fuel Tank Level Input	01 2F	$Y = 0.392 * X$	$\frac{Y}{100} = 0.389 * \frac{X}{100}$
Engine Speed (RPM)	01 0C	$Y = \frac{256 * X_0 + X_1}{4}$	$Y = 64X_0 + 32$
Vehicle Speed (Km/h or Mile/h)	01 0D	$Y = X$ or $Y = 0.621 * X$	$\frac{Y}{100} = 0.619 * \frac{X}{100}$
Engine Coolant Temperature(^oC or ^oF)	01 05	$Y = X - 40$ or $Y = 1.8 * X - 40$	$Y = 1.7 * X - 22$
Intake Manifold Absolute Pressure(KPa or inHg)	01 0B	$Y = X$ or $Y = X/3.39$	$\frac{Y}{10} = \frac{X}{100}/0.335$

Tab 5. Result of reverse engineering the formulas of OBD-II protocol: the request messages, the formulas in ground truth, and the formulas inferred by DP-Reverser.

Evaluation

Result of UDS and KWP 2000 Frames

Car	#ESV (formula)	#Correct ESV	Precision	#ESV (Enum)
Car A	28	28	100.0%	0
Car B	8	7	87.5%	0
Car C	5	5	100.0%	0
Car D	12	12	100.0%	5
Car E	5	5	100.0%	4
Car F	8	8	100.0%	5
Car G	5	4	80.0%	22
Car H	5	5	100.0%	13
Car I	11	9	81.8%	0
Car J	20	20	100.0%	20
Car K	41	41	100.0%	0
Car L	29	28	96.6 %	20
Car M	4	4	100.0%	14
Car N	26	26	100.0%	19
Car O	18	18	100.0%	9
Car P	7	7	100.0%	6
Car Q	18	18	100.0%	17
Car R	40	40	100.0%	2
Total	290	285	98.3%	156

Tab 6. Result of ESV analysis: Number of ESVs with formulas (i.e., column "#ESV (formula)"), number of ESVs that GP can infer formulas correctly (i.e., column "#Correct ESV"), precision of inferring formulas with GP (i.e., column "Precision"), and the number of ESVs without formulas (i.e., column "#ESV (Enum)").

Evaluation

Comparion with Alternative Algorithms for Formula Inferring

Car	#ESV (formula)	#Correct ESV (Linear Reg)	# Correct ESV (Polynomial)
Car A	28	14	20
Car B	8	2	1
Car C	5	1	2
Car D	12	10	8
Car E	5	3	2
Car F	8	4	3
Car G	5	2	2
Car H	5	5	3
Car I	11	9	6
Car J	20	11	8
Car K	41	2	0
Car L	29	25	12
Car M	4	4	2
Car N	26	14	11
Car O	18	11	6
Car P	7	3	3
Car Q	18	7	4
Car R	40	34	28
Total	290	127	93

Tab 7. Precision of inferring formulas of UDS and KWP 2000 with linear regression (i.e., column “#Correct ESV (Linear Reg)”) and polynomial curve fitting (i.e., column “#Correct ESV (Polynomial)”).

Result of Reverse Engineering ECR

Car	#ECR	Service ID	Car	#ECR	Service ID
Car A	11	2F	**Car D**	5	30
Car E	3	30	**Car F**	5	30
Car H	6	2F	**Car I**	10	2F
Car J	27	30	**Car N**	21	2F
Car O	4	2F	**Car Q**	32	30

Tab 8. Number of ECRs extracted from vehicles.

- ***First Request***: The controller sends the "Freeze current state" message. The format is "**2F {DID: 2 bytes} 02**".
- ***Second Request***: The controller sends the "Short term adjustment" message. The format is "**2F {DID: 2 bytes} 03 {control state: n bytes}**".
- ***Third Request***: The controller sends the "Return control to ECU" message. The format is "**2F {DID: 2 bytes} 00**".

Formulas Extracted from Apps

APP Name	Formula Type	# Formula
Carly for VAG	UDS KWP 2000	90 137
Carly for Mercedes	UDS KWP 2000	1624 468
Carly for Toyota	KWP 2000	7
inCarDoc	OBD-II	82
Car Computer - Olivia Drive	OBD-II	74
CarSys Scan	OBD-II	64
Easy OBD	OBD-II	55
inCarDoc Pro	OBD-II	49
OBD Boy(OBD2-ELM327)	OBD-II	45
FordSys Scan Free	OBD-II	42
ChevroSys Scan Free	OBD-II	40
ToyoSys Scan Free	OBD-II	40
Obd Mary	OBD-II	34
OBD2 Boost	OBD-II	34
Obd Harry Scan	OBD-II	28
Obd Arny	OBD-II	27
MOSX	OBD-II	24
Dr Prius Dr Hybrid	OBD-II	22
Dacar Pro OBD2	OBD-II	21
OBD2 Scanner Fault Codes Desc	OBD-II	16
Dacar Pro OBD2	OBD-II	14
Engie Easy Car Repair	OBD-II	8
PHEV Watchdog	OBD-II	8
Torque Lite(OBD2&Car)	OBD-II	5
Kiwi OBD	OBD-II	3
OBDclick	OBD-II	2
Dr Prius Dr Hybrid	OBD-II	1
Fuel Economy for Torque Pro	OBD-II	1

Tab 9. Telematics apps containing formulas.

Thank you!

Please send questions to yulele08@gmail.com

The Hong Kong Polytechnic University 香港理工大學

Zhejiang University 浙江大學

PennState

科恩实验室 KEEN security lab