Bluetooth Low Energy Security Testing with Combinatorial Methods

Dominik Schreiber, Dimitris E. Simos, Manuel Leithner, Jovan Zivanovic

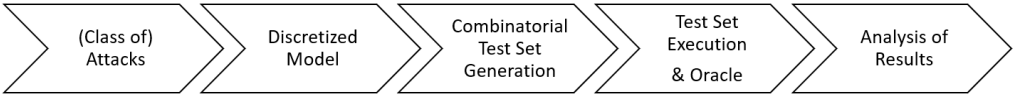
MATRIS Research Group, SBA Research, Austria

July 9^{th}, 2025
USENIX ATC
Boston, USA

Combinatorial Security Testing (CST)

D. Schreiber
D. E. Simos
M. Leithner
J. Zivanovic

- Most errors triggered by low parameter interactions
- Application of combinatorial methods to security testing
- Focus on specially crafted input
- Test for unspecified functionality
- Affects modeling and oracle
- Successfully applied to SQLi, XSS, TLS testing



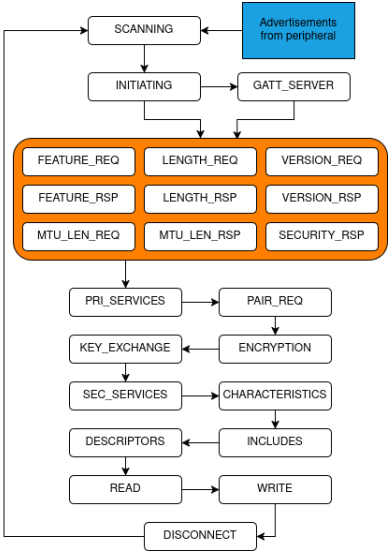
Covering Array (CA) Example

- Provides t-way interaction coverage guarantee over input model
- System with 10 binary parameters
- Exhaustive: 1024 tests
- Example of t=3 Covering Array: 13 tests
- Various tools available
  - CAgen (MATRIS)
  - Pict (Microsoft)
  - ACTS (NIST)

Test	Var → A	B	C	D	E	F	G	H	I	J
1	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	1
3	1	1	1	0	1	0	0	0	0	1
4	1	0	1	1	0	1	0	1	0	0
5	1	0	0	0	1	1	1	0	0	0
6	0	1	1	0	0	1	0	0	1	0
7	0	0	1	0	1	0	1	1	1	0
8	1	1	0	1	0	0	1	0	1	0
9	0	0	0	1	1	1	0	0	1	1
10	0	0	1	1	0	0	1	0	0	1
11	0	1	0	1	1	0	0	1	0	0
12	1	0	0	0	0	0	0	1	1	1
13	0	1	0	0	0	1	1	1	0	1

GreyHound Fuzzer

- Initially developed for WIFI testing
- Also supports Bluetooth and BLE
- Definition of packets in Scapy library
- Custom firmware for NRF52840
- State machine
- Access to all layers
- Fuzzing algorithm guided by heuristic optimization



Input Parameter Model (IPM)

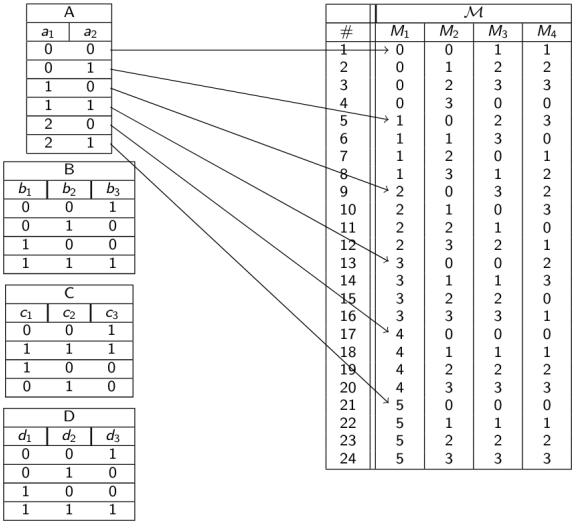
- One model per layer
- Devised by
  - Values from documentation
  - Field Types
- Marked invalid values

Packet Layers
BTLE
BTLE_ADV
BTLE_CONNECT_REQ

```
BTLE_CONNECT_REQ_IPM = {
    "InitA": ["UNCHANGED", "~__SLAVE_ADDR__"],
    "AdvA": ["UNCHANGED", "~__MASTER_ADDR__"],
    "AA": ["UNCHANGED", "~0x8e89bed6", "~0xffffffff",
    ↪  "~0x00000000"],
    "crc_init": ["UNCHANGED", "0", "0xffffff"],
    "win_size": ["UNCHANGED", "0", "2", "0xf", "~0xff",],
    "win_offset": ["UNCHANGED", "0", "1", "2", "~0xffff"],
    "interval": ["UNCHANGED", "0", "16", "32", "~0xffff"],
    "latency": ["UNCHANGED", "0", "1", "10", "~500", "~501",
    ↪  "~0xffff"],
    "timeout": ["UNCHANGED", "~0", "16", "32", "~0xffff"],
    "chM": ["UNCHANGED", "~0x0", "0x11FFFFFFFF",
    ↪  "~0xFFFFFFFFFF"],
    "hop": ["UNCHANGED", "~0", "~4", "7", "~17"]),
    "SCA": ["UNCHANGED", "~1"]
}
```

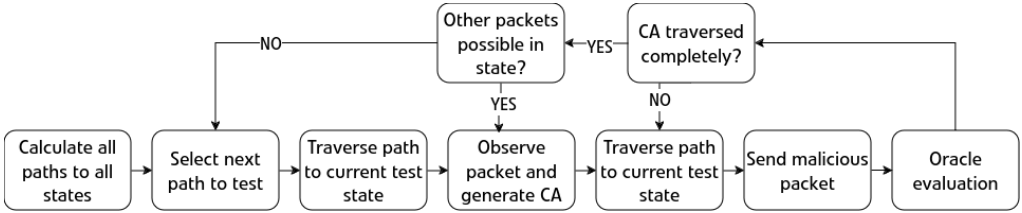
Covering Array (CA) Generation

- Interactions between layers
- Composite system model
- One CA for every layer
- Generated with Pict
- Layer CAs combined to meta CA
- Individual and meta CAs strength t=2



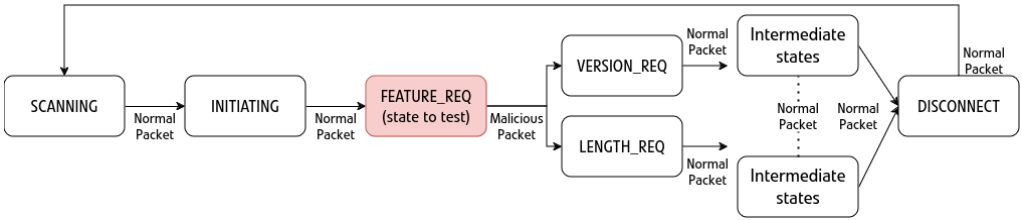
Testing Strategy

- Construct set of paths to states
- Chose one path to a state
- Test if end state is reachable
- Check for new packet
- Generate CA for packet
- Execute Tests
- If no new packets, check next path



Test Execution

- Send normal packets until state to to test
- Replace packet values with CA row
- Send malicious packet
- Send normal packets until disconnect

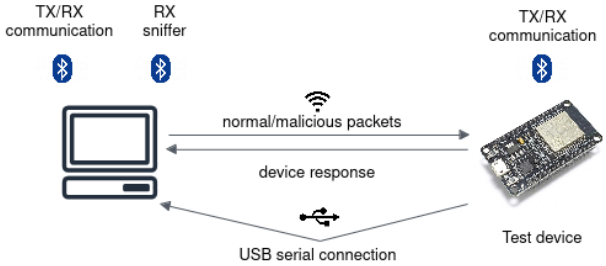


Oracles

- GreyHound fuzzer oracle very complex
  - Many false positives
- Crash oracle
- Advertisement oracle

Test Setup

- 2 x NRF52840 usb dongle
- Laptop
- Target micro controller
- MongoDB
- Replicated 4 times



Target Devices

SoC Vendor	SoC Model	SDK Versions	Sample App
Texas Instruments	CC2640R2	3.30.00.20 (old fw), 5.30.00.03 (new fw)	project zero
Espressif Systems	ESP32	4.1 (old fw), 5.0 (new fw)	nimble/blepprphrl
Nordic Semiconductors	nRF52	15.3.0 (old fw), 17.0.1 (new fw)	ble_app_gatts_c
Bouffalolab	bl602	AI-Thinker WB2 beta v1.1.8	ble_slave
WCH	CH582M	MounRiver Studio community v1.50	Peripheral
Hi-Link	W801	a93b517	wm_ble_client_api-_multi_conn_demo
Realtek	RTL8720DN	Realtek Ameba Boards 3.1.6	BLEBatteryService
Silicon Labs	BG22	simplicity studio SV5.7.1.1	bluetooth_controlling-_led_from_smartphone
Ambiq	Apollo3	sparkfun apollo3 boards v2.2.1	LED_Button
MacroGiga Electronics	MG126	seeed SAMD boards 1.8.4	analog_output

Table: Tested devices and software versions.

Results

Target	Issue	State	Parameter
CC2640R2 new fw	URWF	initiating	interval: 0xffff
	URWF	initiating	timeout: 0
	URWF	initiating	latency: 0xffff
CC2640R2 old fw	URWF	initiating	interval: 0xffff
	URWF	initiating	timeout: 0
	URWF	initiating	latency: 0xffff
	URWF	initiating	interval: 0
	URWF	initiating	latency: 501
ESP32 new fw	URWF	initiating	AA: 0x00000000
	URWF	initiating	AA: 0xffffffff
ESP32 old fw	core dump	initiating	chM: 0x0
	URWF	initiating	AA: 0x00000000
bl602	URN	initiating	chM: 0
	core dump	length_req	max_rx_bytes: 0
W801	URN	initiating	AA: 0x00000000
RTL8720DN	URN	initiating	latency: 0xffff
	TO	initiating	interval: 0xffff
	TO	initiating	win_offset: 0xffff
	TO	initiating	latency: 500
	TO	initiating	interval: 0
Apollo3	URN	pair_request	authentication: 255
MG126	TO	initiating	win_offset: 0xffff
	TO	initiating	latency: 501

Table: TO - Timeout, URWF - Unavailable but Recoverable With Code Fix, URN - Unavailable Reset Needed, core dump - Core dump detected

Result Comparison

	Issues		Anomalies		Tests	
Target	GH	CST	GH	CST	GH	CST
CC2640R2 (new fw)	0	**3**	0	-	2,129	**1,188**
CC2640R2 (old fw)	2	**5**	1	-	**1,000**	1,309
ESP32 (new fw)	0	**2**	1	-	5,979	**1,910**
ESP32 (old fw)	**2**	**2**	0	-	**1,000**	2,115
nRF52 (new fw)	0	0	1	-	8,790	**1,309**
nRF52 (old fw)	0	0	1	-	**1,000**	1,309
bl602	1	**2**	2	-	7,233	**2,985**
CH582M	0	0	2	-	4,237	**1,184**
W801	0	**1**	1	-	**879**	1,132
RTL8720DN	0	**5**	2	-	10,153	**2,664**
BG22	0	0	0	-	8,016	**1,282**
Apollo3	**1**	**1**	0	-	7,935	**594**
MG126	0	**2**	0	-	9,993	**373**

Table: GH = GreyHound fuzzer, CST = Combinatorial Security Testing method

Result Summary

- GreyHound executed more tests
- CST found more issues
- Both found most issues early
- Many DoS possibilities
- Core dump potential for remote code execution

Method	Total Tests	Total Unique Issues
GreyHound	68,344	17
CST	19,354	19

Disclosure

- Issues reported to vendors
- Espressif Systems
  - Acknowledged and fixed the reported issues
  - Bug bounty of 2,229$
- Realtek
  - Stopped responding
- Texas Instruments, Bouffalolab, Hi-Link, Ambiq and MacroGiga Electronics did not responded to our reports.

D. Schreiber
D. E. Simos
M. Leithner
J. Zivanovic

Conclusion

- CST can be applied to BLE testing
- All found issues triggered by single parameters
- Simple models can already find many issues
- Difficult to find compromise between coverage and execution time
- CST is more consistent compared to fuzzing

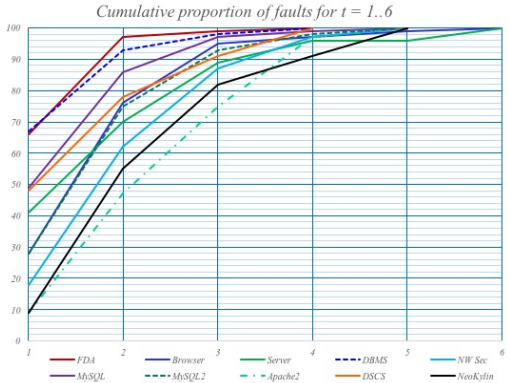
Thank you very much for your attention!

Questions?

matris@sba-research.org
matris.sba-research.org

Combinatorial Testing (CT)

- Based on interaction rule
- Most errors triggered by low parameter interactions
- Efficient and effective testing method
- Provides (quantitative) assurance
- Successfully applied to hardware, OS and application layer testing



Lessons Learned

- Strong masking effects
- Execution time of large CAs becomes
- Constrain layer interactions
- Model sequence interactions

Real Life Setup

D. Schreiber
D. E. Simos
M. Leithner
J. Zivanovic

Future Work

- Sequence testing
- Automatic model extraction
- Testing of other protocols
- Improving oracle reliability

Bluetooth Low Energy

- Separate protocol from Bluetooth classic
- Uses less energy
- Simpler hardware requirements
- Sensors and IoT devices
- Separated host / controller layers