부채널기반 디스어셈블러

Power-based Side-Channel Instruction-level Disassembler
논문 리뷰

https://youtu.be/b-IC7XBQUJo

한성대학교 HANSUNG UNIVERSITY

CryptoCraft LAB
https://crypto.modoo.at

Contents

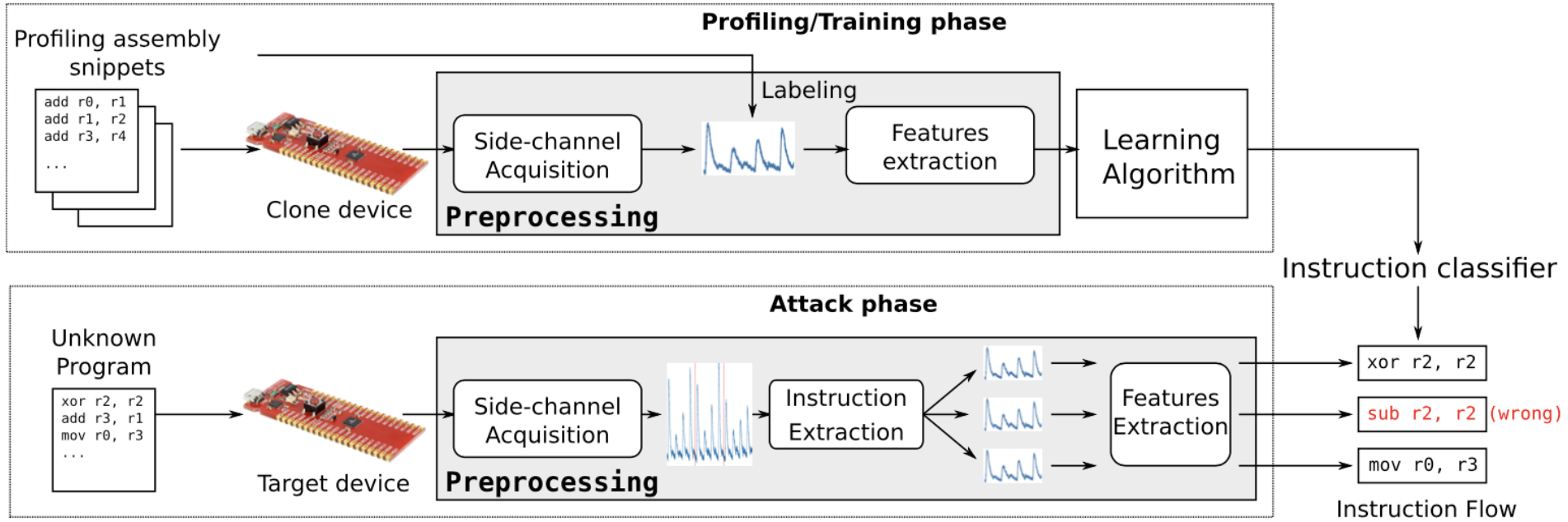
CryptoCraft LAB

SCBD (Side-Channel Based Disassembling)

- 리버스 엔지니어링은 소프트웨어 프로그램의 이진 코드를 가져와 원래 소스 코드를 다시 만드는 프로세스
- 소프트웨어의 리버스 엔지니어링은 불법 복제 또는 저작권 분석을 위한 일반적인 방법
- 통상적으로 Disassembler를 사용하여 기계어를 어셈블리 언어로 변환하는 방법을 사용합니다.
- SCBD란
  부채널 누출이라고하는 물리적 서명을 기반으로 장치에서 실행되는 명령을 복구하는 작업

SCBD (Side-Channel Based Disassembling)

- SCBD 작업은 감독 된 기계 학습 분류 문제
  정확하게 예측할 수 있는 분류기를 구축하기 위해 학습 알고리즘 필요



- 모델을 만들려면 많은 대상 별 지식이 필요하기 때문에 이러한 분류를 교육하는 것은 실제로 어려움
  opcode 분류 접근법은 소형 마이크로 컨트롤러에서 작동하는 것으로 입증되었지만 더 복잡한 프로세서로 확장 될 가능성은 낮음

Power-based Side-Channel Instruction-level Disassembler

- 명령 레벨 단위로 임베디드 시스템의 실시간 작동을 분석
  - 계층 적 분류 프레임 워크
  - 연속 웨이블릿 변환 (CWT)
  - KL (Kullback-Leibler) 발산을 사용
  - PCA 사용
  - 공변량 변화 적응 기술

- 디스어셈블러를 AVR 8 비트 마이크로 컨트롤러에서 구현
- 99.03 % 이상의 성공률을 가진 레지스터 이름을 포함한 테스트 명령을 인식

Grouping

- ATMega 328P의 131 개의 명령어 중 잔류 제어 명령, 곱셈 명령 및 잔류 분기 명령을 제외한 112 개의 명령 인식
- 피연산자를 기준으로 8 개의 그룹으로 나눔
- 큰 클래스를 분류 할 경우 계산 복잡성을 크게 줄임

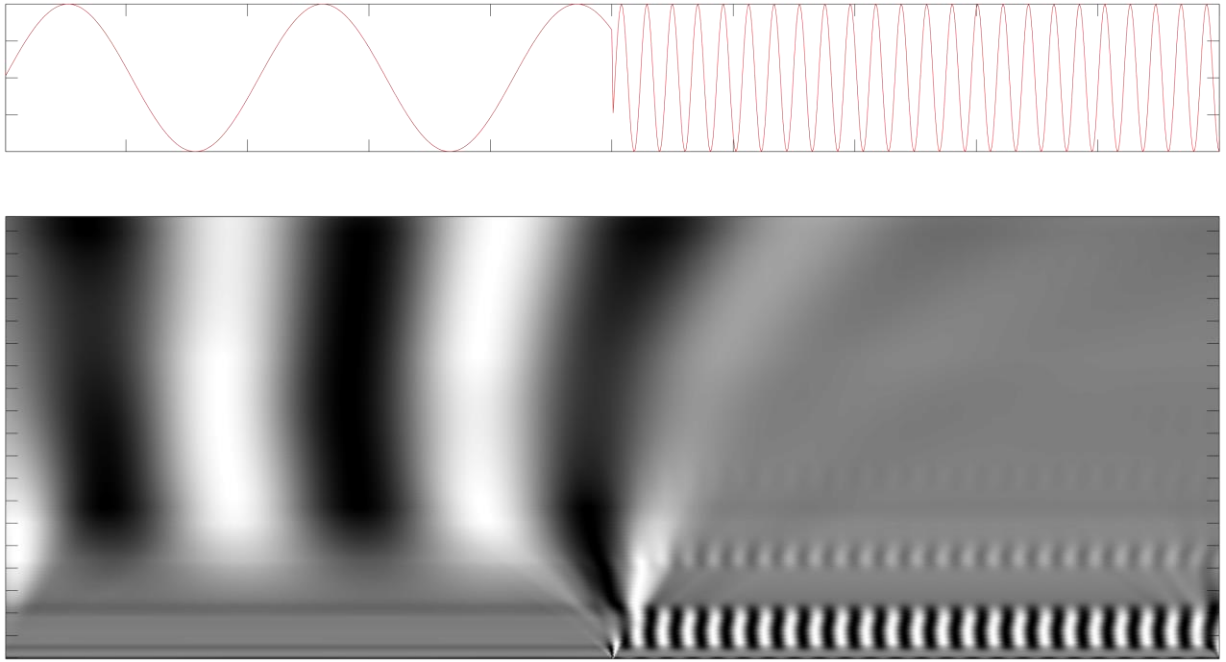
Table 2: Grouping AVR instructions.

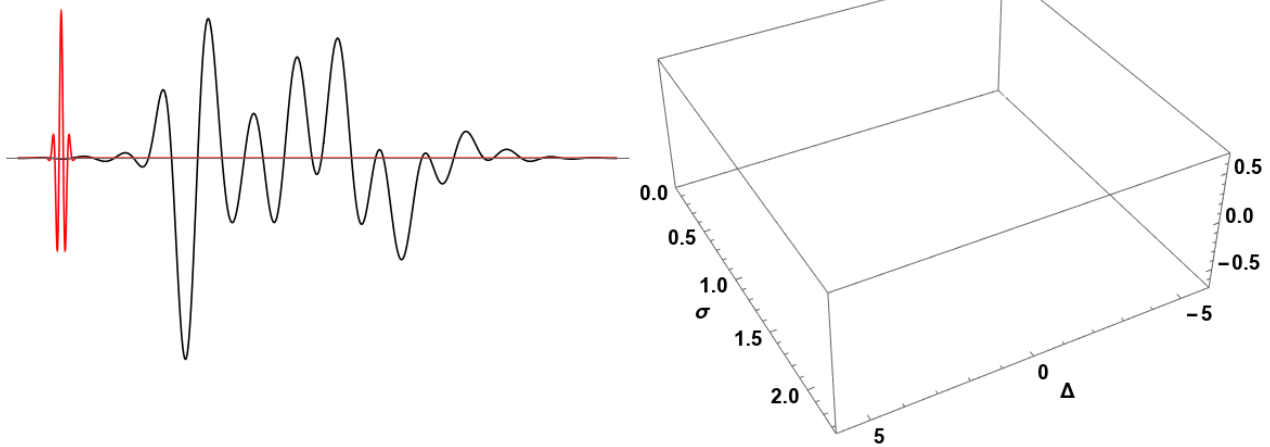
	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8
Insts.	ADD, ADC, SUB SBC, AND, OR EOR, CPSE, CP, CPC, MOV, MOVW	ADIW, SUBI, SBCI SBIW, ANDI, ORI SBR, CBR, CPI LDI	COM, NEG, INC DEC, TST, CLR SER, LSL, LSR ROL, ROR, ASR SWAP	RJMP, JMP, BREQ BRNE, BRCS, BRCC BRSH, BRLO, BRMI BRPL, BRGE, BRLT BRHS, BRHC, BRTS BRTC, BRVS, BRVC BRIE, BRID	LDS LD LDD STS ST STD	SEC, CLC, SEN CLN, SEZ, CLZ SEI, SES, CLS SEV, CLV, SET CLT, SHE, CLH	SBRC, SBRS SBIC, SBIS BRBS, BRBC SBI, CBI BST, BLD BSET, BCLR	LPM ELPM
Operands	Rd, Rr	Rd, K	Rd	k	Rd, k Rd, (−)X(+) Rd, (−)Y(+(q)) Rd, (−)Z(+(q))		Rr(Rd), b A, b s, k s	Rd, Z(+)
# of Insts.	12	10	13	20	24	15	12	6
Description	Arith.[1]	Arith.,[1] Data.[2]	Bit.,[3] Arith.[1]	Bran.[4]	Data.[2]	Bit.[3]	Bran.,[4] Bit.[3]	Data.[2]

[1] Arithmetic and logic instruction, [2] Data instruction, [3] Bit and bit-test instruction, [4] Branch instruction

CWT (Continuous wavelet transform)

- 연속 웨이블릿 변환 (CWT)에 의해 2차원 시간 주파수 영역에 매핑
- 신호의 주파수와 해당 주파수와 관련된 시간을 제공하여 여러 분야에서 사용하기에 매우 편리합니다.
- 예를 들어, 보행 분석을 위한 가속 신호 처리, 결함 탐지, 저전력 심박 조율기 설계 및 초 광대역 (UWB) 무선 통신에 사용





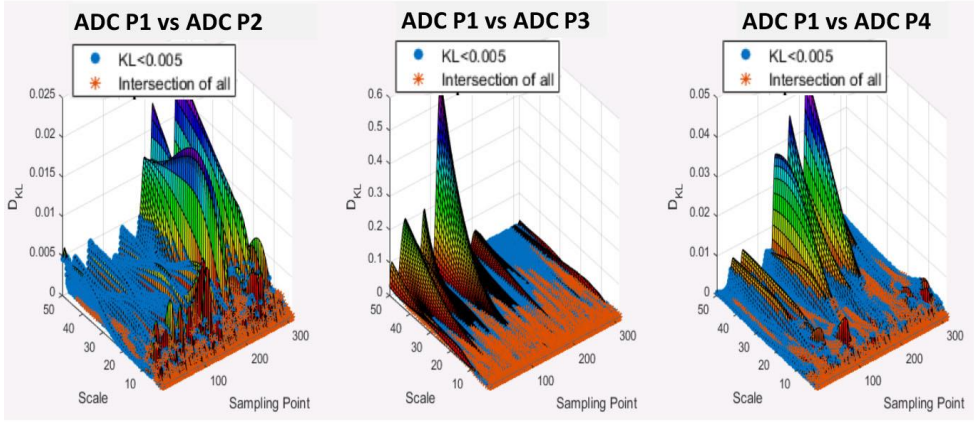
CWT (Continuous wavelet transform)

- 사이드 채널 누설 흔적에서 노이즈를 제거하고 수집 된 흔적을 완벽하게 정렬하는 데 사용됩니다.
- 시간-주파수 영역에서 뚜렷하고 변하지 않는 특징
- 공변량 시프트 문제를 해결하는 데 필요

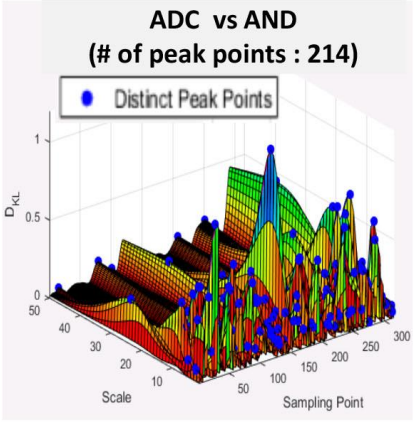
KL (Kullback-Leibler) Divergence

- 구별되고 변하지 않는 포인트를 추출하기 사용
- 두 확률분포의 차이를 계산하는 데에 사용하는 함수
- 어떠한 확률분포 P가 있을 때, 샘플링 과정에서 그 분포를 근사적으로 표현하는 확률분포 Q를 P대신 사용할 경우 엔트로피 변화를 의미
- 특정 샘플링 포인트는 KL- 분산 값이 커야 함

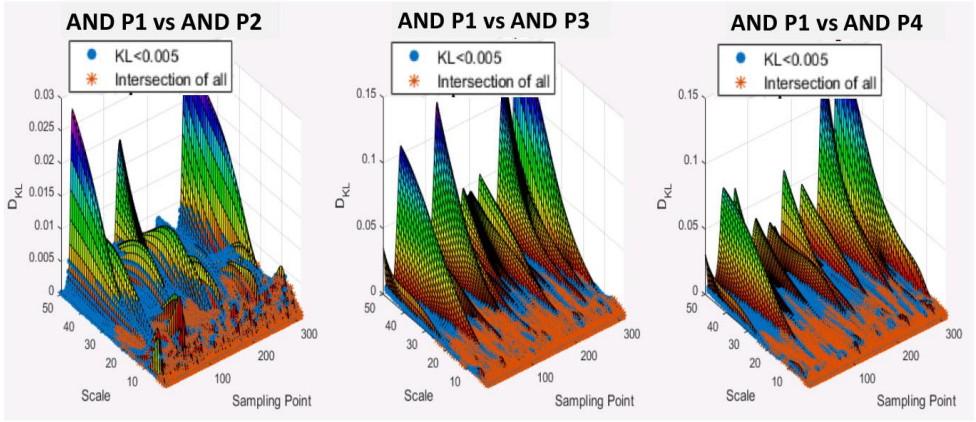
KL (Kullback-Leibler) Divergence



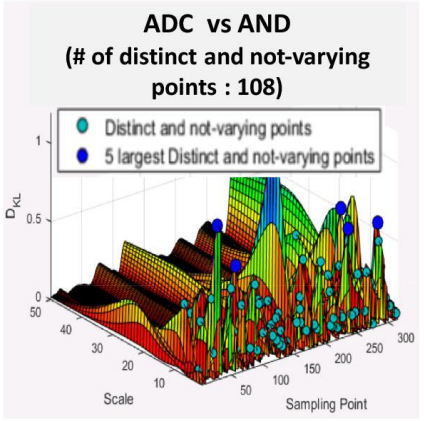
(a) Not-varying points of ADC based on D_{KL}^{W}



(b) Distinct points between ADC and AND based on D_{KL}^{B}



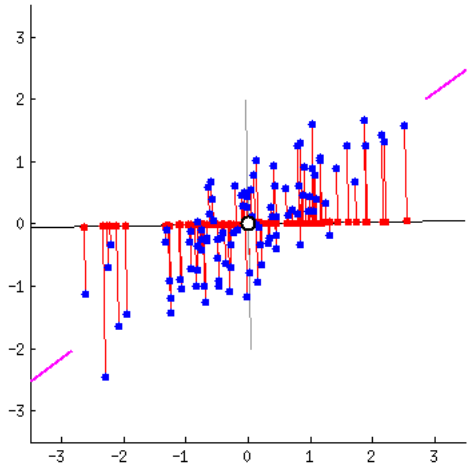
(c) Not-varying points of AND based on D_{KL}^{W}

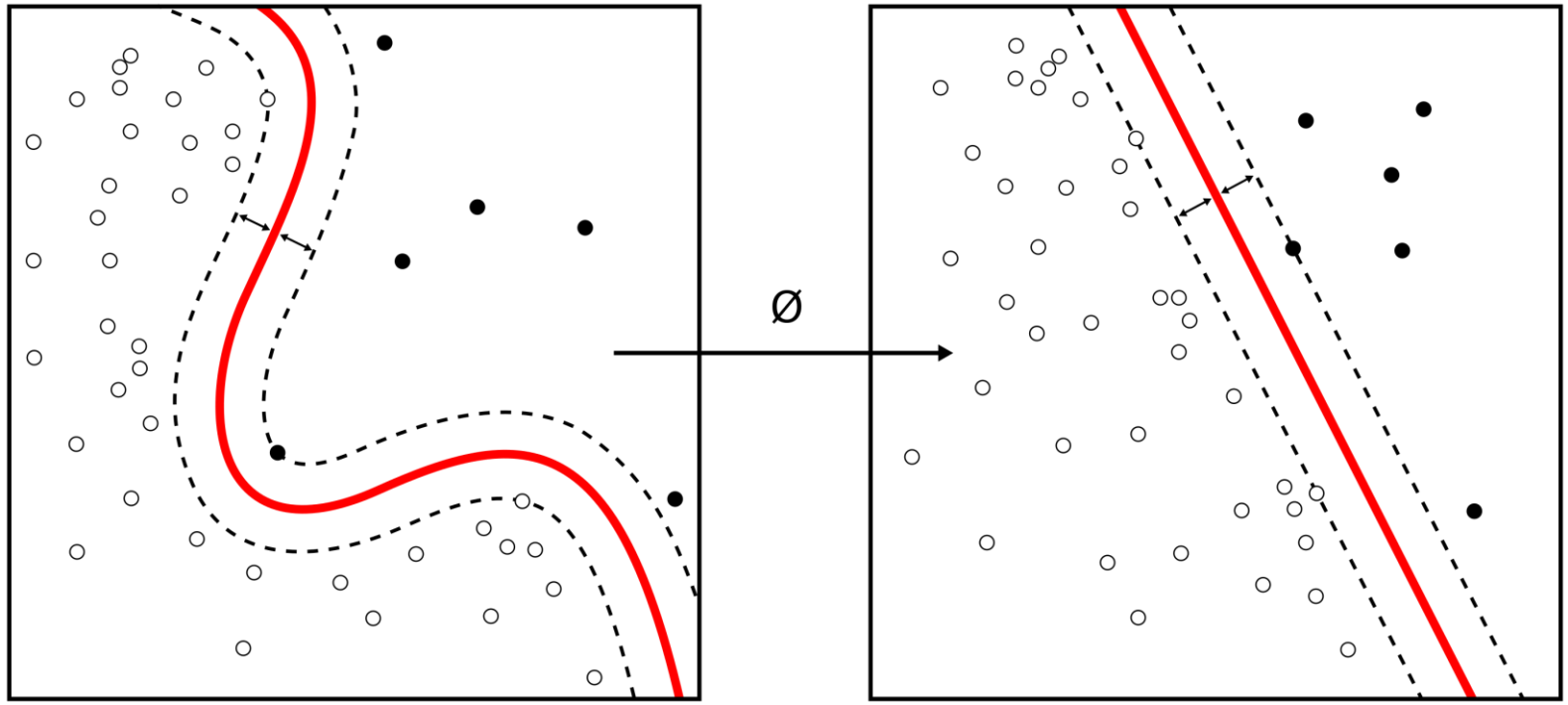


(d) Distinct and not-varying points between ADC and AND

PCA (Principal Component Analysis)

- KL 분기로 선택된 특징점에 PCA 적용
- 서로 연관 가능성이 있는 고차원 공간의 표본들을 선형 연관성이 없는 저차원 공간(주성분)의 표본으로 변환하기 위해 직교 변환을 사용



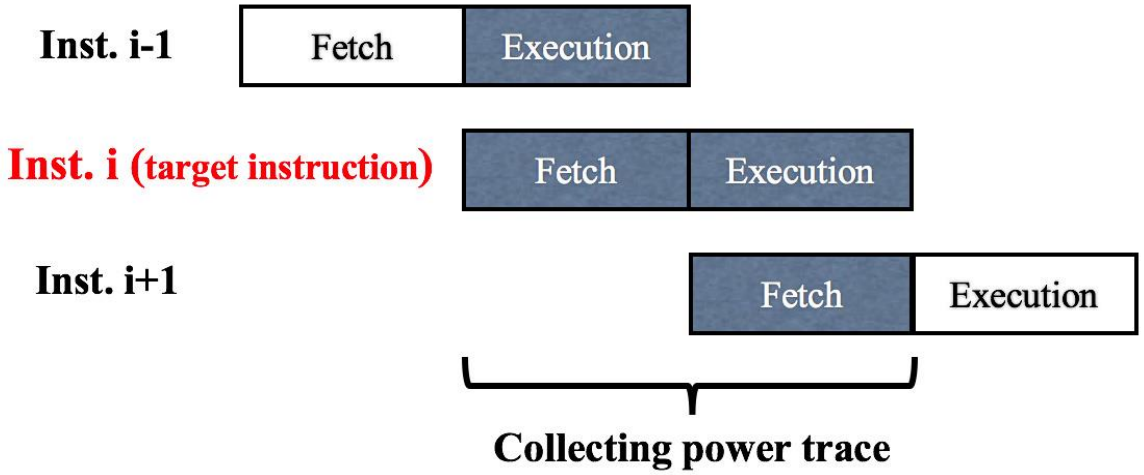


실험환경

- 훈련 또는 프로파일 링을 위해 16MHz의 클럭 주파수를 가진 ATMega 328P 마이크로 컨트롤러에서 전력 트레이스를 수집
- 다른 5 개의 ATMega 328P 마이크로 컨트롤러가 대상 장치로 사용
- Tektronix MDO3102 오실로스코프를 사용하여
  GND 핀과 접지 사이의 저항 (330Ω)의 전압을 측정
- 오실로스코프의 설정은 2.5GS / s, 250MHz 대역폭, 10k 샘플 포인트

실험 설정

- AVR 마이크로 컨트롤러에는 2 개의 파이프 라인 단계가 있으므로
  대상 프로파일 링 명령은 이전 명령과 다음 명령의 영향을받음
- 3000 개의 전력 트레이스가 샘플링됩니다.
  또한 무작위로 선택된 명령어를 사용하여 고유 한 Rd 당 3000 개의 전력 트레이스와
  Rr을 무작위로 선택한 명령어를 사용하여 고유 한 Rr 당 3000 개의 전력 트레이스를 측정합니다



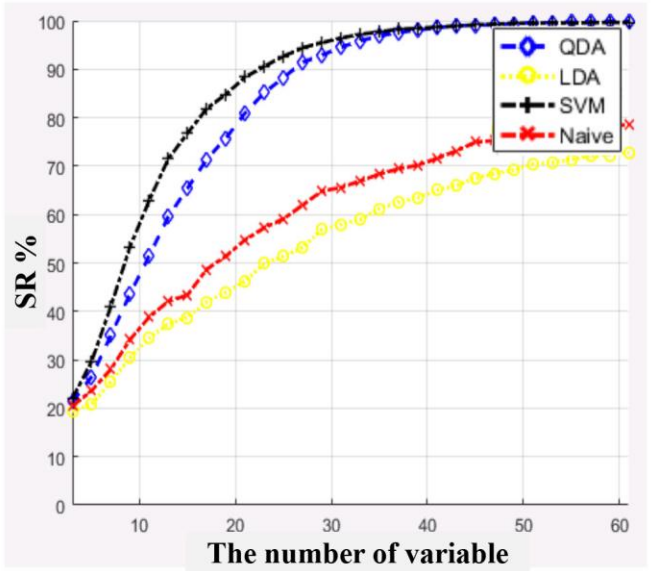

```
"sbi   %0, %1 \n\t"
"nop   \n\t"
"cbr   r27, 100 \n\t"
"add   r21, r18 \n\t"
"dec   r27 \n\t"
"nop   \n\t"
"cbi   %0, %1 \n\t"
```

Randomly selected

Fixed

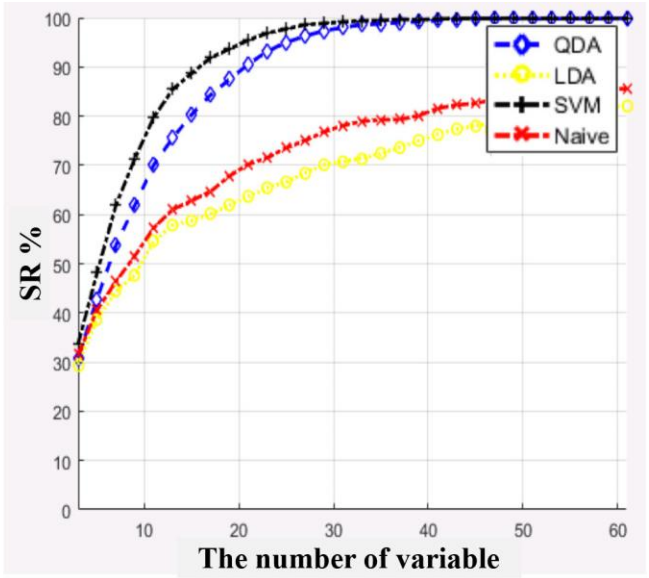
교육 훈련 및 분류

- 차원 축소 후에는 각 클래스 당 2500 개의 전력 트레이스가 훈련에 사용됩니다. 선형 판별 분석 (LDA), 2 차 판별 분석 (QDA), 지원 벡터 머신 (SVM) 및 나이브 베이즈 방법이 테스트



# of var.	QDA	LDA	SVM	Naive
43	99.6	77.35	99.85	83.35

(a) Classification of groups (Group1 ~ Group8)



# of var.	QDA	LDA	SVM	Naive
61	99.93	74.53	99.7	80.22

(b) Classification of the first group's instructions

교육 훈련 및 분류

Table 3: Successful recognition rate (SR) of classification between ADC and AND with covariate shift adaptation (CSA).

Classifier	Without CSA	Without Norm.	With Norm.
QDA	18.5%	54.3%	92%
SVM	19.2%	57.8%	93.2%

Table 4: Successful recognition rate (SR) of classification between ADC and AND in 5 different devices.

Classifier	Dev. 1	Dev. 2	Dev. 3	Dev. 4	Dev. 5
QDA	89.3%	91.5%	88.9%	92.3%	94.5%
SVM	90.4%	92.8%	90.8%	93.4%	95.6%

결론

- 부채널 정보를 사용하여 99.03 % 이상의 성공률을 가진 레지스터 이름을 포함한 테스트 명령을 인식
- 실제 세계 시나리오에서 완전한 리버스 엔지니어링을 모방하기 위해 실제 코드에 대한 방법을 적용 예정
- 또한 고주파 클럭에서 실행되는 최첨단 마이크로 컨트롤러도 테스트 된 디스어셈블러의 성능을 평가 예정

Q & A