데이터 분석 중심의 자동차용 침입 탐지 (NETSEC-KR 2020)

김휘강 (cenda@korea.ac.kr)

고려대학교 정보보호대학원

해킹대응기술연구실

KOREA UNIVERSITY

HCRL Hacking and Countermeasure Research Lab

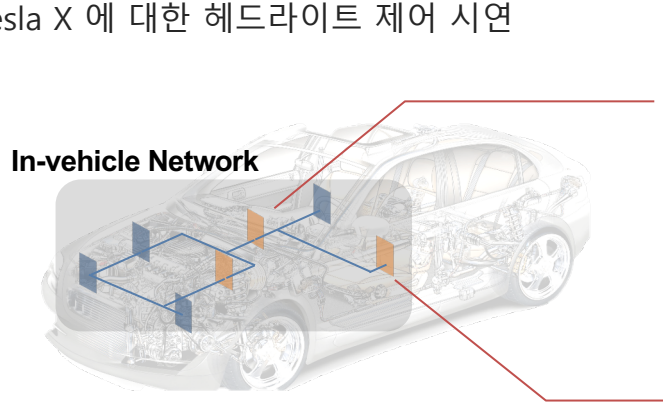
CONTENTS

I. 2019 정보보호 R&D 챌린지 revisited
II. 융합보안 챌린지 결과
III. Lesson learned

자동차 침입 탐지 트랙

□ 차량 내부 네트워크 침입을 통해, 차량을 원격에서 차량 제어

- 차량 원격 제어 – Charlie Miller and Chris Valasek ('15)
  - Jeep Cherokee 차량에 대한 CAN Bus로 악의적인 메시지 주입으로 차량 제어 시연
- 차량 원격 제어 및 취약점 탐지 – Keen Security Lab ('16, '17)
  - BMW 및 Tesla S 차량에 대한 공격을 통해 차량으로부터 12mile (약19km) 떨어진 거리에서 차량 제어 시연
  - ECU 제어를 통해 Tesla X 에 대한 헤드라이트 제어 시연



©2020, Hacking and Countermeasure Research Lab

융합보안 챌린지 – 자동차 침입탐지 트랙 문제 설명

자동차 침입 탐지 트랙

□ **차량 내부 네트워크 보안을 위한, Data-driven 자동차 침입탐지 트랙 개최**

□ **2017년의 차량 이상징후 탐지 트랙 -> 2018 차량주행 데이터 기반 도난탐지 트랙 -> 2019 자동차용 침입탐지 트랙**

- 차량 보안 분야 트랙의 연속성 유지
- 차량용 IDS 개발 분야에 응용 가능
- 기 게재되어 검증된 논문과 dataset 활용, academic follow-up research 활성화

□ **출제 의도: CAN bus dataset 기반의 침입 탐지 알고리즘 및 프로그램 제시**

- Data-Driven 으로 차량 네트워크에서 차량 공격에 대한 침입 탐지 기술 보급/확산
- 현재 차량보안 동향 (대다수의 차량에 침입탐지시스템이 탑재되는 동향) 반영
- UNECE WP 29, ITU-T SG17 X.itssec-4, X.ipscv 등 관련 국제 표준에 대응

©2020, Hacking and Countermeasure Research Lab

대회 취지/ 문제 설명

자동차 침입 탐지 트랙

□ **KU-Survival Analysis Dataset – key reference**

- Mee Lan Han, Byung Il Kwak, and Huy Kang Kim. "Anomaly intrusion detection method for vehicular networks based on survival analysis." Vehicular Communications 14 (2018): 52-63.

□ **예선: KIA Soul, HYUNDAI Sonata, CHEVORLET Spark 차량 3개에 대한 정상 (Attack free), 공격 (Flooding, Fuzzy, Malfunction) 데이터셋**

□ **본선: KIA Soul, HYUNDAI Sonata 차량 2개에 대한 정상 (Attack free), 공격 (Flooding, Fuzzy, Malfunction) 데이터셋 (주행시 데이터 수집)**

©2020, Hacking and Countermeasure Research Lab

대회 취지/ 문제 설명

자동차 침입 탐지 트랙

□ 차량 내부의 CAN Bus Message들을 수집

- 제공될 CAN Message는 Timestamp, CAN ID, DLC, Payload 로 구성됨
- 학습용 데이터셋에는 라벨링이 이뤄져 있음

[데이터 추출 방법]

```
Timestamp: 1481192898.078149    ID: 04b1 DLC: 8    00 00 00 00 00 00 00 00
Timestamp: 1481192898.078396    ID: 0164 DLC: 8    00 08 00 00 00 00 05 0d
Timestamp: 1481192898.078637    ID: 0370 DLC: 8    ff 20 00 80 ff 00 00 ec
Timestamp: 1481192898.078873    ID: 043f DLC: 8    10 50 64 ff 51 54 09 00
Timestamp: 1481192898.079115    ID: 0440 DLC: 8    ff f0 00 00 ff 54 09 00
Timestamp: 1481192898.079357    ID: 04f2 DLC: 8    00 00 90 38 00 00 00 a1
Timestamp: 1481192898.079597    ID: 0110 DLC: 8    e0 3c 30 09 00 00 00 00
```

CAN Message 예시

- 차량 3종에서 정상 (Attack Free status) 및 공격 (Under Attack status) 데이터셋을 수집
- YF Sonata, KIA Soul, Chevrolet SPARK 차량에 대한 정상 상태 및 공격 상태에 대한 CAN 데이터셋 수집
- 공격에 대한 차량 돌발상황 대처를 위해, 정차상태에서 데이터셋 수집

YF Sonata

KIA Soul

Chevrolet SPARK

©2020, Hacking and Countermeasure Research Lab

자동차 침입 탐지 트랙

□ 데이터셋 구성

- 차량 공격시도 패킷이 포함된 차량 네트워크 데이터셋 (Labeling 패킷 단위로 적용)
  - 예선: (Flooding, Fuzzing, Malfunction) , 차종 A, B, C
  - 본선: 1차(Fuzzing + Malfunction, 차종 (A + B), 주행 시 CAN data 추출)
  - 본선: 2차(Malfunction + Replay, 차종 (A + B), 주행 시 CAN data 추출)

구분	사용용도	Dataset	공격 유형	차종수	Labeling	난이도
예선	Train dataset	정상 데이터셋 + 공격 데이터셋 3개	Flooding, Fuzzing, Malfunction	3개 차량	O	중
	Release dataset	정상 데이터셋 + 공격 데이터셋 3개	Flooding, Fuzzing, Malfunction	3개 차량	X	중
본선	1st session - Train dataset	정상 데이터셋 + 공격 데이터셋 2개	Fuzzing, Malfunction	2개 차량	O	중
	1st session - Release dataset	공격 데이터셋 2개	Fuzzing, Malfunction	2개 차량	X	중
	2nd session - Train dataset	공격 데이터셋 2개	Malfunction, Replay	2개 차량	O	상
	2nd session - Release dataset	공격 데이터셋 2개	Malfunction, Replay	2개 차량	X	상

©2020, Hacking and Countermeasure Research Lab

자동차 침입 탐지 트랙

□ 평가방법 (F1-score)

- Positive, Negative가 편향될 수 있어 Precision과 Recall의 조화평균으로 평가하는 방법

카테고리		실제결과	
		Positive	Negative
실험결과	Positive	*True Positive(TP)*	*False Positive(FP)*
	Negative	*False Negative(FN)*	*True Negative(TN)*

- True Positive는 실제 True를 True로 정확하게 예측한 상황을 의미함
- True Negative는 실제 False를 False로 정확하게 예측한 상황을 의미함
- False Positive는 실제 False을 True로 예측한 상황을 의미함 (오탐)
- False Negative는 실제 True를 False로 예측한 상황을 의미함 (미탐)

$$Precision = \frac{TP}{TP + FP} \qquad Recall = \frac{TP}{TP + FN} \qquad \boldsymbol{F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}}$$

©2020, Hacking and Countermeasure Research Lab

융합보안 챌린지 결과

©2020, Hacking and Countermeasure Research Lab

©2020, Hacking and Countermeasure Research Lab

융합보안 챌린지 결과

참가팀들의 역량이 매년 뚜렷이 향상되는 추세

- Data-Driven Security 의 활성화 추세
- 데이터셋에 대한 이해도 (data exploration analysis) 증가
- Deep learning 의 보편화
- Deep learning 외에도 본인들 고유의 아이디어를 많이 가미함
- 다만, 알고리즘 적용에 대한 정당화 (justification) 노력은 여전히 부족

자동차 연구용 테스트 베드가 있을 경우, 그간 차량보안 관련 연구 진입장벽을 손쉽게 뛰어넘을 수 있을 것으로 판단됨

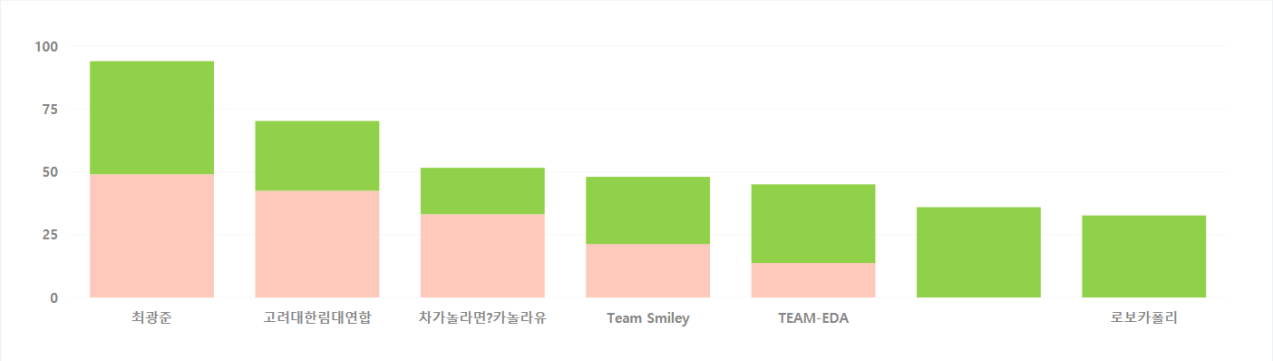
©2020, Hacking and Countermeasure Research Lab

융합보안 챌린지 결과

2019 자동차 탐지 트랙 – 순위 및 정확도

자동차용 침입 탐지트랙

최광준	고려대한림대연합	차가놀라면?카놀라유	Team Smiley	TEAM-EDA	자동차 없는 자동차 침입 탐지	로보카폴리
93.994%	70.2%	51.574%	47.975%	44.99%	35.903%	32.612%



자동차용 침입 탐지트랙

팀명	1차 탐지율	2차 탐지율
최광준	98.008%	89.98%
고려대한림대연합	84.875%	55.525%
차가놀라면?카놀라	66.266%	36.881%
Team Smiley	42.504%	53.445%
TEAM-EDA	27.377%	62.603%
자동차 없는 자동차	0%	71.806%
로보카폴리	0%	65.224%

©2020, Hacking and Countermeasure Research Lab

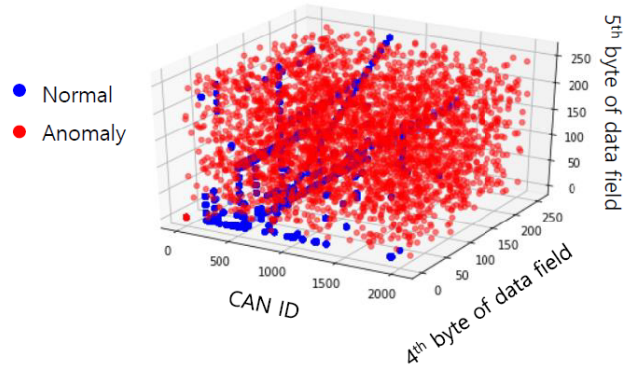
융합보안 챌린지 결과

우승팀 – 발표자료 (발췌)

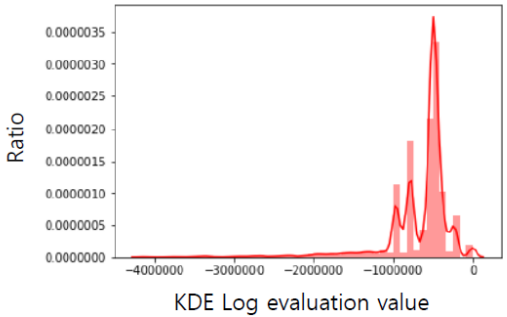
□ EDA and Feature engineering

- 정상 상태의 CAN message 우선순위 및 데이터 내용이 비정상 상태와 다르게 나타남
- Gaussian Distribution, Kernel Density Estimator (KDE)를 통한 CAN message 의 data field 분포 추정 결과를 Feature vector로 구성

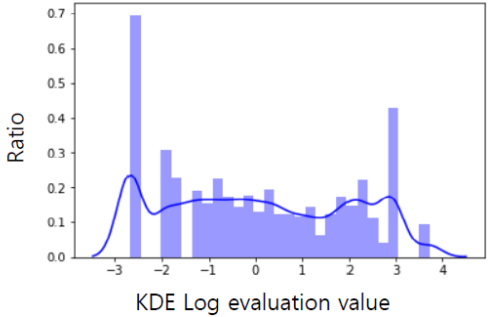
Spark 차량의 정상/침입 상태에서 CAN ID와 Data 분포



Spark 차량의 침입 상태에서 KDE 확률분포



Spark 차량의 정상 상태에서 KDE 확률분포



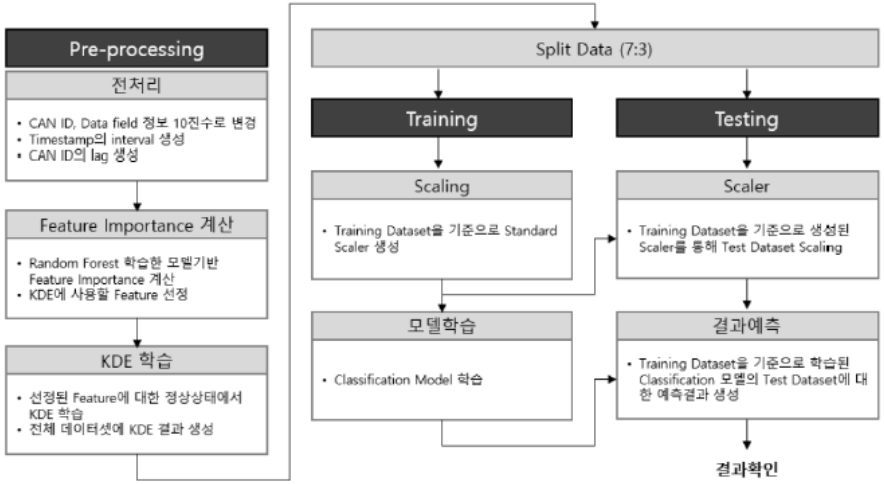
©2020, Hacking and Countermeasure Research Lab

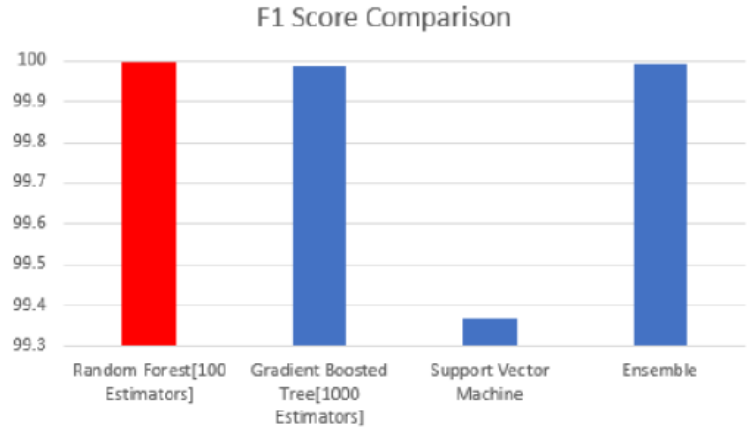
융합보안 챌린지 결과

우승팀 – 발표자료 (발췌)

□ Random Forest and Gradient Boosted Tree

- 정상 상태의 CAN ID Sequence, Time interval, and Frequency 를 기반으로 데이터 분포와 데이터 발생 확률을 추가 Feature로 구성하여 학습 데이터로 사용
- Random Forest 와 Gradient Boosted Tree 알고리즘을 공격 식별





Classifier에 따른 성능 비교

©2020, Hacking and Countermeasure Research Lab

융합보안 챌린지 결과

자동차용 침입탐지 트랙 - 2019

□ **대학생 연합, 대학원생 및 실무 데이터 분석 팀들과 같이 다양한 그룹에서 대회 참여**

- 해당 트랙에 다양한 조합의 팀들이 지원함 – 자동차용 침입탐지시스템 개발의 진입장벽을 낮춤

©2020, Hacking and Countermeasure Research Lab

융합보안 챌린지 결과

자동차용 침입탐지 트랙 - 2019

□ 예선 및 본선 경연

- 충분한 분석 시간이 주어지는 예선에서는 탐지 정확도가 높게 나타났음
- 예선과 유사하지만 다른 CAN message를 대상으로 하는 공격 또는 예선과 다른 이상징후 발생했을 경우, 이를 탐지하는데 어려움을 나타냄
  - 학습용 데이터의 비정상 데이터에 대한 Overfitting 발생
  - 이전에 적용된 공격데이터에 대해서만 알고리즘들이 잘 적용됨 (새로운 공격 주입 또는, 약간의 변형된 공격들은 성공됨)
  - 대다수의 본선 진출팀들이 예선에 비해 탐지 정확도가 확연히 떨어졌으며, 당일 데이터 분석 시간 동안에 이를 효과적으로 해결한 팀이 본선에서 1위 우승을 확보

©2020, Hacking and Countermeasure Research Lab

융합보안 챌린지 결과

자동차용 침입탐지 트랙 - 2019

□ 예선 및 본선 경연

- 본선에서 3번의 제출 시도 제한
  - 1, 2번째의 답안지 제출시 높은 정확도를 받지 못할 경우, 심리적 압박으로 제일 좋은 성적이 나왔던 답안지로 마지막 시도에 제출함
  - 정상 데이터에 대한 학습이 잘 이뤄졌을 경우, 학습 시간이 적더라도 Cost-effectiveness (가성비) 좋은 탐지 알고리즘 및 모델을 생성 가능

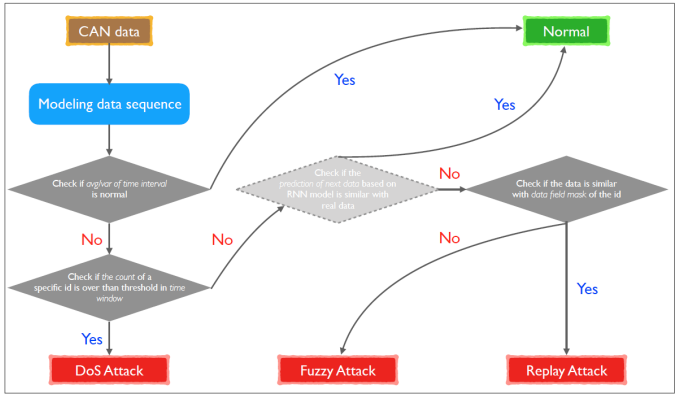
©2020, Hacking and Countermeasure Research Lab

기존 대회결과

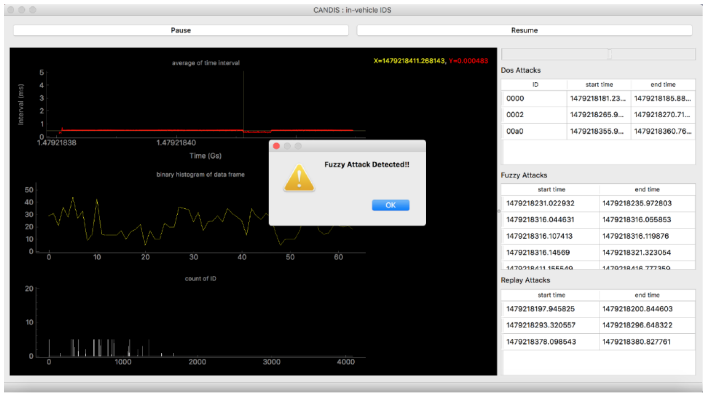
자동차용 침입탐지 트랙 – 2017 년 대회 사례

□ 예선 및 본선 경연

- 본선 1위 팀 - 방법론
  - 데이터 분석 및 CAN Data sequence modeling 을 통한 Recurrent Neural Network (RNN) 알고리즘을 이용하여 차량 이상징후를 탐지
    - 각 ID는 발생빈도가 고정되어 있기 때문에 정상적인 패킷의 빈도는 일정함
    - 공격시, 특정 ID와 데이터가 포함된 패킷이 고정된 주기를 벗어나 발생하며, 일부 ID의 sequence가 변경됨



[방법론]



[개발한 시각화 프로그램]

©2020, Hacking and Countermeasure Research Lab

기존 대회결과

자동차용 침입탐지 트랙 (2017 – 2019)

□ 예선 및 본선 경연

- 2019년 자동차용 침입탐지 트랙과는 다르게, 2017년에는 시각화 프로그램을 함께 제출 받음
  - 공격 또는 이상징후 탐지시, 이를 알려주는 알람 메시지 또는 그래프 분석에 대한 시각화 프로그램을 함께 제출하여, 공격에 대한 높은 정확도와 가시성을 확보함

- 경연 문제의 난이도 향상
  - 2017년도 경연 문제는 Time window 단위로 이상징후 여부를 탐지
    - 특정 시간대에 공격이 있었을 경우, 모두 공격으로 정의
  - 2019년도 경연 문제는 개별 CAN message 단위로 공격 여부를 탐지
    - 테스트 시, CAN message 에 대한 개별적인 공격 여부를 탐지/Labeling 을 수행해야 하므로, 더 어려운 문제가 됨 (Rule 기반으로 탐지 시, "정상,공격,정상"일 경우에 각각 CAN message의 Labeling이 어려워짐)
  - 예선에 학습 및 테스트용으로 사용되는 데이터의 크기를 2017년도 대비 1/8 수준으로 축소시켜 학습 시간 또는 탐지 알고리즘 적용 시간을 간략화함

©2020, Hacking and Countermeasure Research Lab

LESSONS LEARNED

©2020, Hacking and Countermeasure Research Lab

Lessons learned

제출 횟수 제한 및 시간 제한을 통한 현실적인 모델링 수행

□ 무제한 제출 방지 및 학습 시간을 제한적으로 운영

- 학습에 많은 시간이 걸리지 않고, 변형된 공격에 유효하게 적용될 수 있는 light-weight feature vector 발굴과 탐지 알고리즘 개발이 요구됨

- 무거운 탐지 모델 (Ex - deep learning 기반의 모델들) 들에 대한 제한된 시간내에서의 침입 탐지
  - Random Forest 로 98% 의 성능 / 10 mins vs. Deep learning 으로 99.9% 의 성능 / 60mins
  - 차량 환경에서는 어떤 결과물이 더 가치가 있을지 고려해야할 사항

©2020, Hacking and Countermeasure Research Lab

다양한 조건의 데이터셋

□ 3개 차종 및 4개 공격에 대한 탐지 알고리즘

- 차종의 수를 추가함으로써, 탐지 알고리즘의 일반성을 강화함
- 유사한 공격이지만, 조금의 변화를 준 공격에 대해서도 탐지 정확도가 높은 강건한 모델 개발이 요구됨

□ 정지 상태 및 주행 상태에 대한 데이터셋

- 본선 공격에 적용된 Replay 공격은 Data field의 값이 정상 상태에서의 Data field 값과 거의 유사하게 나타남. => Malfunction 및 Fuzzing 공격 탐지를 위한 Data field 값의 변화에 대한 탐지로는 한계가 있음
- 주행 상태의 데이터셋은 정상 상태의 데이터셋에 비해 Data field의 값이 크게 변화함.
  - 정상 상태에서의 Rule 기반 탐지 알고리즘은 한계가 있음 -> Machine Learning 기반의 알고리즘 요구

©2020, Hacking and Countermeasure Research Lab

향후 고려 사항

공격 데이터셋 변화

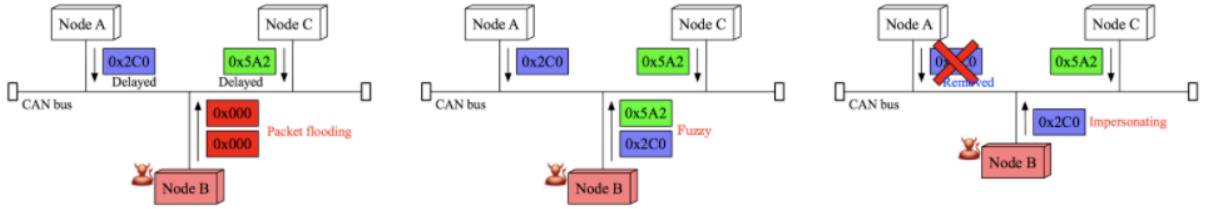
□ CAN bus에서의 새로운 공격을 개발할 필요

- 기존 공격 데이터셋은 다년간 국내외에서 reference case 로 여겨지고 있음
- 참가팀이 직접 다양한 공격 기법을 개발하여 release 할 경우 산업계에 기여하는 바가 커질 것으로 판단됨

1. Dataset

We provide datasets which include DoS attack, fuzzy attack, impersonation attack, and attack free states. Datasets were constructed by logging CAN traffic via the OBD-II port from a real vehicle while message injection attacks were performing.
We extracted the in-vehicle data from KIA SOUL.

1. DoS Attack : Injecting messages of '0x000' CAN ID in a short cycle.
2. Fuzzy Attack : Injecting messages of spoofed random CAN ID and DATA values.
3. Impersonation Attack : Injecting messages of Impersonating node, arbitration ID = '0x164'.
4. Attack Free State: Normal CAN messages.



(a) DoS attack (b) Fuzzy attack (c) Impersonation attack

향후 고려 사항

공격 / 방어 팀 구성을 통한 챌린지 구조 변화

- 각자 개발해 온 본선 진출 팀들간의 공격과 방어(탐지)를 통해 새로운 공격 패턴 생성
- 새로운 공격들에 대한 탐지가 가능하도록, 탐지 알고리즘의 고도화 가능

최신 차량에 대한 데이터셋 필요

- 노후된 차량의 데이터셋보다, 최신 차량의 내부 네트워크 데이터셋을 통한, 최신 차량의 데이터 패턴 확보 필요

©2020, Hacking and Countermeasure Research Lab

2020 챌린지형 과제 – 개념

참가팀이 직접 공격 기법 및 방어 기법 개발

- □ 개발한 공격 패턴을 운영진에 제출, 대회 시 참가팀들에 공격 진행
- □ 참가팀들은 방어자의 입장에서 타팀이 개발한 공격 패턴을 탐지할 수 있도록 함
- □ 실차에 차량용 IDS 와 공격 노트북을 동시에 연결하여 공/방을 진행

향후 자율주행 차량 (모형 또는 실차)으로 공격/방어 기법 확장

- □ Network based attack 이 있는 환경 하에서 방어를 성공적으로 수행하며 트랙 완주
- □ Network based Attack 과 Adversarial Machine Learning attack 과의 조합
- □ 강인함과 복원성 (Robustness & Resilience) 를 갖는 자율/무인이동체 보안 개발에 기여할 수 있음

©2020, Hacking and Countermeasure Research Lab

데이터셋 Access

차량 주행 데이터셋 (Know Your Master dataset)

- Original dataset: http://ocslab.hksecurity.net/Datasets/datachallenge2019/car
- Description (예선): http://datachallenge.kr/challenge19/convergence-security/vehicle/introduction/
- 예선/본선 데이터셋 및 정답지 등: http://ocslab.hksecurity.net/Datasets/datachallenge2019/car

©2020, Hacking and Countermeasure Research Lab

Thank You

KOREA UNIVERSITY

HCRL
Hacking and Countermeasure Research Lab