CAN Anomaly Detection Method using Frequency Analysis and Random Forest

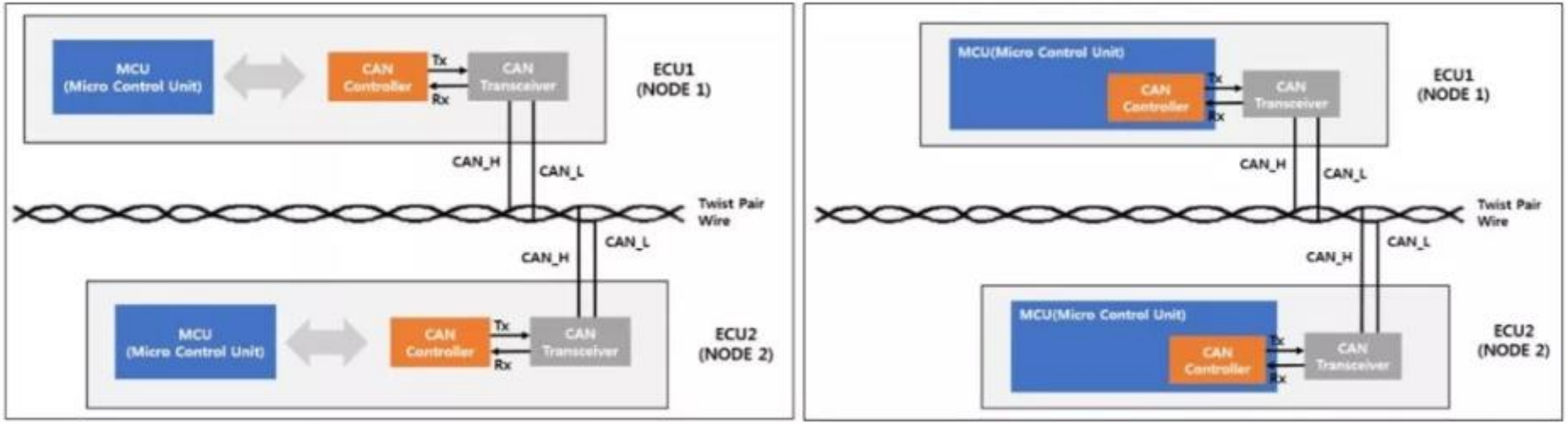
2019.12.05

최광준

목차

1. What is CAN(Controller Area Network)?

- 차량 내에서 호스트 컴퓨터 없이 마이크로 컨트롤러나 장치들이 서로 통신하기 위해 설계된 표준 통신규격
- 차량내 ECU(Electronic Control Unit)들은 **CAN 프로토콜**을 사용해 통신



Case1. MCU에 CAN Controller가 내장되지 않은 경우

Case2. MCU에 CAN Controller가 내장된 경우

출처: https://www.fescaro.com/ko/archives/249

2. Features of CAN

- ID는 메시지의 우선순위를 의미하며, 낮은 ID 번호가 더 높은 우선순위를 가지며 전송됨

1. 메시지 지향성 프로토콜
    I. CAN은 노드의 주소에 의해 데이터 교환 X
    II. 메시지의 우선순위에 따라 ID를 할당하고, 이 **ID를 이용해 메시지를 구별**하는 방식
    III. 임의의 한 노드 A가 메시지를 전송했다면, A를 제외한 나머지 노드들은 A가 **전송한 메시지가 자신에게 필요한지 여부를 ID기반으로 판단**
2. 보완적인 에러 감지 메커니즘
3. 멀티 마스터 능력
    I. …동시전송시, **더 낮은 ID 번호가 더 높은 우선순위를 가지며 우선 전송**됨
4. 결점이 있는 노드의 감지와 비활성화
    I. 실시간으로 결함있는 노드를 감지해 해당 노드를 비활성화 함으로써 네트워크의 신뢰성 보장
5. 전기적 노이즈에 강함
6. 저렴한 가격 및 구성의 용이성

출처: https://www.fescaro.com/ko/archives/249

3. Background and Related works

- 침입상태에서는, CAN 메시지의 분포가 달라진다.
- 침입상태에서는, CAN 메시지의 우선순위를 결정하는 CAN ID의 Time interval이 달라진다.
- 침입상태에서는, CAN ID의 Sequence가 달라진다.

I. … flow-based method evaluates several parameters using **frequency and the average of CAN message occurrence**
II. … the **time interval of the CAN ID** under injection attacks shorter than in the normal status
III. … detected anomalous status by using a transition matrix defined as patter of the **reiterative CAN ID sequence**
IV. … proposed an anomaly detector based on **long short term memory using a RNN** trained to predict the **CAN ID**

출처: 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)

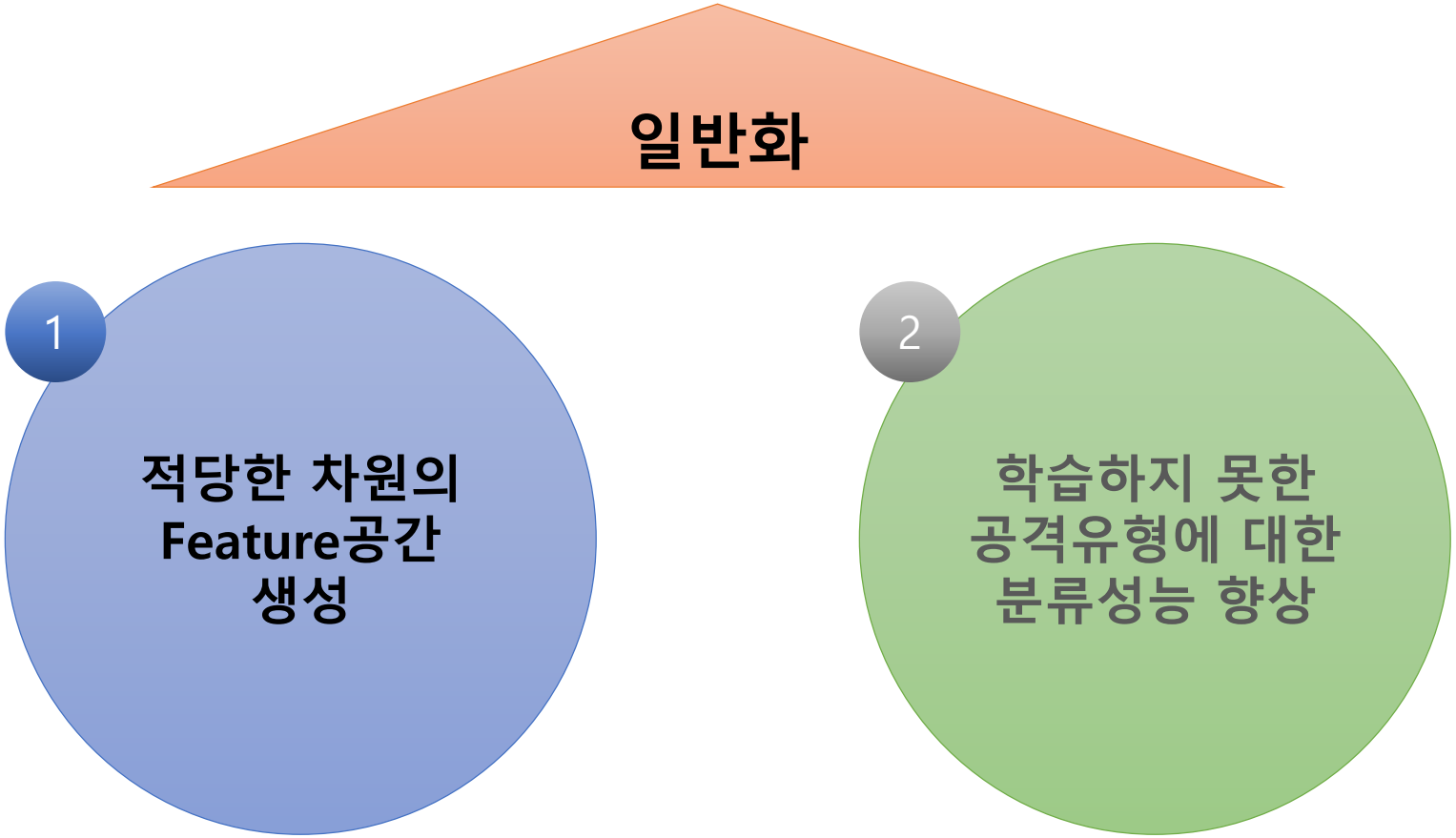
3. Background and Related works

- 침입상태에서는, 정상 상태에서 발생하는 CAN 메시지와 그 우선순위와 다른 형태를 보일 수 있다.

I. … the fuzzy attack is an attack that **randomly injects compromised ID, DLC, and Data fields** …
II. … the replay attack causes a problem by **injecting a set of CAN messages extracted and logged in a certain order** into the vehicle networks.
III. … the flooding attack … injecting a large number of messages **with the CAN ID set to 0x000** into the vehicle networks.
IV. … the malfunction attack **targets a selected CAN ID** from among the extractable CAN IDs of a certain vehicle.

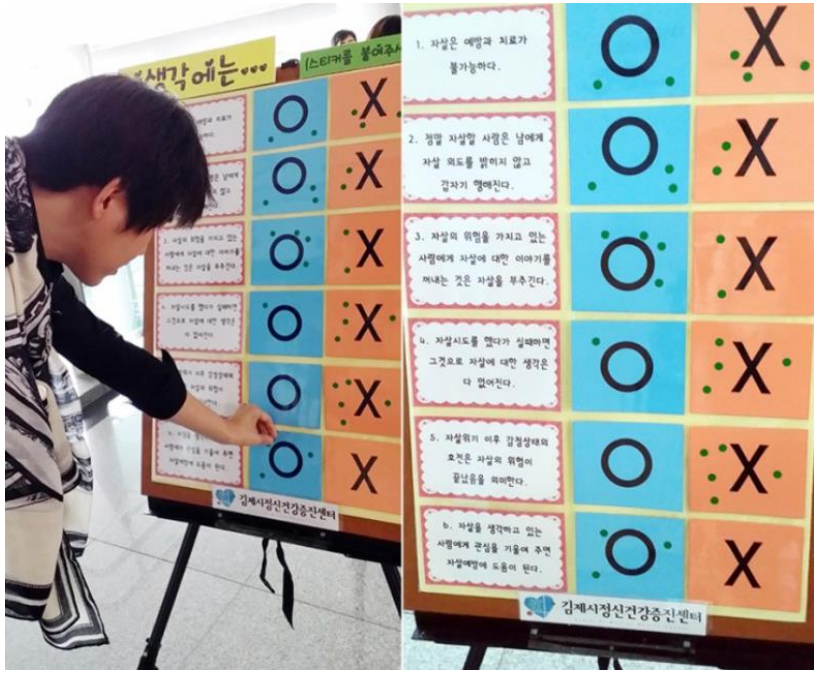
출처: Mee Lan Han, Byung Il Kwak, and Huy Kang Kim. "Anomaly intrusion detection method for vehicular networks based on survival analysis." Vehicular Communications 14 (2018)

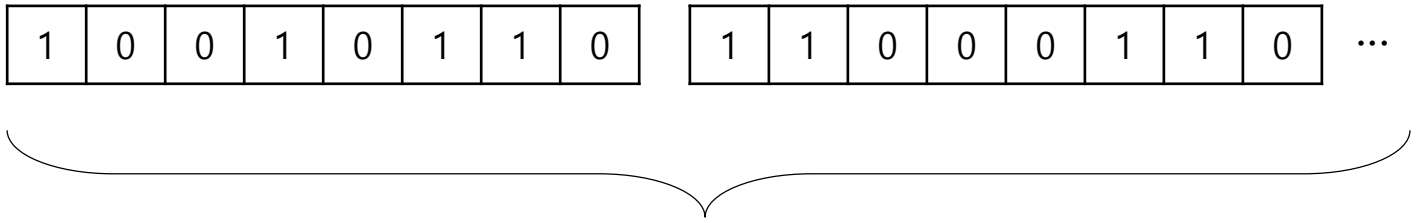
4. EDA and Feature engineering



4. EDA and Feature engineering

- 만약에 Data field의 정보를 bit 단위 Feature로 생성한다면?





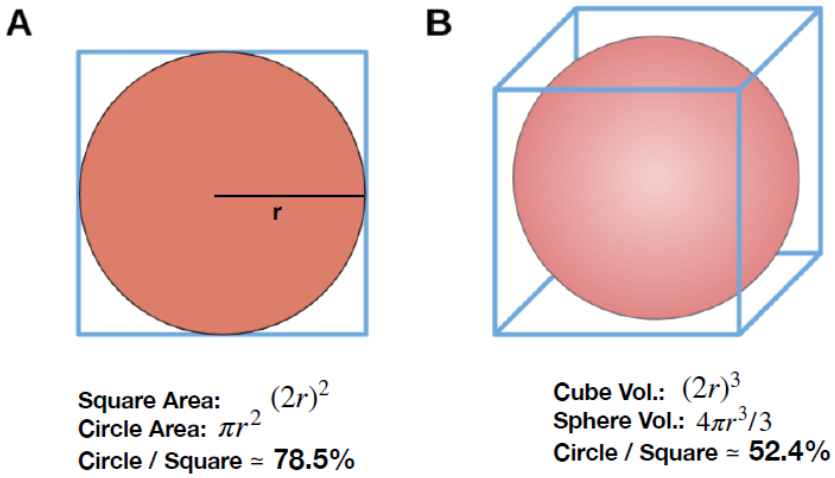
Data Field

출처: https://m.blog.naver.com/PostView.nhn?blogId=jbgokr&logNo=130189048126&proxyReferer=https%3A%2F%2Fwww.google.com%2F

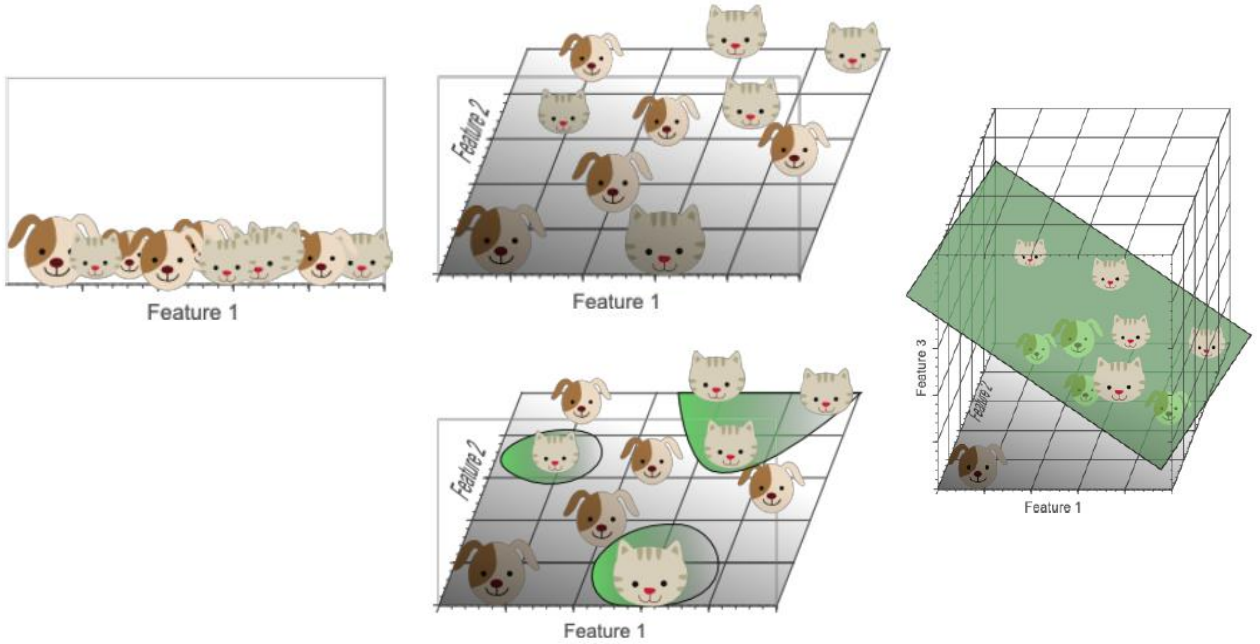
4. EDA and Feature engineering

- 차원을 높일수록 분류는 쉬워지나, 훈련데이터에 과적합 될 가능성이 큼

“Geometric Insanity”



Cat vs. Dog Classifier



출처: https://www.freecodecamp.org/news/the-curse-of-dimensionality-how-we-can-save-big-data-from-itself-d9fa0f872335/

4. EDA and Feature engineering

- 정상상태에서의 CAN 메시지는 그 내용에 맞는 우선순위를 가지며, 침입 상황에서는 다른 형태를 보일 것이다.
- 정상상태에서 일반적인 CAN ID Sequence 및 Time Interval이 존재하며, 침입 상황에서는 다른 형태를 보일 것이다.
- 제출용 데이터에 차량 정보가 없으므로, 차량 dependent한 모델 생성은 불가하다.

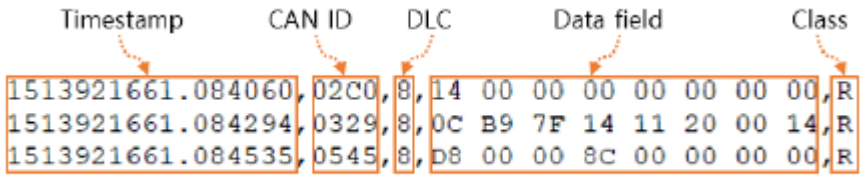


Figure 3. CAN Data – 학습용

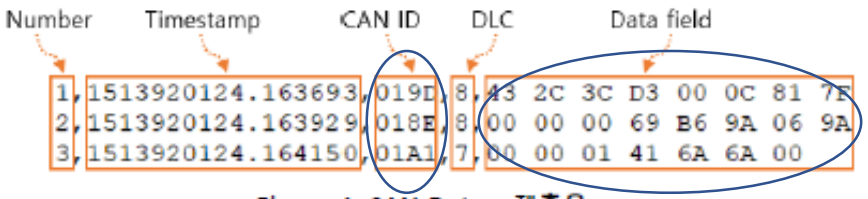


Figure 4. CAN Data - 제출용

✓ CAN ID와 Data Field를 10진수 숫자로 변경해 Feature 공간에 표현하자

✓ CAN ID의 Sequence를 반영하기 위해 Lag값을 Feature로 사용하자

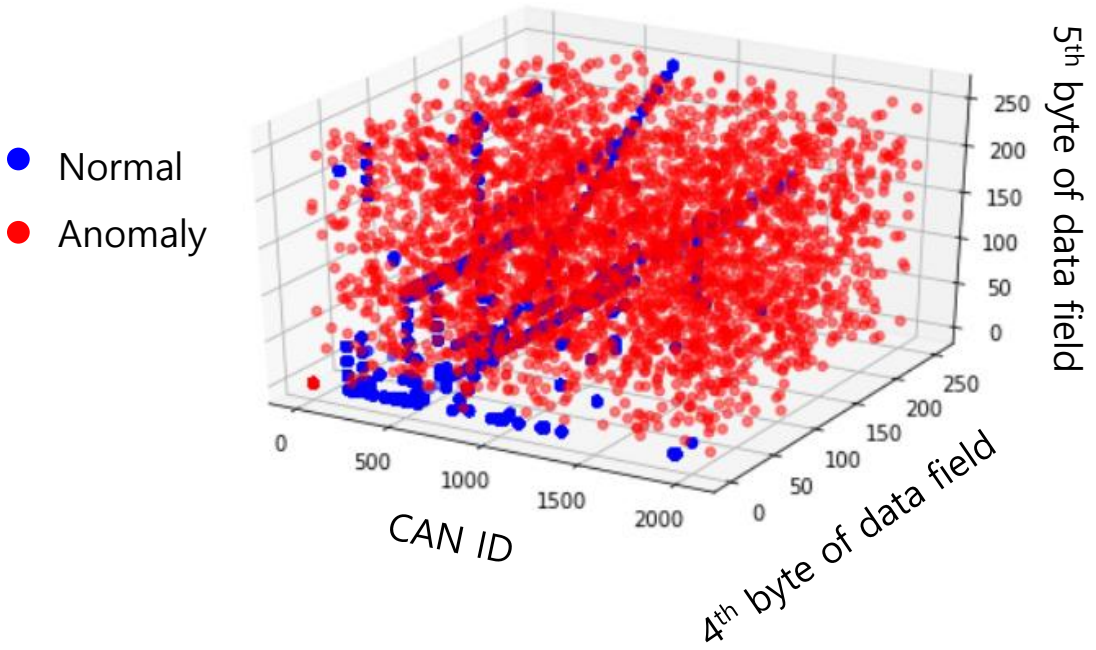
✓ Communication Interval을 반영하자

출처: 침입-탐지-트랙-데이터셋-설명.pdf, Competition 제공

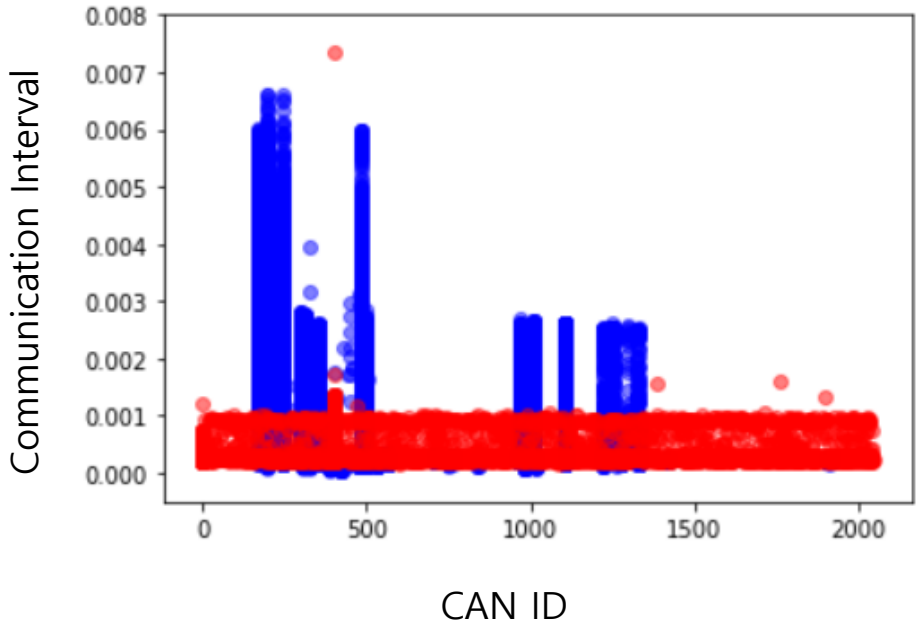
4. EDA and Feature engineering

- 정상상태에서의 CAN 메시지는 그 내용에 맞는 우선순위를 가지며, 침입 상황에서는 다른 형태를 보인다.
- 침입 상황에서는 대체로 통신 interval이 짧다.

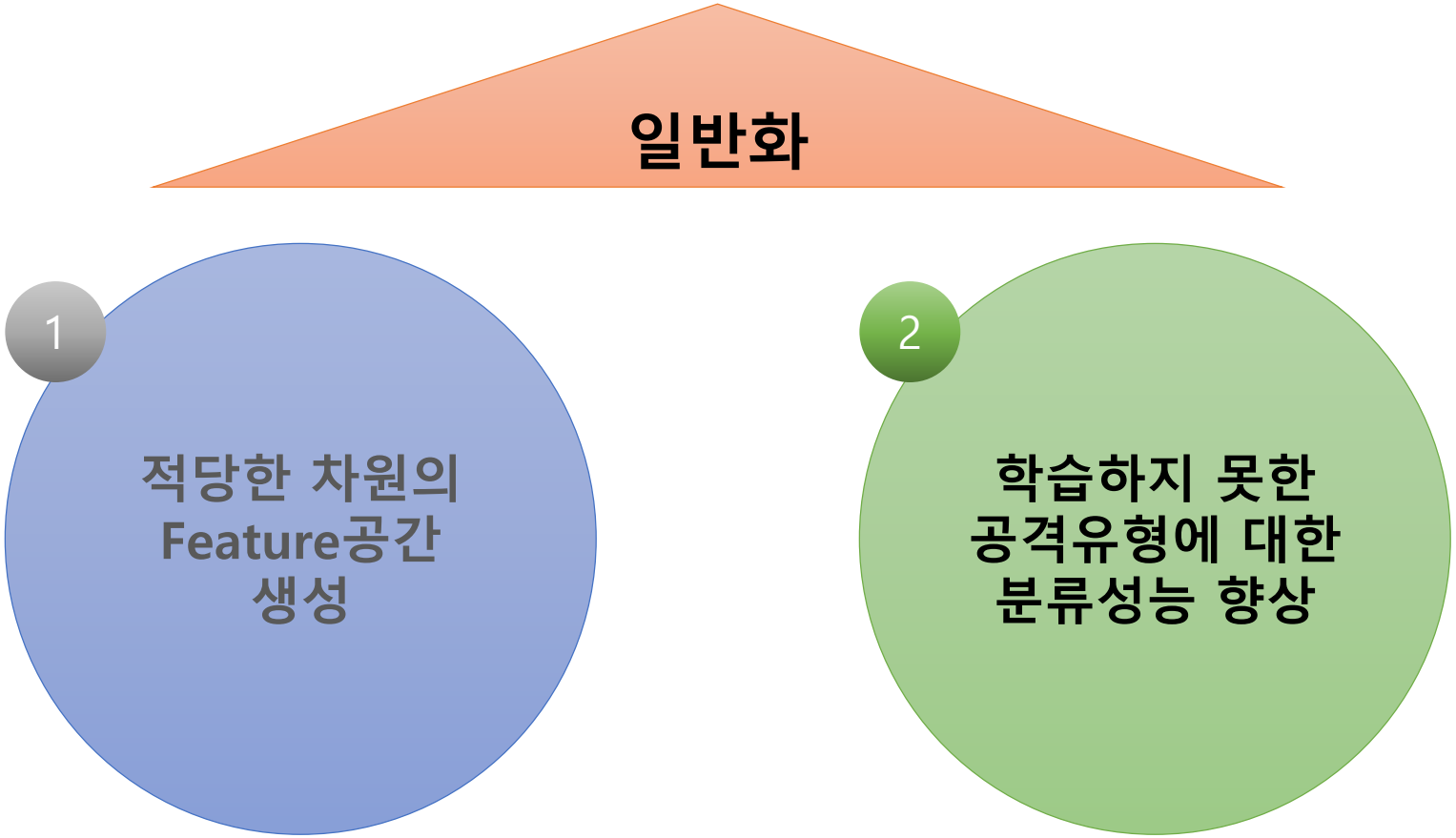
Spark 차량의 정상/침입 상태에서 CAN ID와 Data 분포



Spark 차량의 정상/침입 상태에서 CAN ID에 따른 통신 Interval



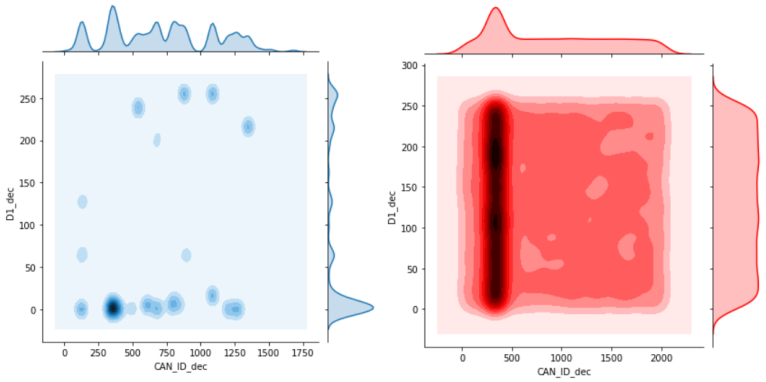
4. EDA and Feature engineering



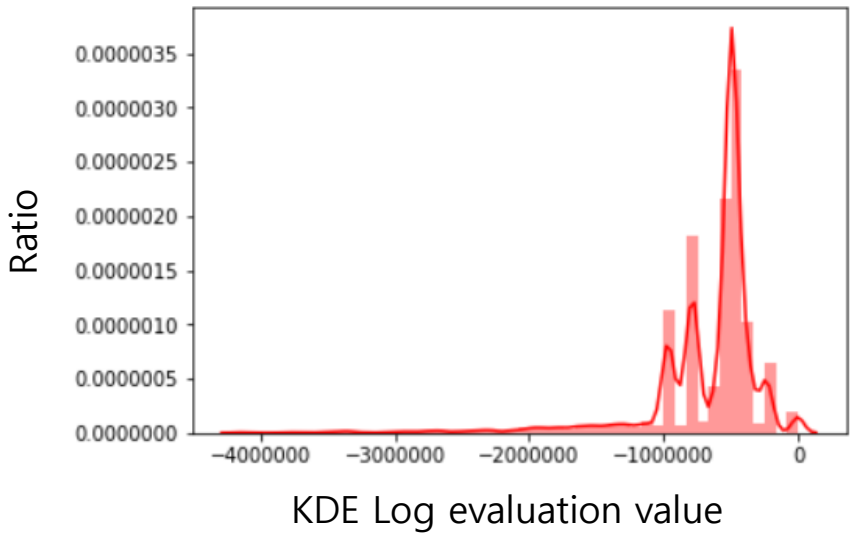
4. EDA and Feature engineering

- 모델이 경험하지 못한 공격 유형에 대해서도 탐지성능을 높일 수는 없을까?
- 동일한 공격 유형에 대해서도 탐지 성능을 높이기 위해 무엇을 할 수 있을까?
- Kernel density estimator를 사용

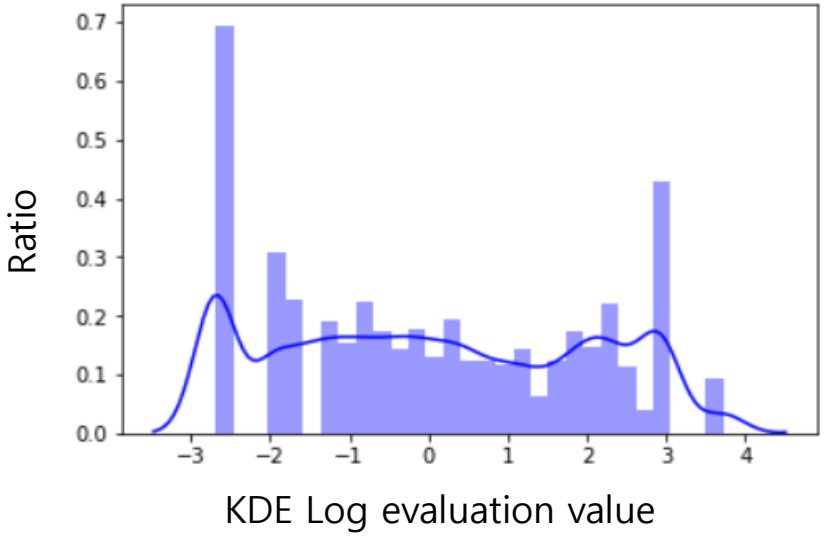
정상/침입에서 CAN ID와 Data Frame 분포



Spark 차량의 침입 상태에서 KDE density 분포

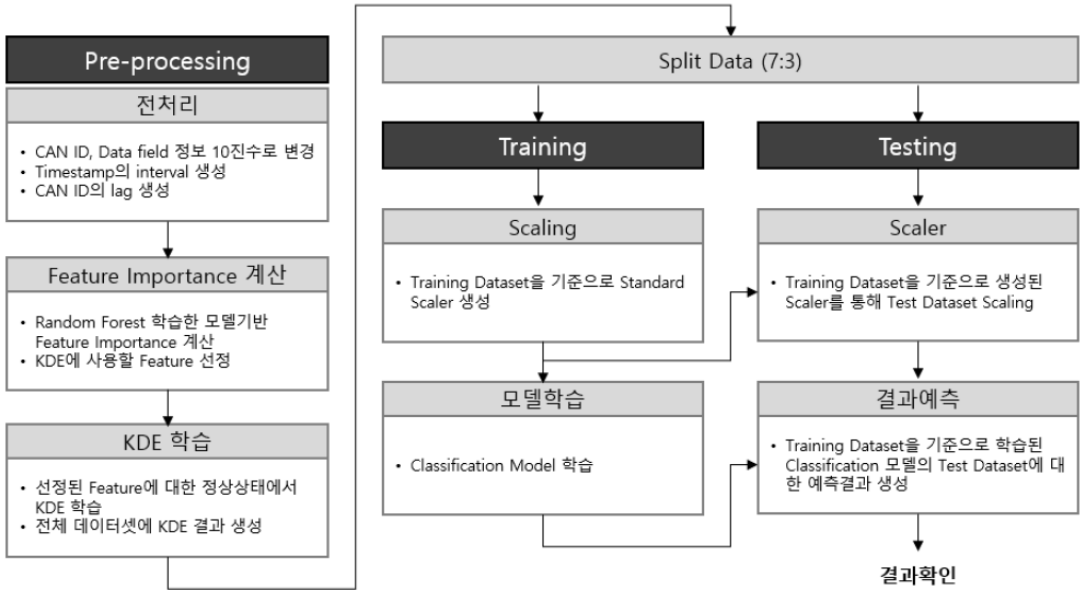


Spark 차량의 정상 상태에서 KDE density 분포

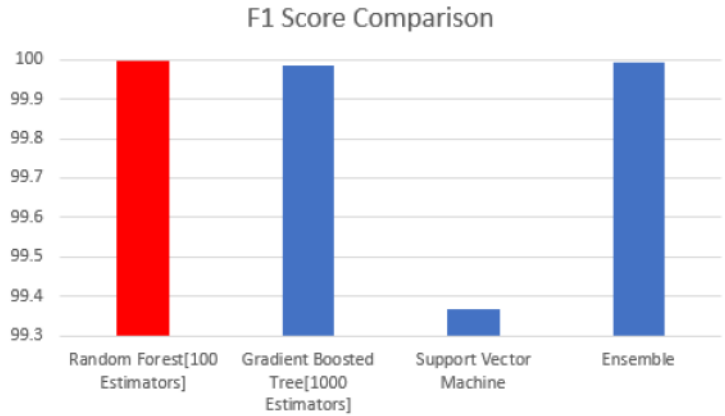


5. Results Comparison

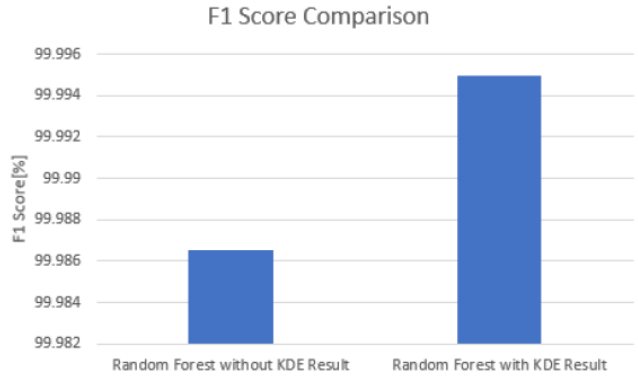
- 분류 성능이 가장 높은 Random Forest를 사용
- KDE 결과를 Feature로 추가한 경우 분류성능이 Slightly 향상되었으나 학습데이터에서 보지못한 공격유형을 분류하는 데 유용하게 작용할 수 있을 것으로 예상



모델학습 및 결과확인 프로세스



Classifier에 따른 성능 비교

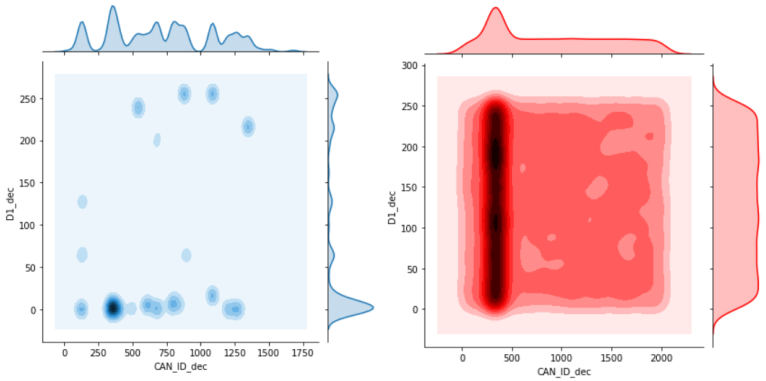


KDE 사용 유무에 따른 분류성능 비교

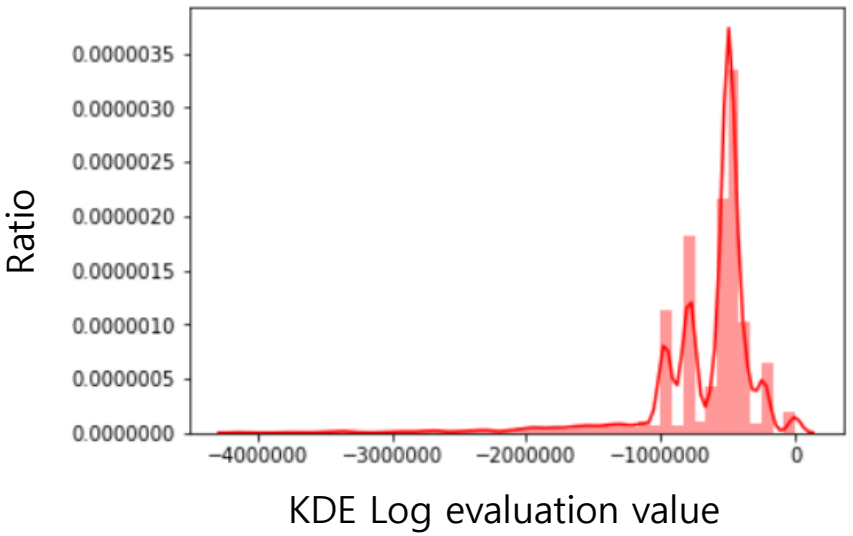
6. Conclusion and Future works

- 정상 데이터만을 기반으로 정상상태를 판단하는 모델을 생성
- 예를 들어, KDE를 사용한 밀도추정 방법 등

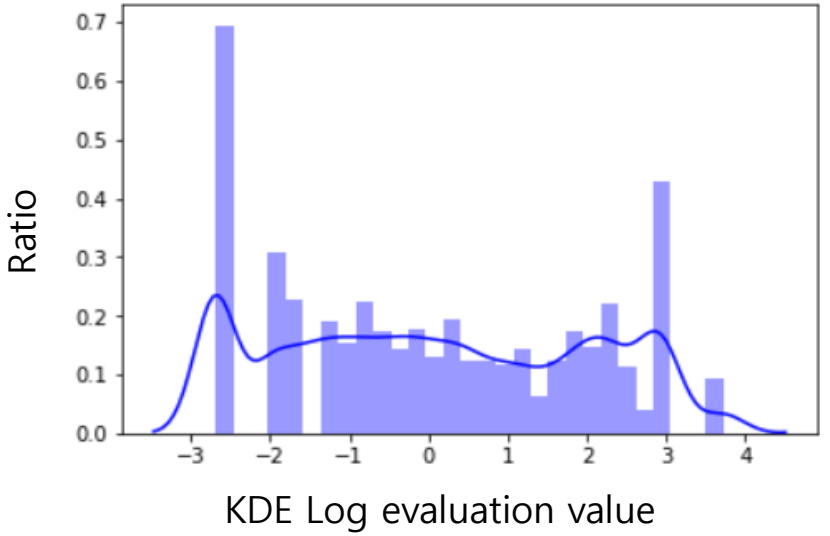
정상/침입에서 CAN ID와 Data Frame 분포 및 발생 빈도



Spark 차량의 침입 상태에서 KDE 밀도분포



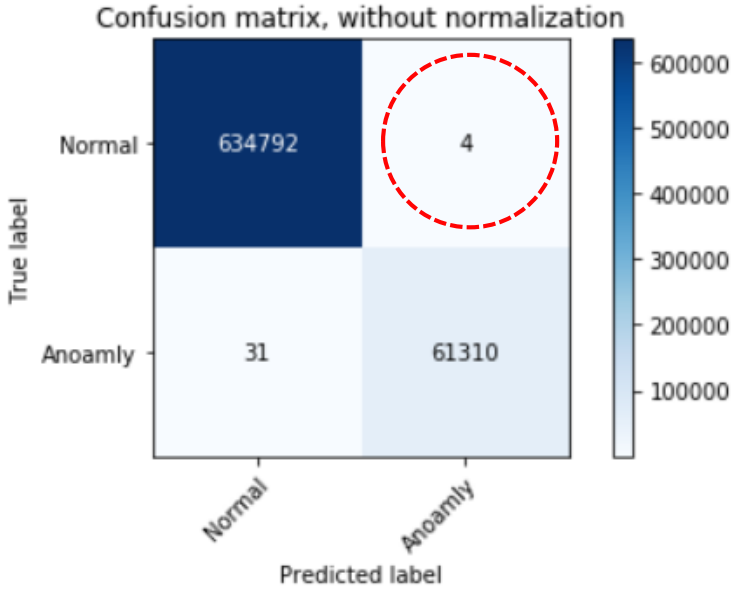
Spark 차량의 정상 상태에서 KDE 밀도분포



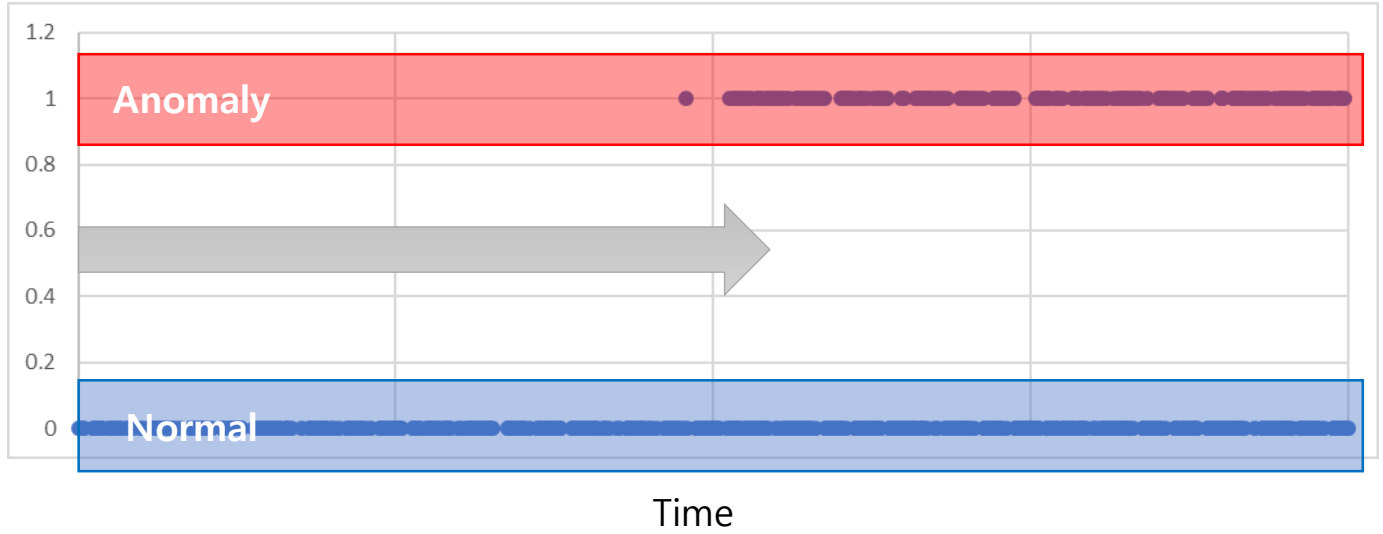
6. Conclusion and Future works

- Sequence 데이터라면 False Alarm을 제거하기 위해 Filtering을 적용할 수 있지 않을까?
- 모델 성능의 한계로 인한 False Alarm은 Random하게 발생하지만, 실제 이상은 반복적으로 발생하지 않을까?
- 시간축에서 이상발생 분포를 기반으로 실제 이상과 가짜이상을 구분해보면 어떨까?
- 실제 Application에서는 활용 가능할 것으로 예상되나, 감지시간 지연으로 Competition에서는 사용하지 않음

본선 1차 테스트셋 분류결과

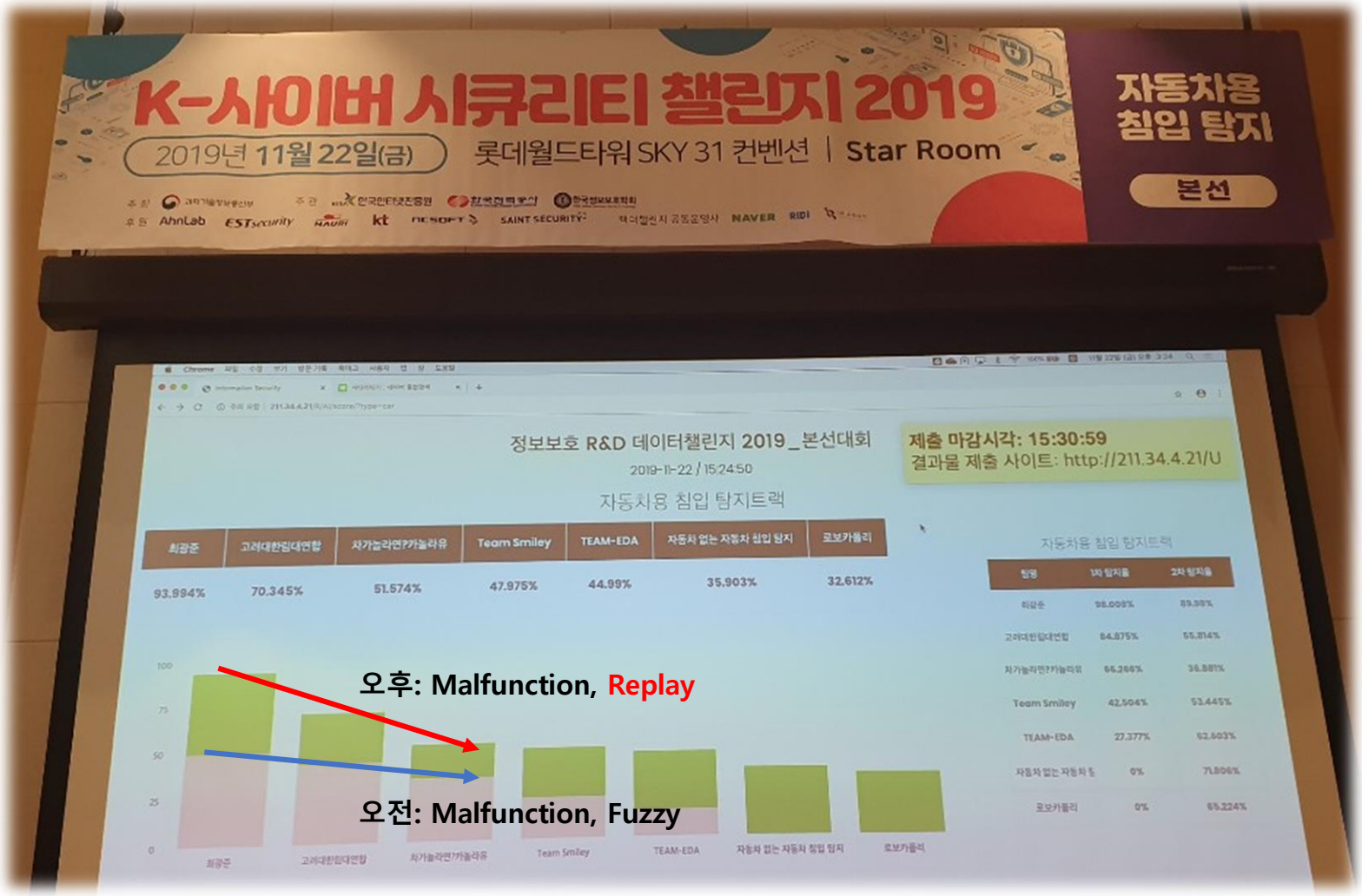


시간축에서 Soul 차량의 Fuzzy 공격발생 현황



7. Review

- 본선 시간의 한계 + 예선에 제공된 공격 유형에 최적화된 분석 프로세스 및 모델 한계
- KDE 사용이 상대적으로 유리하게 작용했을 것으로 예상됨



7. Review

- CAN Message의 Sequence를 고려할 수 있는 Feature 생성과 모델을 사용하면 Replay 공격에 대한 분류 정확도 향상 시킬 수 있을 것으로 예상됨

"The **replay attack** causes a problem by injecting a set of **CAN messages** extracted and logged **in a certain order** into the vehicle network"

출처: 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)

7. Q&A

Question & Answer