(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년03월09일
(11) 등록번호 10-1714520
(24) 등록일자 2017년03월03일

(51) 국제특허분류(Int. Cl.)
H04L 29/06 (2006.01) *H04L 12/40* (2006.01)
H04L 29/08 (2006.01)
(52) CPC특허분류
H04L 63/1416 (2013.01)
H04L 12/40104 (2013.01)
(21) 출원번호 10-2015-0151648
(22) 출원일자 2015년10월30일
심사청구일자 2015년10월30일
(56) 선행기술조사문헌
US20140250531 A1
JP2004312083 A

(73) 특허권자
현대자동차주식회사
서울특별시 서초구 헌릉로 12 (양재동)
기아자동차주식회사
서울특별시 서초구 헌릉로 12 (양재동)
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
정호진
서울특별시 금천구 시흥대로 165, 210동 904호(시흥동, 남서울힐스테이트아파트)
안현수
서울특별시 송파구 올림픽로35길 104, 27동 407호 (신천동, 장미아파트)
(뒷면에 계속)
(74) 대리인
박영복

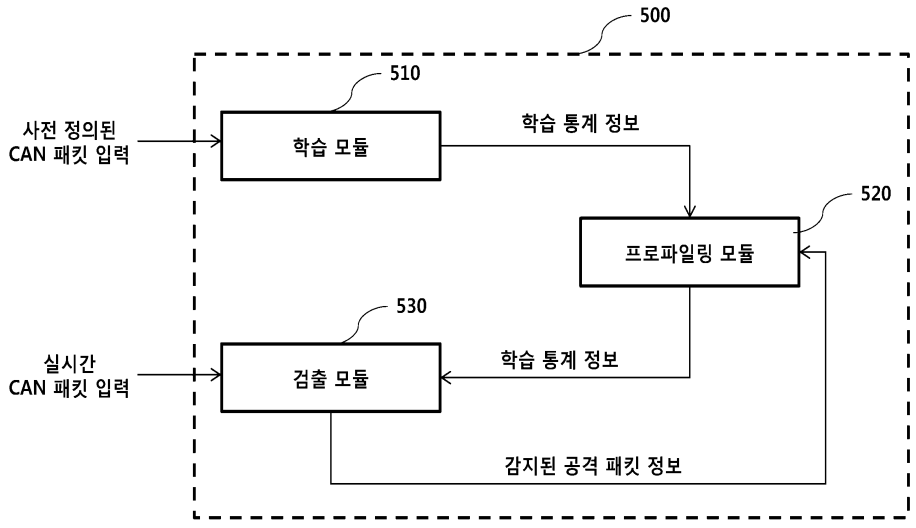
전체 청구항 수 : 총 19 항 심사관 : 문형섭

(54) 발명의 명칭 **차량 내 네트워크 공격 탐지 방법 및 장치**

(57) 요 약

본 발명은 차량 CAN 통신 공격 탐지 방법 및 장치에 관한 것으로서, 차량 CAN(Controller Area Network) 통신 공격을 탐지하는 장치는 미리 정의된 CAN 패킷이 입력되면 CAN ID를 식별하고, 상기 식별된 CAN ID에 대응되는 데이터 필드의 비트 단위 표현의 학습 통계 정보를 생성하는 학습 모듈과 상기 학습 모듈로부터 수신되는 상기 학습 통계 정보 및 미리 수집된 공격 유형이 유지되는 프로파일링 모듈과 실시간 입력된 CAN 패킷으로부터 CAN ID 및 데이터 필드의 비트 열을 추출하고, 상기 추출된 CAN ID에 대응되는 상기 학습 통계 정보를 상기 프로파일링 모듈로부터 획득하여 신경망을 구성하고, 상기 구성된 신경망에 상기 추출된 비트 열을 입력하고, 상기 신경망의 출력 값에 기반하여 공격 여부를 판단하는 검출 모듈을 포함할 수 있으며, 그에 따라 보다 효과적이고 정확하게 CAN 통신 공격을 탐지할 수 있는 장점이 있다.

대 표 도 - 도5



(52) CPC특허분류
H04L 67/30 (2013.01)
H04L 2012/40215 (2013.01)
H04L 2012/40273 (2013.01)
(72) 발명자
강제원
서울특별시 마포구 독막로 145, 110동 502호(창전동, 서강쌍용예가)
강민주
서울특별시 서대문구 이화여대2다길 15, 봉하숙 지하2호(대현동)

명 세 서

청구범위

청구항 1

차량 CAN(Controller Area Network) 통신 공격을 탐지하는 장치에 있어서,

미리 정의된 CAN 패킷이 입력되면 CAN ID를 식별하고, 상기 식별된 CAN ID에 대응되는 데이터 필드의 비트 단위 표현의 학습 통계 정보를 생성하는 학습 모듈;

상기 학습 모듈로부터 수신되는 상기 학습 통계 정보 및 미리 수집된 공격 유형이 유지되는 프로파일링 모듈; 및

실시간 입력된 CAN 패킷으로부터 CAN ID 및 데이터 필드의 비트 열을 추출하고, 상기 추출된 CAN ID에 대응되는 상기 학습 통계 정보를 상기 프로파일링 모듈로부터 획득하여 신경망을 구성하고, 상기 구성된 신경망에 상기 추출된 비트 열을 입력하고, 상기 신경망의 출력 값에 기반하여 공격 여부를 판단하고, 상기 출력 값에 따라 공격인 것으로 판단되면, 상기 실시간 입력된 CAN 패킷을 상기 프로파일링 모듈에 전송하는 검출 모듈

을 포함하고, 상기 프로파일링 모듈이 상기 공격이 감지된 CAN 패킷을 이용하여 공격 유형을 갱신하는, 차량 CAN 통신 공격 탐지 장치.

청구항 2

삭제

청구항 3

제1항에 있어서,

상기 검출 모듈은 상기 실시간 입력된 CAN 패킷이 상기 공격 유형에 해당되는 CAN 패킷인지 여부를 판단하여 상기 공격 여부를 판단하는 수단을 더 포함하는, 차량 CAN 통신 공격 탐지 장치.

청구항 4

제1항에 있어서,

상기 학습 모듈은 상기 비트 단위 표현의 학습 통계 정보에 기반하여 특징 점을 추출하는, 차량 CAN 통신 공격 탐지 장치.

청구항 5

제4항에 있어서,

상기 학습 모듈은 상기 데이터 필드 내 비트 열 중 적어도 하나의 일부 특정 비트 열을 상기 특징 점으로 선택하는, 차량 CAN 통신 공격 탐지 장치.

청구항 6

제5항에 있어서,

상기 검출 모듈은 상기 선택된 특징 점에 상응하는 비트 열을 상기 구성된 신경망에 입력하는, 차량 CAN 통신 공격 탐지 장치.

청구항 7

제5항에 있어서,

상기 특징 점은 모드 정보를 표현하는 비트 열 및 수치 정보를 표현하는 비트 열 중 적어도 하나를 포함하는, 차량 CAN 통신 공격 탐지 장치.

청구항 8

제4항에 있어서,

상기 학습 모듈은 시간에 따른 상기 데이터 필드 내 비트 열의 비트 단위 변화 량에 기반하여 상기 특징 점을 비트 단위로 선택하는, 차량 CAN 통신 공격 탐지 장치.

청구항 9

제8항에 있어서,

상기 비트 단위 변화 량은 연속된 두 CAN 패킷의 데이터 필드 영역을 비트 단위 exclusive-OR 연산을 통해 산출되는, 차량 CAN 통신 공격 탐지 장치.

청구항 10

제1항에 있어서,

상기 미리 정의된 CAN 패킷은 정상 상태의 CAN 패킷과 비정상 상태의 CAN 패킷을 포함하는, 차량 CAN 통신 공격 탐지 장치.

청구항 11

차량 CAN(Controller Area Network) 통신 공격을 탐지하는 장치에 있어서,

사전 정의된 CAN 패킷이 입력되면, CAN 데이터 프레임을 구성하는 필드 단위로 설정된 템플릿 영역에 매칭되는 상기 사전 정의된 CAN 패킷의 비트 열을 추출하여 비트 단위 표현의 학습 통계 정보를 생성하는 학습 모듈;

상기 생성된 학습 통계 정보 및 미리 수집된 공격 유형이 유지되는 프로파일링 모듈; 및

실시간 입력된 CAN 패킷으로부터 상기 템플릿 영역에 매칭되는 비트 열을 추출하고, 상기 템플릿 영역에 대응하여 기 생성된 상기 학습 통계 정보를 이용하여 신경망을 구성하고, 상기 구성된 신경망에 상기 추출된 비트 열을 입력하고, 상기 신경망의 출력 값에 기반하여 공격 여부를 판단하는 검출 모듈

을 포함하는, 차량 CAN 통신 공격 탐지 장치.

청구항 12

차량 CAN(Controller Area Network) 통신 공격을 탐지하는 방법에 있어서,

미리 정의된 CAN 패킷이 입력되면 CAN ID를 식별하고, 상기 식별된 CAN ID에 대응되는 데이터 필드의 비트 단위 표현의 학습 통계 정보를 생성하여 저장하는 단계;

실시간 입력된 CAN 패킷으로부터 CAN ID 및 데이터 필드의 비트 열을 추출하는 단계;

상기 추출된 CAN ID에 대응되는 상기 학습 통계 정보에 기반하여 신경망을 구성하는 단계;

상기 구성된 신경망에 상기 추출된 비트 열을 입력하고, 상기 신경망의 출력 값에 기반하여 공격 여부를 판단하는 단계; 및

소정 기록 영역에 기 수집된 공격 유형에 대한 정보가 유지되되, 상기 출력 값에 따라 공격인 것으로 판단되면, 상기 실시간 입력된 CAN 패킷을 이용하여 상기 공격 유형을 갱신되는 단계

를 포함하는, 차량 CAN 통신 공격 탐지 방법.

청구항 13

삭제

청구항 14

제12항에 있어서,

상기 실시간 입력된 CAN 패킷이 상기 공격 유형에 해당되는 CAN 패킷인지 여부를 판단하여 공격을 감지하는 단

계를 더 포함하는, 차량 CAN 통신 공격 탐지 방법.

청구항 15

제12항에 있어서,

상기 비트 단위 표현의 학습 통계 정보에 기반하여 특징 점을 추출하는 단계를 더 포함하는, 차량 CAN 통신 공격 탐지 방법.

청구항 16

제15항에 있어서,

상기 특징 점을 추출하는 단계는

상기 데이터 필드 내 비트 열 중 적어도 하나의 일부 특정 비트 열을 상기 특징 점으로 선택하는 단계; 및

시간에 따른 상기 데이터 필드 내 비트 열의 비트 단위 변화 량에 기반하여 상기 특징 점을 비트 단위로 선택하는 단계

중 적어도 하나를 포함하는, 차량 CAN 통신 공격 탐지 방법.

청구항 17

제16항에 있어서,

상기 특징 점으로 선택된 비트 열 또는 비트가 상기 구성된 신경망에 입력되는, 차량 CAN 통신 공격 탐지 방법.

청구항 18

제16항에 있어서,

상기 일부 특정 비트 열을 상기 특징 점으로 선택되는 경우, 상기 특징 점은 모드 정보를 표현하는 비트 열 및 수치 정보를 표현하는 비트 열 중 적어도 하나를 포함하는, 차량 CAN 통신 공격 탐지 방법.

청구항 19

제16항에 있어서,

상기 비트 단위 변화 량은 연속된 두 개의 CAN 패킷의 데이터 필드 영역을 비트 단위 exclusive-OR 연산하여 산출되는, 차량 CAN 통신 공격 탐지 방법.

청구항 20

차량 CAN(Controller Area Network) 통신 공격을 탐지하는 방법에 있어서,

사전 정의된 CAN 패킷이 입력되면, CAN 데이터 프레임을 구성하는 필드 단위로 설정된 템플릿 영역에 매칭되는 상기 사전 정의된 CAN 패킷의 비트 열을 추출하여 비트 단위 표현의 학습 통계 정보를 생성하여 저장하는 학습 단계; 및

실시간 입력된 CAN 패킷으로부터 상기 템플릿 영역에 매칭되는 비트 열을 추출하고, 상기 템플릿 영역에 대응하여 기 생성된 상기 학습 통계 정보를 이용하여 신경망을 구성하고, 상기 구성된 신경망에 상기 추출된 비트 열을 입력하고, 상기 신경망의 출력 값에 기반하여 공격 여부를 판단하는 검출 단계

를 포함하는, 차량 CAN 통신 공격 탐지 방법.

청구항 21

제12항, 제14항 내지 제20항 중 어느 한 항에 기재된 방법을 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

발명의 설명

기 술 분 야

[0001] 본 발명은 차량 내 통신 보안에 관한 것으로서, 상세하게, 차량 내 통신 네트워크상에 존재하는 패킷의 통계적 특성에 기반하여 공격을 탐지하는 방법 및 그를 위한 장치에 관한 것이다.

배 경 기 술

[0002] 미래형 자동차는 차량 내/외부에 탑재된 각종 센서들과 연동되는 다수의 전자 제어 장치(ECU: Electronic Control Unit)를 장착하여 CAN(Controller Area Network)과 같은 차량 내 네트워크를 이용하여 통신을 수행함으로써 차량의 각종 기능을 스마트하게 제어하여 사용자의 편의성과 주행 효율성을 증진시키는 것이 주된 목적이다.

[0003] 최근 정보통신 분야의 사물인터넷 (IoT: Internet of Things, IoT) 기술 분야에서는 플랫폼으로 스마트카 등을 이용하기 위한 연구가 활발히 진행 중에 있다.

[0004] 또한, 최근에는 지능형 교통망 서비스를 제공하기 위한 차량간 통신(Vehicle to Vehicle Communication), 차량 기간 망 통신(Vehicle to Infrastructure Communication) 등에 대한 표준화 및 기술 개발이 활발히 진행되고 있다.

[0005] 이에 따라, 미래의 스마트 차량에 대한 공격 형태는 종래에 비해 보다 많은 통신 패킷이 존재할 뿐만 아니라 다양한 외부 장치와의 통신을 수행하므로, 해킹 유형이 다양해지고 분산적인 형태로 진화할 것이 예상되고 있다.

[0006] 하지만, 종래에는 차량 내 사용 가능한 자원은 한정적인 반면, 해킹 유형은 보다 다양하고 분산적이므로 효과적인 대응이 쉽지 않은 문제점이 있었다.

[0007] 특히, 차량의 정지 상태뿐만 아니라 주행 상태에서의 보안 해킹 및 공격은 운전자의 프라이버시뿐만 아니라 주행 안전에 치명적인 영향을 미칠 수 있다. 일 예로, 해킹에 의해 주행 중 엔진이 정지하거나, 조향 핸들/브레이크 등이 오동작하는 경우 운전자의 안전에 치명적일 수 있다.

[0008] 따라서, 상기와 같은 다양한 차량에 대한 해킹 공격을 조기에 감지하고, 악의적인 공격에 대해 적절히 대처하기 위한 효율적인 알고리즘 및 시스템 개발이 요구되고 있다.

발명의 내용

해결하려는 과제

[0009] 본 발명은 상술한 종래 기술의 문제점을 해결하기 위해 고안된 것으로, 본 발명의 목적은 차량 내 네트워크 공격 탐지 방법 및 장치를 제공하는 것이다.

[0010] 본 발명의 다른 목적은 사전 학습된 CAN 패킷의 통계적 특성에 기반하여 적응적으로 공격을 탐지하는 방법 및 그를 위한 장치를 제공하는 것이다.

[0011] 본 발명의 또 다른 목적은 CAN 통신 패킷을 복호하지 않고 이진 데이터의 통계적 특성에 기반하여 적응적으로 공격을 탐지하는 것이 가능한 CAN 공격 탐지 방법 및 그를 위한 장치를 제공하는 것이다.

[0012] 본 발명에서 이루고자 하는 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급하지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0013] 본 발명은 차량 CAN 통신 공격 탐지 방법 및 장치를 제공한다.

[0014] 본 발명의 일 실시예에 따른 차량 CAN(Controller Area Network) 통신 공격을 탐지하는 장치는 미리 정의된 CAN 패킷이 입력되면 CAN ID를 식별하고, 상기 식별된 CAN ID에 대응되는 데이터 필드의 비트 단위 표현의 학습 통계 정보를 생성하는 학습 모듈과 상기 학습 모듈로부터 수신되는 상기 학습 통계 정보 및 미리 수집된 공격 유형이 유지되는 프로파일링 모듈과 실시간 입력된 CAN 패킷으로부터 CAN ID 및 데이터 필드의 비트 열을 추출하고, 상기 추출된 CAN ID에 대응되는 상기 학습 통계 정보를 상기 프로파일링 모듈로부터 획득하여 신경망을 구성하고, 상기 구성된 신경망에 상기 추출된 비트 열을 입력하고, 상기 신경망의 출력 값에 기반하여 공격 여부

를 판단하는 검출 모듈을 포함할 수 있다.

[0015] 또한, 상기 검출 모듈은 상기 출력 값에 따라 공격인 것으로 판단되면, 해당 공격 CAN 패킷을 프로파일링 모듈에 전송하되, 상기 공격 CAN 패킷을 이용하여 상기 프로파일링 모듈이 상기 공격 유형을 갱신할 수 있다.

[0016] 또한, 상기 검출 모듈은 상기 실시간 입력된 CAN 패킷이 상기 공격 유형에 해당되는 CAN 패킷인지 여부를 판단하여 공격을 감지할 수 있다.

[0017] 또한, 상기 학습 모듈은 상기 비트 단위 표현의 학습 통계 정보에 기반하여 특징 점을 추출할 수 있다.

[0018] 여기서, 상기 학습 모듈은 상기 데이터 필드 내 비트 열 중 적어도 하나의 일부 특정 비트 열을 상기 특징 점으로 선택할 수 있다.

[0019] 또한, 상기 검출 모듈은 상기 선택된 특징 점에 상응하는 비트 열을 상기 구성된 신경망에 입력할 수 있다.

[0020] 또한, 상기 특징 점은 모드 정보를 표현하는 비트 열 및 수치 정보를 표현하는 비트 열 중 적어도 하나를 포함할 수 있다.

[0021] 또한, 상기 학습 모듈은 시간에 따른 상기 데이터 필드 내 비트 열의 비트 단위 변화 량에 기반하여 상기 특징 점을 비트 단위로 선택할 수도 있다.

[0022] 여기서, 상기 비트 단위 변화 량은 연속된 두 CAN 패킷의 데이터 필드 영역을 비트 단위 exclusive-OR 연산을 통해 산출될 수 있다.

[0023] 또한, 상기 미리 정의된 CAN 패킷은 정상 상태의 CAN 패킷과 비정상 상태의 CAN 패킷을 포함할 수 있다.

[0024] 본 발명의 다른 일 실시예에 따른 차량 CAN(Controller Area Network) 통신 공격을 탐지하는 장치는 사전 정의된 CAN 패킷이 입력되면, CAN 데이터 프레임을 구성하는 필드 단위로 설정된 템플릿 영역에 매칭되는 상기 사전 정의된 CAN 패킷의 비트 열을 추출하여 비트 단위 표현의 학습 통계 정보를 생성하는 학습 모듈과 상기 생성된 학습 통계 정보 및 미리 수집된 공격 유형이 유지되는 프로파일링 모듈과 실시간 입력된 CAN 패킷으로부터 상기 템플릿 영역에 매칭되는 비트 열을 추출하고, 상기 템플릿 영역에 대응하여 기 생성된 상기 학습 통계 정보를 이용하여 신경망을 구성하고, 상기 구성된 신경망에 상기 추출된 비트 열을 입력하고, 상기 신경망의 출력 값에 기반하여 공격 여부를 판단하는 검출 모듈을 포함할 수 있다.

[0025] 본 발명의 또 다른 일 실시예에 따른 차량 CAN(Controller Area Network) 통신 공격을 탐지하는 방법은 미리 정의된 CAN 패킷이 입력되면 CAN ID를 식별하고, 상기 식별된 CAN ID에 대응되는 데이터 필드의 비트 단위 표현의 학습 통계 정보를 생성하여 저장하는 단계와 실시간 입력된 CAN 패킷으로부터 CAN ID 및 데이터 필드의 비트 열을 추출하는 단계와 상기 추출된 CAN ID에 대응되는 상기 학습 통계 정보에 기반하여 신경망을 구성하는 단계와 상기 구성된 신경망에 상기 추출된 비트 열을 입력하고, 상기 신경망의 출력 값에 기반하여 공격 여부를 판단하는 단계를 포함할 수 있다.

[0026] 이때, 상기 차량 CAN 통신 공격 탐지 방법은 소정 기록 영역에 기 수집된 공격 유형에 대한 정보가 유지되되, 상기 출력 값에 따라 공격인 것으로 판단되면, 상기 실시간 입력된 CAN 패킷을 이용하여 상기 공격 유형을 갱신되는 단계를 더 포함할 수 있다.

[0027] 또한, 상기 차량 CAN 통신 공격 방법은 상기 실시간 입력된 CAN 패킷이 상기 공격 유형에 해당되는 CAN 패킷인지 여부를 판단하여 공격을 감지하는 단계를 더 포함할 수 있다.

[0028] 또한, 상기 차량 CAN 통신 공격 방법은 상기 비트 단위 표현의 학습 통계 정보에 기반하여 특징 점을 추출하는 단계를 더 포함할 수 있다.

[0029] 여기서, 상기 특징 점을 추출하는 단계는 상기 데이터 필드 내 비트 열 중 적어도 하나의 일부 특정 비트 열을 상기 특징 점으로 선택하는 단계 및 시간에 따른 상기 데이터 필드 내 비트 열의 비트 단위 변화 량에 기반하여 상기 특징 점을 비트 단위로 선택하는 단계 중 적어도 하나를 포함할 수 있다.

[0030] 또한, 상기 특징 점으로 선택된 비트 열 또는 비트가 상기 구성된 신경망에 입력될 수 있다.

[0031] 또한, 상기 일부 특정 비트 열을 상기 특징 점으로 선택되는 경우, 상기 특징 점은 모드 정보를 표현하는 비트 열 및 수치 정보를 표현하는 비트 열 중 적어도 하나를 포함할 수 있다.

[0032] 또한, 상기 비트 단위 변화 량은 연속된 두 개의 CAN 패킷의 데이터 필드 영역을 비트 단위 exclusive-OR 연산

하여 산출될 수 있다.

[0033] 본 발명의 또 다른 일 실시예에 따른 차량 CAN(Controller Area Network) 통신 공격을 탐지하는 방법은 사전 정의된 CAN 패킷이 입력되면, CAN 데이터 프레임을 구성하는 필드 단위로 설정된 템플릿 영역에 매칭되는 상기 사전 정의된 CAN 패킷의 비트 열을 추출하여 비트 단위 표현의 학습 통계 정보를 생성하여 저장하는 학습 단계와 실시간 입력된 CAN 패킷으로부터 상기 템플릿 영역에 매칭되는 비트 열을 추출하고, 상기 템플릿 영역에 대응하여 기 생성된 상기 학습 통계 정보를 이용하여 신경망을 구성하고, 상기 구성된 신경망에 상기 추출된 비트 열을 입력하고, 상기 신경망의 출력 값에 기반하여 공격 여부를 판단하는 검출 단계를 포함할 수 있다.

[0034] 본 발명의 또 다른 일 실시예는 상기한 차량 CAN 통신 공격 탐지 방법들 중 어느 하나의 방법을 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공할 수 있다.

[0035] 상기 본 발명의 양태들은 본 발명의 바람직한 실시예들 중 일부에 불과하며, 본원 발명의 기술적 특징들이 반영된 다양한 실시예들이 당해 기술분야의 통상적인 지식을 가진 자에 의해 이하 상술할 본 발명의 상세한 설명을 기반으로 도출되고 이해될 수 있다.

발명의 효과

[0036] 본 발명에 따른 방법 및 장치에 대한 효과에 대해 설명하면 다음과 같다.

[0037] 본 발명은 차량 내 네트워크 공격 탐지 방법 및 장치를 제공하는 장점이 있다.

[0038] 또한, 본 발명은 사전 학습된 CAN 패킷의 통계적 특성에 기반하여 적응적으로 공격을 탐지하는 방법 및 그를 위한 장치를 제공하는 장점이 있다.

[0039] 또한, 본 발명은 CAN 통신 패킷을 복호하지 않고 이진 데이터의 통계적 특성만을 고려하여 적응적으로 공격을 탐지하므로, 제어기내에서의 공격 탐지를 위한 계산량을 최소화시킬 수 있는 장점이 있다.

[0040] 또한, 본 발명은 입력된 CAN 통신 패킷에 대한 통계적 특성을 비트 단위 학습을 통해 추출하므로 CAN 프레임의 신텍스에 대한 고려가 최소화되어, 학습 과정에서 사용자의 수동 입력을 최소화시킬 수 있는 장점이 있다.

[0041] 본 발명에서 얻을 수 있는 장점 및 효과는 이상에서 언급한 장점 및 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0042] 이하에 첨부되는 도면들은 본 발명에 관한 이해를 돕기 위한 것으로, 상세한 설명과 함께 본 발명에 대한 실시예들을 제공한다. 다만, 본 발명의 기술적 특징이 특정 도면에 한정되는 것은 아니며, 각 도면에서 개시하는 특징들은 서로 조합되어 새로운 실시예로 구성될 수 있다.

도 1은 본 발명에 따른 차량 CAN 통신 네트워크의 구조를 설명하기 위한 도면이다.

도 2는 국제 표준에 정의된 CAN 데이터 프레임의 구조이다.

도 3 내지 4는 본 발명의 일 실시예에 따른 CAN 통신에 사용되는 전송 방식을 설명하기 위한 도면이다.

도 5는 본 발명의 일 실시예에 따른 CAN 공격 탐지 장치의 구조를 설명하기 위한 블록도이다.

도 6은 본 발명의 일 실시예에 따른 CAN 공격 탐지 장치에서의 학습 통계 정보 생성 절차를 설명하기 위한 순서도이다.

도 7은 본 발명의 일 실시예에 따른 CAN 공격 탐지 장치에서의 CAN 공격 감지 방법을 설명하기 위한 순서도이다.

도 8은 본 발명의 다른 일 실시예에 따른 CAN 공격 탐지 장치에서의 CAN 공격 감지 방법을 설명하기 위한 순서도이다.

도 9는 본 발명의 일 실시예에 따른 비트 단위 확률 통계로 구성된 CAN 패킷의 데이터 필드에 대한 특징 점 추출 방법을 설명하기 위한 도면이다.

도 10은 본 발명의 다른 일 실시예에 따른 비트 단위 확률 통계로 구성된 CAN 패킷의 데이터 필드에 대한 특징

점 추출 방법을 설명하기 위한 도면이다.

도 11은 본 발명의 일 실시예에 따른 신경망 학습을 통해 생성된 학습 통계 정보에 기반하여 공격을 탐지하는 절차를 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0043] 이하, 본 발명의 실시예들이 적용되는 장치 및 다양한 방법들에 대하여 도면을 참조하여 보다 상세하게 설명한다. 이하의 설명에서 사용되는 구성요소에 대한 접미사 "모듈" 및 "부"는 명세서 작성의 용이함만이 고려되어 부여되거나 혼용되는 것으로서, 그 자체로 서로 구별되는 의미 또는 역할을 갖는 것은 아니다.

[0044] 이상에서, 본 발명의 실시예를 구성하는 모든 구성 요소들이 하나로 결합되거나 결합되어 동작하는 것으로 설명되었다고 해서, 본 발명이 반드시 이러한 실시예에 한정되는 것은 아니다. 즉, 본 발명의 목적 범위 안에서라면, 그 모든 구성 요소들이 하나 이상으로 선택적으로 결합하여 동작할 수도 있다. 또한, 그 모든 구성 요소들이 각각 하나의 독립적인 하드웨어로 구현될 수 있지만, 각 구성 요소들의 그 일부 또는 전부가 선택적으로 조합되어 하나 또는 복수 개의 하드웨어에서 조합된 일부 또는 전부의 기능을 수행하는 프로그램 모듈을 갖는 컴퓨터 프로그램으로서 구현될 수도 있다. 그 컴퓨터 프로그램을 구성하는 코드들 및 코드 세그먼트들은 본 발명의 기술 분야의 당업자에 의해 용이하게 추론될 수 있을 것이다. 이러한 컴퓨터 프로그램은 컴퓨터가 읽을 수 있는 저장매체(Computer Readable Media)에 저장되어 컴퓨터에 의하여 읽혀지고 실행됨으로써, 본 발명의 실시예를 구현할 수 있다. 컴퓨터 프로그램의 저장매체로서는 자기 기록매체 및 광 기록매체 등이 포함될 수 있다.

[0045] 또한, 이상에서 기재된 "포함하다", "구성하다" 또는 "가지다" 등의 용어는, 특별히 반대되는 기재가 없는 한, 해당 구성 요소가 내재될 수 있음을 의미하는 것이므로, 다른 구성 요소를 제외하는 것이 아니라 다른 구성 요소를 더 포함할 수 있는 것으로 해석되어야 한다. 기술적이거나 과학적인 용어를 포함한 모든 용어들은, 다르게 정의되지 않는 한, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 사전에 정의된 용어와 같이 일반적으로 사용되는 용어들은 관련 기술의 문맥 상의 의미와 일치하는 것으로 해석되어야 하며, 본 발명에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0046] 또한, 본 발명의 구성 요소를 설명하는 데 있어서, 제 1, 제 2, A, B, (a), (b) 등의 용어를 사용할 수 있다. 이러한 용어는 그 구성 요소를 다른 구성 요소와 구별하기 위한 것일 뿐, 그 용어에 의해 해당 구성 요소의 본질이나 차례 또는 순서 등이 한정되지 않는다. 어떤 구성 요소가 다른 구성 요소에 "연결", "결합" 또는 "접속"된다고 기재된 경우, 그 구성 요소는 그 다른 구성 요소에 직접적으로 연결되거나 또는 접속될 수 있지만, 각 구성 요소 사이에 또 다른 구성 요소가 "연결", "결합" 또는 "접속"될 수도 있다고 이해되어야 할 것이다.

[0047] 도 1은 본 발명에 따른 차량 CAN 통신 네트워크의 구조를 설명하기 위한 도면이다.

[0048] 도 1을 참조하면, 본 발명에 적용되는 CAN(Controller Area Network) 통신 시스템은 크게 차량용 게이트웨이(Gateway, 140), 제1 내지 제n 제어기(110), CAN 버스(120), OBD 단말(130), OBD 컨텍터(131), 차량용 게이트웨이(140), 차량 텔레매틱스 단말(150) 중 적어도 하나를 포함하여 구성될 수 있다.

[0049] 일반적으로, CAN 버스상에서는 각각의 노드(ECU)가 CAN 버스에 흘러 다니는 데이터를 읽거나 쓰기 위해 액세스(access)할 때 전체 노드를 제어하는 마스터(master)가 존재하지 않는다. 따라서, 각 노드에서 데이터를 전송할 준비가 되면, 먼저 전송 준비가 되었는지를 버스상에서 확인하고 그 후 CAN 프레임을 CAN 네트워크에 전송한다. 전송되는 CAN 프레임에는 전송 노드 및 수신 노드에 대한 주소 정보가 포함되어 있지 않으며, 대신, 각각의 노드는 CAN 프레임에 포함된 고유한 ID를 통해 프레임을 분류하여 데이터를 수신한다.

[0050] 차량용 게이트웨이(140)는 CAN상에 연결된 제어기(110)들에 대한 인증 절차를 통해 해당 제어기가 안전한 제어기인지 여부를 판단할 수 있다. 또한, 차량용 게이트웨이(140)는 차량 텔레매틱스 단말(150) 및 OBD 단말(130)과 유선 또는 무선으로 연결될 수 있다. 일 예로, 사용자는 OBD 단말(130)을 OBD 컨넥터(131)에 연결한 후 OBD 단말(130)상의 화면을 통해 CAN 주선 버스(122)에 연결된 제어기들(110)의 상태 정보를 확인하거나, CAN 주선 버스(122)를 통해 송수신되는 각종 제어 신호를 모니터링할 수 있다. 또한, 사용자는 OBD 단말(130)을 통해 제어기들(110)에 의해 수집된 차량 상태 정보를 확인할 수도 있다. 이때, 제어기들(110)의 상태 정보, CAN 주선 버스(122)상의 제어 신호, 제어기들(100)에 의해 수집된 차량 상태 정보 등은 차량용 게이트웨이(140)를 통해 OBD 단말(130)에 전송될 수 있다.

[0051] 다른 일 예로, OBD 컨넥터(131)는 CAN 주선 버스(122)에 직접 연결될 수도 있다. 이 경우, OBD 단말(130)은 차량용 게이트웨이(140)를 경유하지 않고, 직접 CAN 주선 버스(122)상에 송수신되는 신호를 모니터링하거나 소정 제어 명령을 통해 제어기들(110)로부터 제어기 상태 정보 및 차량 상태 정보를 획득할 수 있다.

[0052] 또한, 차량용 게이트웨이(140)는 OBD 단말(130)로부터의 소정 제어 신호에 따라 차량에 탑재된 제어기-즉, ECU(Electric Control Unit)-들에 설치된 소프트웨어 버전 정보를 수집하고, 수집된 소프트웨어 버전 정보를 OBD 단말(130)에 전송할 수도 있다. 또한, 차량용 게이트웨이(140)는 OBD 단말(130)의 소정 소프트웨어 업데이트 요청 신호에 따라 해당 제어기를 위한 소프트웨어 파일을 OBD 단말(130)로부터 수신한 후 해당 제어기에 설치할 수도 있다.

[0053] CAN 주선 버스(122)는 Twisted pair wire를 사용하며, 2개의 선은 서로 다른 신호(CAN_HI, CAN_LO)에 의해 구동된다. CAN 주선 버스(122)의 양 종단에는 종단 저항(121)이 구비될 수 있다. CAN 주선 버스(122)상의 전송 속도는 버스의 길이-즉, 주선의 길이-에 따라 달라질 수 있다.

[0054] 제1 내지 제N 제어기(110)는 CAN 주선 버스(122)상의 소정의 CAN 조인트 컨넥터(Joint Connector) 또는 CAN 허브(Hub)-미도시-와 CAN 지선 버스(123)를 통해 연결될 수 있으며, 이론적으로 하나의 CAN 네트워크에 연결될 수 있는 최대 제어기의 개수는 2032이다. 또한, 하나의 CAN 허브에는 복수개의 제어기가 CAN 지선 버스(123)를 통해 연결될 수 있다.

[0055] 이하에서는 도면 번호 110 내지 115를 참조하여 일반적인 CAN 주선 버스에 연결되는 제어기의 구조를 살펴보기로 한다.

[0056] 제어기(110)는 CAN 드라이버(111), CAN 컨트롤러(113), 마이크로컨트롤러(Microcontroller, 115)로 구성될 수 있다.

[0057] CAN 드라이버(111)는 CAN 지선 버스(123) 및 CAN 컨넥터 또는 CAN 허브(미도시)를 통해 CAN 주선 버스(122)와 연결되며, 각각의 제어기의 물리 계층을 구성한다. CAN 드라이버(111)는 CAN 주선 버스(122)의 장애를 감지하고 이를 관리하는 기능 및 메시지의 송수신 기능을 제공할 수 있다.

[0058] CAN 컨트롤러(113)는 CAN 프로토콜 메시지를 송수신하고 수신된 메시지에 대한 메시지 필터링 기능을 수행한다. 또는, CAN 컨트롤러(113)는 재전송 제어를 위한 메시지 버퍼 및 마이크로컨트롤러(115)와의 인터페이스 기능을 제공한다.

[0059] 마이크로컨트롤러(115)는 CPU가 탑재될 수 있으며, 상위 계층 프로토콜을 제공하고 다양한 응용들을 제공할 수 있다.

[0060] 상기한 도 1에는 도시되어 있지는 않으나, 제어기는 우선 순위 정보 및 설치된 소프트웨어 버전 정보, 센싱 정보 등이 기록된 소정 메모리를 포함할 수 있다.

[0061] CAN 주선 버스(122)에 연결된 제어기(110)들은 표준에 정의된 CAN 프레임을 통해 제어 신호 및 데이터를 전송할 수 있다.

[0062] CAN 프레임의 구조는 후술할 도 2를 통해 상세히 설명하기로 한다.

[0063] 도 2는 국제 표준에 정의된 CAN 데이터 프레임의 구조이다.

[0064] 도 2를 참조하면, CAN 데이터 프레임(200)은 SOF(Start Of Frame, 201) 필드, ID(Identifier, 202) 필드, RTR(Remote Transmission Request, 203) 필드, IDE(IDentifier Extension, 204) 필드, R(Reserved, 205) 필드, DLC(Data Length Code, 206) 필드, Data(207) 필드, CRC(Cyclic Redundancy Check, 208) 필드, ACK(ACKnowledgement, 209) 필드, EOF(End Of Frame, 210) 필드 및 IFS(InterFrame Space, 211) 필드를 포함하여 구성될 수 있다.

[0065] SOF(201)는 1비트의 길이를 가지며, 해당 프레임의 시작을 표시하기 위해 사용된다.

[0066] ID(202)는 메시지의 종류를 식별하고 메시지의 우선 순위를 지정하기 위한 정보이다. 본 예에서는 ID(202)의 길이가 11비트인 표준 CAN 데이터 프레임 포맷이 도시되어 있으나, 표준에는 ID(202)의 길이가 29비트인 확장 CAN 데이터 프레임 포맷도 정의되어 있다.

[0067] IDE(204)는 해당 프레임이 표준 프레임인지 확장 프레임인지를 식별하기 위해 사용되며, 1비트의 길이를 가진다. 예를 들어, IDE(203)의 값이 0이면 표준 프레임, 1이면 확장 프레임을 의미할 수 있다.

[0068] RTR(203)은 해당 프레임이 원격 프레임인지 데이터 프레임인지를 구별하기 용도로 사용된다. 예를 들어, RTR(203)의 값이 0이면 데이터 프레임을 의미하고, 1이면 원격 프레임을 의미할 수 있다.

[0069] R(205) 필드는 현재 표준에 용도가 정의되지 않은 추후 사용을 위해 예약된 필드로서 1비트의 길이를 가진다.

[0070] DLC(206)은 해당 프레임에 포함된 데이터의 길이를 바이트 단위로 식별하기 위한 코드 정보로서, 4비트의 길이를 가진다.

[0071] Data(207)는 0바이트(Byte)에서 8바이트까지의 가변 길이를 가질 수 있다.

[0072] CRC(208)는 15비트의 주기적인 중복 체크 코드와 1비트의 역행 델리미터(Delimiter)로 구성되며, 수신된 프레임에 오류가 있는지를 확인하기 위한 용도로 사용된다.

[0073] ACK(209)는 수신단에서 해당 프레임을 정상적으로 수신하였는지를 확인하기 위한 필드로서, 2비트의 길이를 가진다. CAN 데이터 프레임을 정확하게 수신한 모든 CAN 컨트롤러는 해당 프레임의 말미에 위치한 ACK 비트를 전송한다. 전송 노드는 버스 상에 ACK 비트 유무를 확인하고, 만약, ACK가 발견되지 않을 경우, 전송 노드는 해당 프레임에 대한 재전송을 시도한다.

[0074] EOF(210)은 해당 CAN 프레임의 종료를 표시하기 필드로서, 7비트의 길이를 가진다.

[0075] IFS(211)는 CAN 컨트롤러가 연속된 프레임을 처리하기 위해 요구하는 시간을 제공하고, 메시지 버퍼 영역에서 적절한 위치로 정확하게 수신된 프레임을 이동시키는데 필요한 시간을 확보하기 위한 용도로 사용될 수 있다.

[0076] 이상에서 살펴본 바와 같이, CAN 데이터 프레임은 총 47비트에서 111비트의 가변 길이를 갖는다. 만약, Data(207)의 크기가 8바이트인 경우, 전체 CAN 데이터 프레임에서 Data(207)가 차지하는 비율은 58%에 해당된다.

[0077] CAN 통신 메시지는 데이터 프레임, 원격 프레임, 에러 프레임 등 여러 종류의 프레임 형식이 제공된다.

[0078] 도 3 내지 4는 본 발명의 일 실시예에 따른 CAN 통신에 사용되는 전송 방식을 설명하기 위한 도면이다.

[0079] 상세하게, 도 3은 특정 전송 노드에서 일정 주기마다 메시지를 전송하는 주기적 전송 모드에 관한 도면이고, 도 4는 전송할 데이터가 변하는 경우에만 메시지를 전송하는 이벤트 전송 모드에 관한 도면이다.

[0080] 도 3을 참조하면, 전송 노드는 미리 정의된 시간 간격마다 주기적으로 메시지-즉, CAN 데이터 프레임-을 생성하여 CAN 네트워크에 전송한다. 이때, 주기적 전송 모드에서는 전송 메시지에 포함된 데이터의 변화 여부가 체크되지 않을 수 있다.

[0081] 특히, 주기적 전송 모드에서 전송되는 주기적인 CAN 데이터 프레임은 수신 노드 입장에서 송신 노드가 정상 동작하고 있는지를 확인하기 위한 Keep Alive 프레임으로 사용될 수 있다.

[0082] 도 4를 참조하면, 이벤트 전송 모드에서 전송 노드는 전송 대상 데이터를 이전 전송한 데이터와 비교하고, 비교 결과, 서로 상이한 경우에 전송 대상 데이터를 CAN 데이터 프레임으로 구성하여 CAN 버스상에 전송한다.

[0083] 이벤트 전송 모드는 동일 데이터를 불필요하게 전송하지 않으므로 CAN 통신 부하를 줄일 수 있는 장점이 있다.

[0084] 도 5는 본 발명의 일 실시예에 따른 CAN 공격 탐지 장치의 구조를 설명하기 위한 블록도이다.

[0085] 도 5를 참조하면, CAN 공격 탐지 장치(500)는 크게 학습 모듈(510), 프로파일링 모듈(520) 및 검출 모듈(530)을 포함하여 구성될 수 있다.

[0086] 학습 모듈(510)은 사전 정의된 정상 상태의 CAN 패킷과 비정상 상태의 CAN 패킷을 분석하여 비트 단위 표현의 학습 통계 정보를 생성하고, 생성된 학습 통계 정보를 프로파일링 모듈(520)에 전송할 수 있다.

[0087] 일 예로, 학습 모듈(510)은 CAN ID 별 학습 통계 정보를 생성할 수 있다. 다른 일 예로, 학습 모듈(510)은 미리 정의된 템플릿 영역에 포함된 필드의 비트 열에 대한 학습 통계 정보를 생성할 수도 있다. 여기서, 템플릿 영역은 상기 도 2의 CAN 데이터 프레임(200)에서, 데이터(207) 필드를 제외한 필드들 중 적어도 하나의 필드를 조합하여 구성될 수 있으나 이에 한정되지는 않으며, 데이터(207) 필드도 템플릿 영역에 포함될 수 있다. 일 예로, 템플릿 영역은 ID(202) 필드와 DLC(206) 필드로 구성될 수 있으나 이에 한정되지는 않는다.

[0088] 만약, 학습 모듈(510)에서 템플릿 영역에 대한 학습 통계 정보가 생성되지 않는 경우, 후술할 검출 모듈(520)은 템플릿 영역에 포함된 필드 값을 미리 설정된 소정 필드 테이블을 참조하여 해당 값의 유효성만을 확인할 수도

있다. 이 경우, 개별 필드 값에 대한 유효성만이 확인되므로 필드 조합-예를 들면, ID(202) 필드와 DLC(206) 필드의 조합-에 대한 유효성은 확인될 수 없는 단점이 있다. 따라서, 템플릿 영역에 대한 학습 통계 정보를 생성하고, 이를 기반으로 템플릿 영역의 비트 열에 대한 유효성을 확인하는 것이 보다 고도화된 해킹 공격을 탐지할 수 있는 장점이 있다. 반면, 필드 단위 유효성 확인 방법은 학습 모듈(510)의 학습 통계 정보를 생성하기 위한 계산 시간이 현저히 줄어들 수 있는 장점이 있다.

[0089] 본 발명의 일 실시예에 따른 학습 통계 정보는 CAN 데이터 프레임(200)의 데이터(207) 필드에 대해서만 생성될 수도 있으나, 이에 한정되지는 않으며, 본 발명의 또 다른 일 실시예는 상술한 템플릿 영역 및 데이터(207) 필드 영역 모두에 대해서 학습 통계 정보를 생성할 수도 있다.

[0090] 프로파일링 모듈(520)에는 학습 모듈(510)에 의해 생성된 학습 통계 정보 및 기 수집된 공격 유형에 관한 정보가 유지될 수 있다.

[0091] 여기서, 공격 유형은 검출 모듈(530)에 의한 공격 검출 결과에 따라 동적으로 갱신될 수 있다. 일 예로, 공격 유형에 관한 정보는 공격으로 판단된 CAN 패킷 리스트일 수 있으나 이에 한정되지는 않는다.

[0092] 일 예로, 검출 모듈(530)은 실시간 입력된 CAN 패킷이 공격 패킷인 것으로 판단되면, 해당 공격 패킷에 관한 정보를 프로파일링 모듈(520)에 전송할 수 있다.

[0093] 검출 모듈(530)은 CAN 패킷이 입력되면, CAN ID를 추출하고, 추출된 CAN ID에 대응되는 학습 통계 정보를 프로파일링 모듈(520)로부터 획득할 수 있다.

[0094] 연이어, 검출 모듈(530)은 획득된 학습 통계 정보를 이용하여 기계 학습을 위한 신경망(Neural Network)를 구성하고, 입력된 CAN 패킷으로부터 데이터 필드의 비트열을 추출하여 신경망에 입력할 수 있다. 검출 모듈(530)은 신경망의 출력 값에 기반하여 해당 CAN 패킷이 공격 패킷인지 여부를 판단할 수 있다.

[0095] 따라서, 학습 통계 정보에는 신경망 구성에 필요한 각종 파라메터 정보-예를 들면, 후술할 계층 및 노드 간 최적화된 가중치 파라메터(Weighting Parameter), 은닉 계층의 개수 등을 포함할 수 있으나, 이에 한정되지는 않는다.

[0096] 도 6은 본 발명의 일 실시예에 따른 CAN 공격 탐지 장치에서의 학습 통계 정보 생성 절차를 설명하기 위한 순서도이다.

[0097] 도 6을 참조하면, CAN 공격 탐지 장치는 CAN 패킷이 입력되면, 미리 구성된 CAN ID 매핑 테이블을 참조하여 CAN ID를 식별할 수 있다(S610 내지 S620).

[0098] CAN 공격 탐지 장치는 식별된 CAN ID 별 데이터 필드에 대한 기계 학습을 통해 비트 단위 표현의 학습 통계 정보를 생성할 수 있다(S630). 여기서, 기계 학습은 신경망을 포함할 수 있으나, 이에 한정되지는 않는다.

[0099] 연이어, CAN 공격 탐지 장치는 CAN ID 별 생성된 비트 단위 표현의 학습 통계 정보를 저장할 수 있다(S640). 여기서, 학습 통계 정보는 상기 도 5의 프로파일링 모듈(520)에 유지될 수 있다.

[0100] 도 7은 본 발명의 일 실시예에 따른 CAN 공격 탐지 장치에서의 CAN 공격 감지 방법을 설명하기 위한 순서도이다.

[0101] 도 7을 참조하면, CAN 공격 탐지 장치는 CAN 패킷이 입력되면, 입력된 CAN 패킷으로부터 CAN ID 및 데이터 필드의 비트 열을 추출할 수 있다(S710 내지 S720).

[0102] CAN 공격 탐지 장치는 추출된 CAN ID에 대응하여 미리 저장된 비트 단위 표현의 학습 통계 정보를 추출할 수 있다(S730).

[0103] 연이어, CAN 공격 탐지 장치는 추출된 학습 통계 정보를 이용하여 신경망을 구성하고, 추출된 비트 열을 신경망에 입력할 수 있다(S740).

[0104] CAN 공격 탐지 장치는 신경망 출력 값에 기반하여 공격 여부를 판단할 수 있다(S750). 일 예로, 신경망 출력 값은 0 또는 1일 수 있으며, 여기서, 0은 정상 패킷, 1은 공격 패킷을 각각 의미할 수 있다.

[0105] 도 8은 본 발명의 다른 일 실시예에 따른 CAN 공격 탐지 장치에서의 CAN 공격 감지 방법을 설명하기 위한 순서도이다.

[0106] 도 8을 참조하면, CAN 공격 탐지 장치는 소정 템플릿 영역을 설정하고, 설정된 템플릿 영역에 포함된 필드의 비

트 열에 대응되는 비트 단위 표현의 학습 통계 정보를 소정 기록 영역에 저장할 수 있다(S810). 여기서, 기록 영역은 상기 도 5의 프로파일링 모듈(520)일 수 있다. 본 발명의 일 실시예에 따른 CAN 공격 탐지 장치는 복수의 템플릿 영역이 설정될 수 있으며, 설정된 템플릿 영역 별 비트 단위 표현의 학습 통계 정보를 생성하여 프로파일링 모듈(520)에 유지될 수 있다. 일 예로, 제1 템플릿 영역은 ID(202) 필드만으로 구성되고, 제2 템플릿 영역은 ID(202) 필드와 DLC(206) 필드로 구성될 수 있으나, 이에 한정되지는 않으며, CAN 데이터 프레임(200)에 포함된 적어도 하나의 필드를 조합하여 템플릿 영역을 설정할 수 있다.

[0107] CAN 공격 탐지 장치는 CAN 패킷이 입력되면, 설정된 템플릿 영역에 상응하는 필드의 비트 열을 입력된 CAN 패킷으로부터 추출하고, 추출된 비트 열-즉, 해당 템플릿 영역-에 상응하는 학습 통계 정보를 해당 기록 영역으로부터 추출할 수 있다(S820 내지 S830).

[0108] CAN 공격 탐지 장치는 추출된 학습 통계 정보를 이용하여 신경망을 구성하고, 추출된 비트 열을 신경망에 입력할 수 있다(S840).

[0109] 연이어, CAN 공격 탐지 장치는 신경망의 출력 값에 기반하여 해당 CAN 패킷이 공격 패킷인지 여부를 판단할 수 있다(S850).

[0110] 도 9는 본 발명의 일 실시예에 따른 비트 단위 확률 통계로 구성된 CAN 패킷의 데이터 필드에 대한 특징 점 추출 방법을 설명하기 위한 도면이다.

[0111] 입력된 CAN 패킷으로부터 정상 상태/해킹 상태의 통계 특성을 표현하기 위한 특징점을 추출하기 위해, CAN 패킷을 구성하는 비트 단위로 직접 통계 특성을 추출할 수 있다. 패킷을 복호하여 원 디지털 신호를 복원하여 해석한 후 해당 신호로부터 통계 특성을 추출하는 기존 방식 대비 비트열로부터 직접 통계 특성을 추출하는 방식은 복호를 수행하지 않아도 되기 때문에 제어기의 계산량을 감소시킬 수 있는 장점이 있다. 입력된 CAN 패킷의 비트 열로부터 통계 특성을 추출하는 방식의 대표적인 예로 암맹(Blind) 방식이 적용될 수 있다.

[0112] 암맹 방식은 CAN 패킷의 이진 열에서 각 인풋 비트의 정보를 특징점으로 사용할 수 있다. 예를 들어 총 64비트의 CAN 데이터 프레임의 데이터 필드가 입력되는 경우, 암맹 방식은 64차원의 특징점을 활용할 수 있다. 예를 들어, 도 9에 도시된 바와 같이, 8Byte (8*8=64 bits)로 구성된 데이터 필드의 비트 열은 64차원의 특징 점(비트)으로 구성될 수 있다. 이때, 각 특징 점(비트)에 대해 심볼 '1'이 나올 확률 또는 각 특징 점(비트)에 대해 심볼 '1'이 나온 회수에 대한 통계 정보를 생성하고, 각 특징점에 수집된 통계 데이터를 공격 여부에 대한 판단 기준으로 활용할 수 있다. 암맹 방식은 패킷의 신텍스에 독립적으로 적용할 수 있으므로 학습 및 해킹 인지 과정에서 사용자의 수동적인 입력이 없이 수행할 수 있는 장점이 있다.

[0113] 도 10은 본 발명의 다른 일 실시예에 따른 비트 단위 확률 통계로 구성된 CAN 패킷의 데이터 필드에 대한 특징 점 추출 방법을 설명하기 위한 도면이다.

[0114] 상술한 도 9의 암맹 방식과 다르게 BoI(Bit-of-Interest) 방식은 해당 필드에 대한 비트열의 신텍스 정보에 기반하여 특정 비트 열을 선택적으로 취하여 특징 점을 구성하는 방식이다. 일 예로, CAN 데이터 프레임의 데이터 필드에 포함되는 정보는 크게 모드 정보와 수치 정보로 구성될 수 있다. 여기서, 모드 정보는 창문의 개폐 여부, 헤드라이트의 ON/OFF 여부 등과 같은 ECU 혹은 일부 장치의 상태를 표현하는 비트 열에 기록될 수 있다. 반면, 수치 정보는 RPM 속도, 연료량, 바퀴 각도 등과 같이 특정 장치의 상태 값을 표현하는 비트 열에 기록될 수 있다. BoI 방식에서는 모드 정보 및 수치 정보만을 선택하여 학습을 위한 특징점으로 이용할 수 있다. 상기한 도 10에 도시된 바와 같이, BoI 방식의 일예로 데이터 필드의 비트 열 중 9번째부터 24번째까지 비트 열은 모드 정보에 해당하고, 40번째부터 64번째 비트 열은 수치 정보에 해당할 때 모드 정보와 수치 정보에 할당된 비트 열만이 특징 점으로 활용될 수 있다.

[0115] 일반적으로, 모드 정보가 활성화된 경우, 수치 정보는 활성화되지 않고, 반대로, 수치 정보가 활성화된 경우, 모드 정보는 활성화되지 않을 수 있다. 따라서, 해당 필드에서의 특정 비트 열에 대한 학습 통계 정보를 유지하고, 이를 입력된 해당 비트 열과 비교하여 공격 여부를 판단하는 것은 계산 속도 측면 및 메모리 사용 측면에서 효과적일 수 있다.

[0116] 본 발명의 또 다른 일 예로 시간에 따른 변화량에 기반한 특징점 추출 방식이 적용될 수도 있다.

[0117] 일 예로, 입력된 CAN 데이터 프레임의 데이터 필드에 기록된 비트 열의 시간에 따른 변화량은 하기 수식(1)에 의해 산출될 수 있다.

[0118] $$dP(T) = \mathrm{P}(T) \otimes \mathrm{P}(T-1)$$ 수식(1)

[0119] 여기서, $\otimes$ 는 component-wise exclusive OR (CEO) operator로 정의한다. 즉, 시간 T와 T-1에서의 입력된 비트 열을 비트 단위로 exclusive-OR 연산을 수행하여 시변 특징 점을 추출할 수 있다.

[0120] 일 예로, T시점의 비트 열과 T-1 시점의 비트 열의 exclusive-OR 연산을 수행한 결과가 1의 값을 갖는 비트가 특징 점으로 결정될 수 있다.

[0121] 따라서, 시변 특징점에 대해 학습된 통계 정보는 이벤트(단말)성 또는 주기적인 패킷 가로채기에 대한 공격을 효과적으로 검출할 수 있는 장점이 있다.

[0122] 이하에서는 본 발명의 일 실시예에 따른 기계 학습의 방법으로 사용되는 신경망 학습 방법을 후술할 도 11을 참조하여 상세히 설명하기로 한다.

[0123] 도 11은 본 발명의 일 실시예에 따른 신경망 학습을 통해 생성된 학습 통계 정보에 기반하여 공격을 탐지하는 절차를 설명하기 위한 도면이다.

[0124] 도 11을 참조하면, 신경망에서 입력 계층에서 입력된 비트 열은 적어도 하나의 은닉 계층을 거쳐 출력 계층에 전달될 수 있다. 일 예로, 도 11을 참조하면, 은닉 계층(i)의 출력 값들-X0 내지 Xd-이 비선형 결합되어 산출되는 최종 출력 값 y^i는 다음 수식(2)에 의해 산출될 수 있다.

[0125] $$y^i = \sigma(w_{0m} + w_m x^i), m = 1, \ldots, M$$ 수식(2)

[0126] 이 때, $\sigma()$ 는 sigmoid 함수로 활성 함수라고 부르며 출력 값이 0과 1을 갖을 수 있다. w는 각 노드 간 연결 선의 weight parameter이고, M은 각 계층에서의 총 노드의 개수를 의미한다. 각 계층의 출력 값은 다음 계층의 입력 값이 된다.

[0127] 본 실시예에 따른 신경망 학습 방식은 감독 학습 방식이므로 학습 단계에서 사용하는 패킷의 해킹 유무를 미리 알고 있다. 즉, 각 패킷마다 최종 출력 노드의 값(y)은 0 또는 1일 수 있다. 신경망에서 각 계층의 파라메터 w는 최종 노드의 값 오차를 최소화하는 방향으로 학습될 수 있다. 이때, 각 계층의 파라메터 w는 수차례의 반복 학습을 통해 최적 값이 결정될 수 있다. 상기 각 계층의 파라메터 w는 학습 통계 정보에 포함되어 프로파일링 모듈에 유지될 수 있다.

[0128] 본 발명은 본 발명의 정신 및 필수적 특징을 벗어나지 않는 범위에서 다른 특정한 형태로 구체화될 수 있음은 당업자에게 자명하다.

[0129] 따라서, 상기의 상세한 설명은 모든 면에서 제한적으로 해석되어서는 아니되고 예시적인 것으로 고려되어야 한다. 본 발명의 범위는 첨부된 청구항의 합리적 해석에 의해 결정되어야 하고, 본 발명의 등가적 범위 내에서의 모든 변경은 본 발명의 범위에 포함된다.

부호의 설명

[0130] 500: CAN 공격 탐지 장치

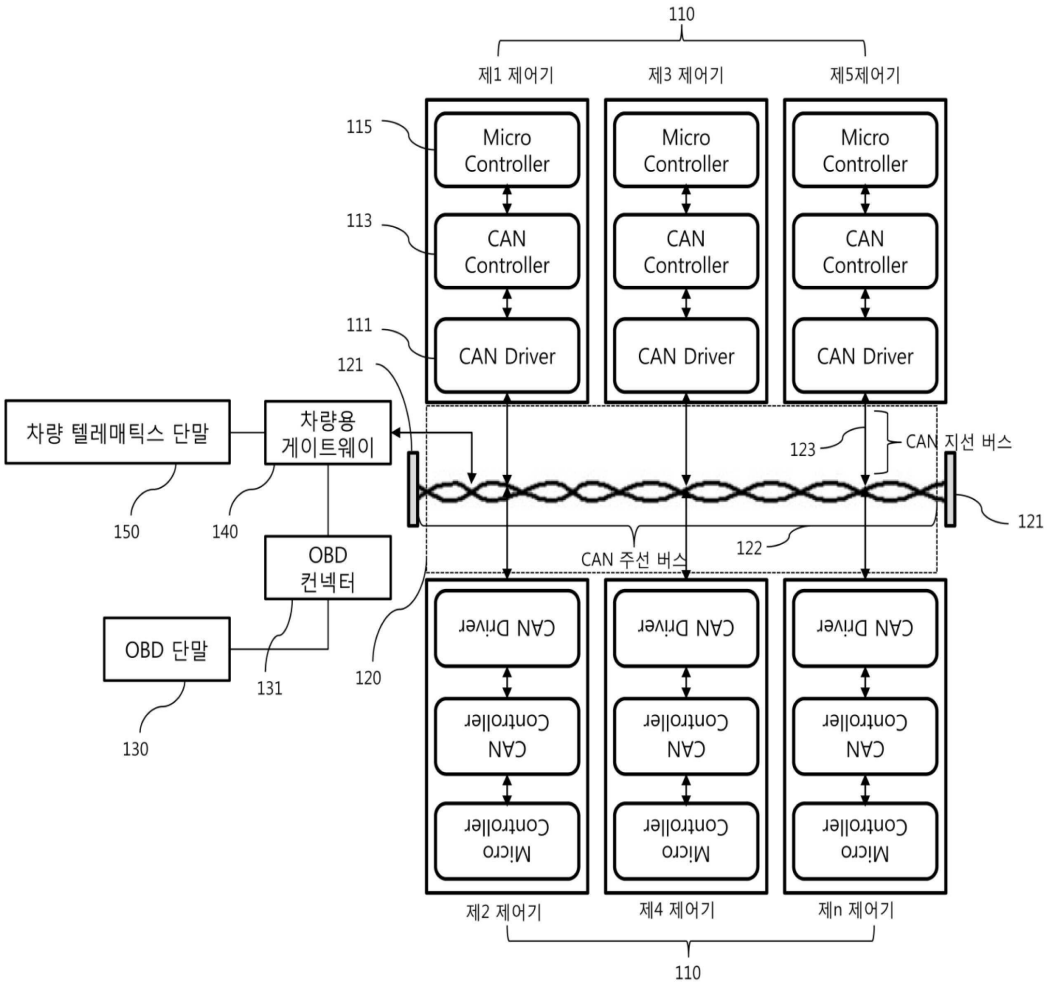
510: 학습 모듈

520: 프로파일링 모듈

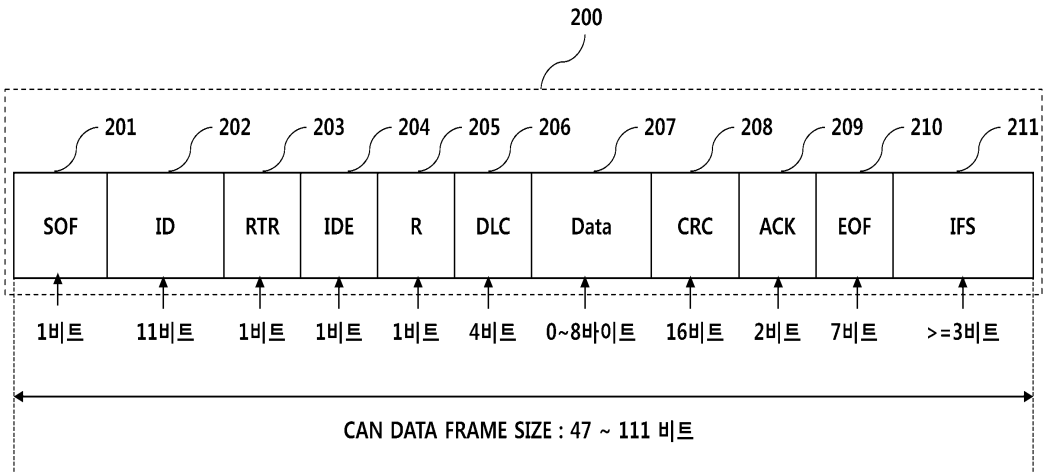
530: 검출 모듈

도면

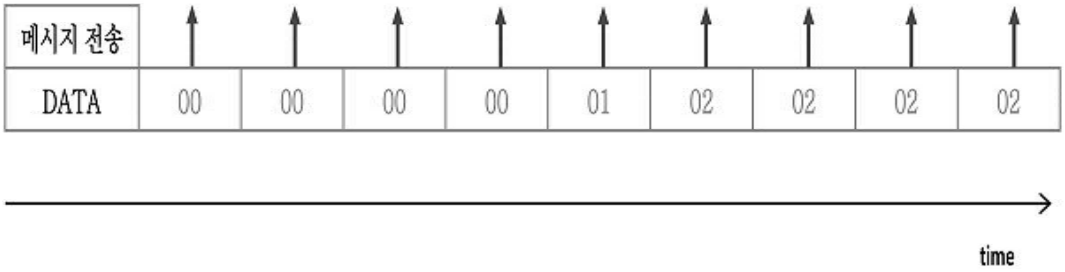
도면1


110
제1 제어기
제3 제어기
제5제어기
115
Micro Controller
113
CAN Controller
111
CAN Driver
121
차량 텔레매틱스 단말
차량용 게이트웨이
123
CAN 지선 버스
150
140
121
OBD 컨넥터
CAN 주선 버스
122
OBD 단말
131
120
130
제2 제어기
제4 제어기
제n 제어기
110

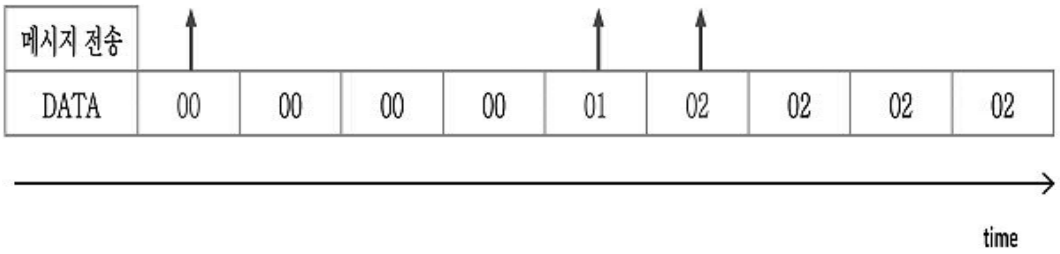
도면2


200
201
202
203
204
205
206
207
208
209
210
211
SOF
ID
RTR
IDE
R
DLC
Data
CRC
ACK
EOF
IFS
1비트
11비트
1비트
1비트
1비트
4비트
0~8바이트
16비트
2비트
7비트
>=3비트
CAN DATA FRAME SIZE : 47 ~ 111 비트

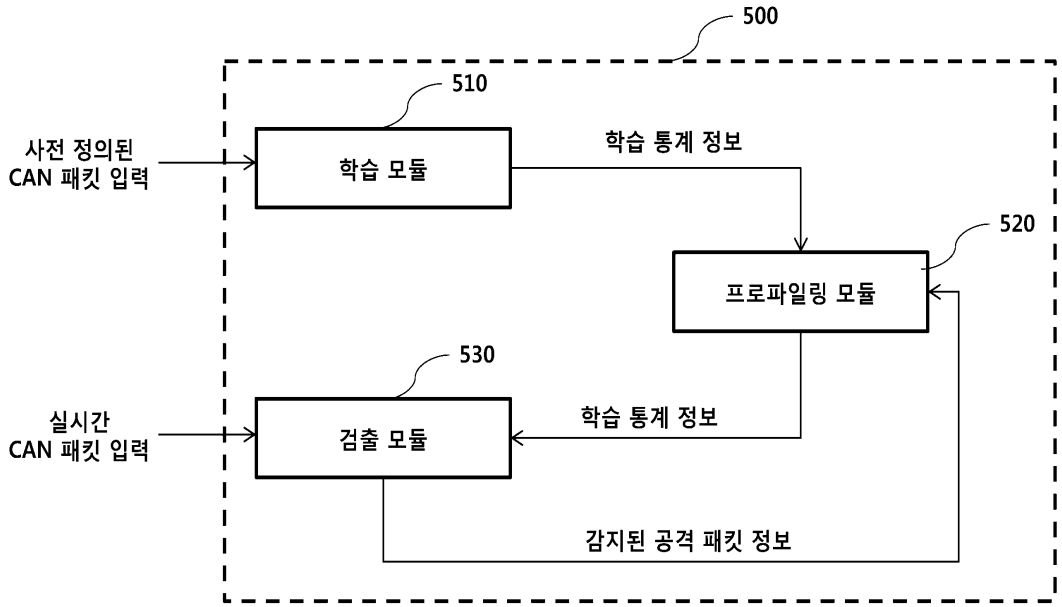
도면3


메시지 전송
DATA
00
00
00
00
01
02
02
02
02
time

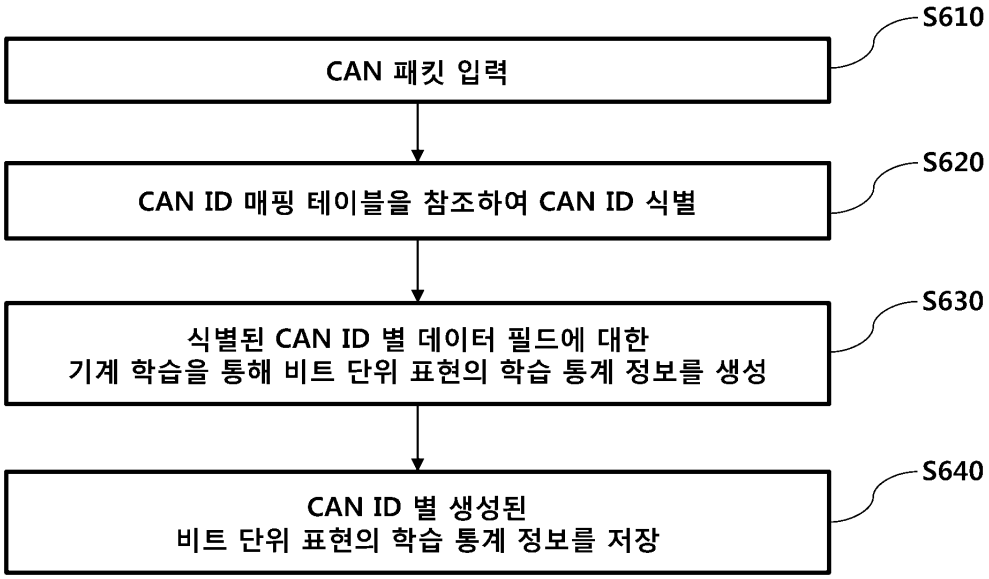
도면4


메시지 전송
DATA
00
00
00
00
01
02
02
02
02
time

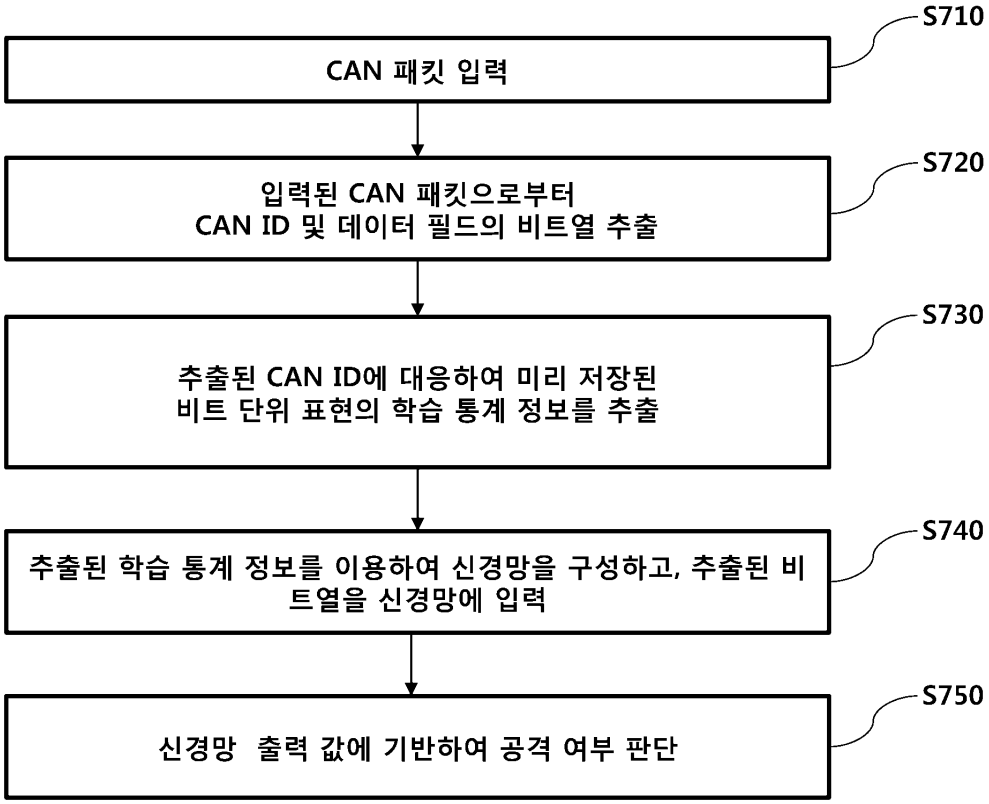
도면5


500
510
사전 정의된
CAN 패킷 입력
학습 모듈
학습 통계 정보
520
프로파일링 모듈
530
실시간
CAN 패킷 입력
검출 모듈
학습 통계 정보
감지된 공격 패킷 정보

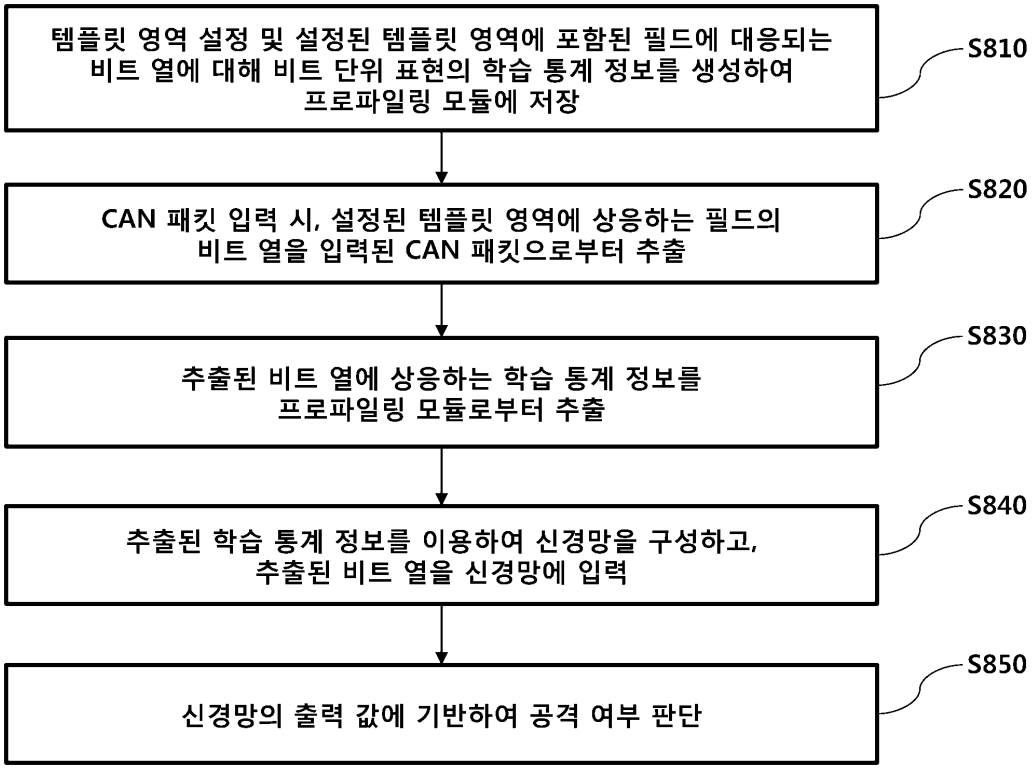
도면6


S610
CAN 패킷 입력
S620
CAN ID 매핑 테이블을 참조하여 CAN ID 식별
S630
식별된 CAN ID 별 데이터 필드에 대한
기계 학습을 통해 비트 단위 표현의 학습 통계 정보를 생성
S640
CAN ID 별 생성된
비트 단위 표현의 학습 통계 정보를 저장

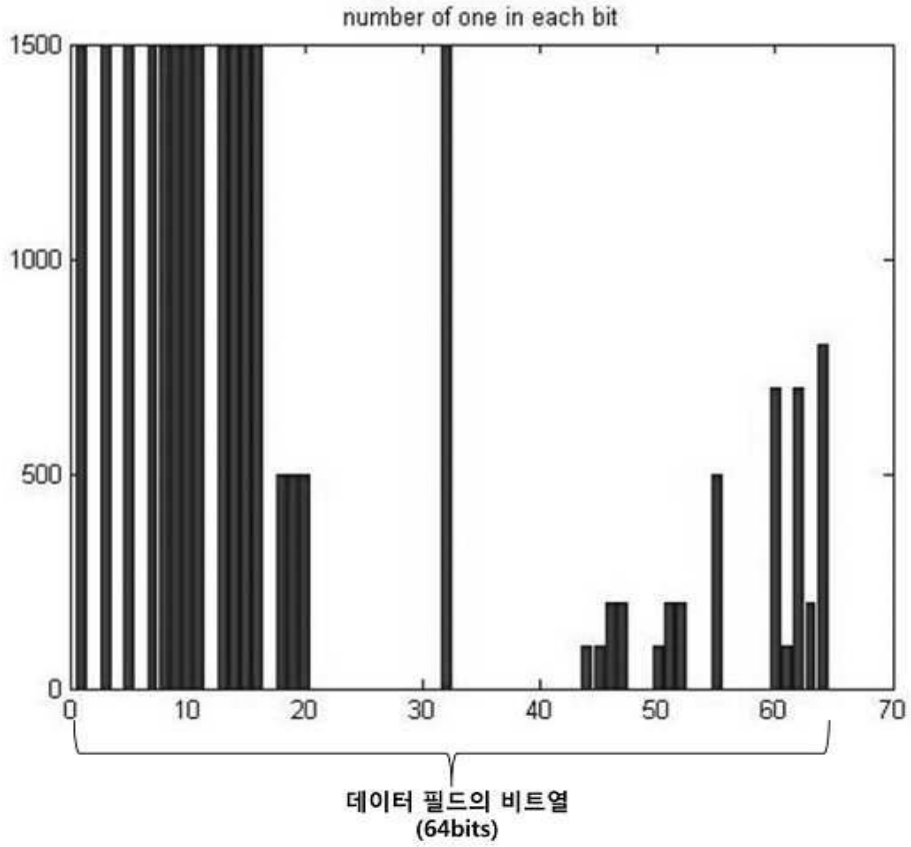
도면7


S710
CAN 패킷 입력
S720
입력된 CAN 패킷으로부터
CAN ID 및 데이터 필드의 비트열 추출
S730
추출된 CAN ID에 대응하여 미리 저장된
비트 단위 표현의 학습 통계 정보를 추출
S740
추출된 학습 통계 정보를 이용하여 신경망을 구성하고, 추출된 비
트열을 신경망에 입력
S750
신경망 출력 값에 기반하여 공격 여부 판단

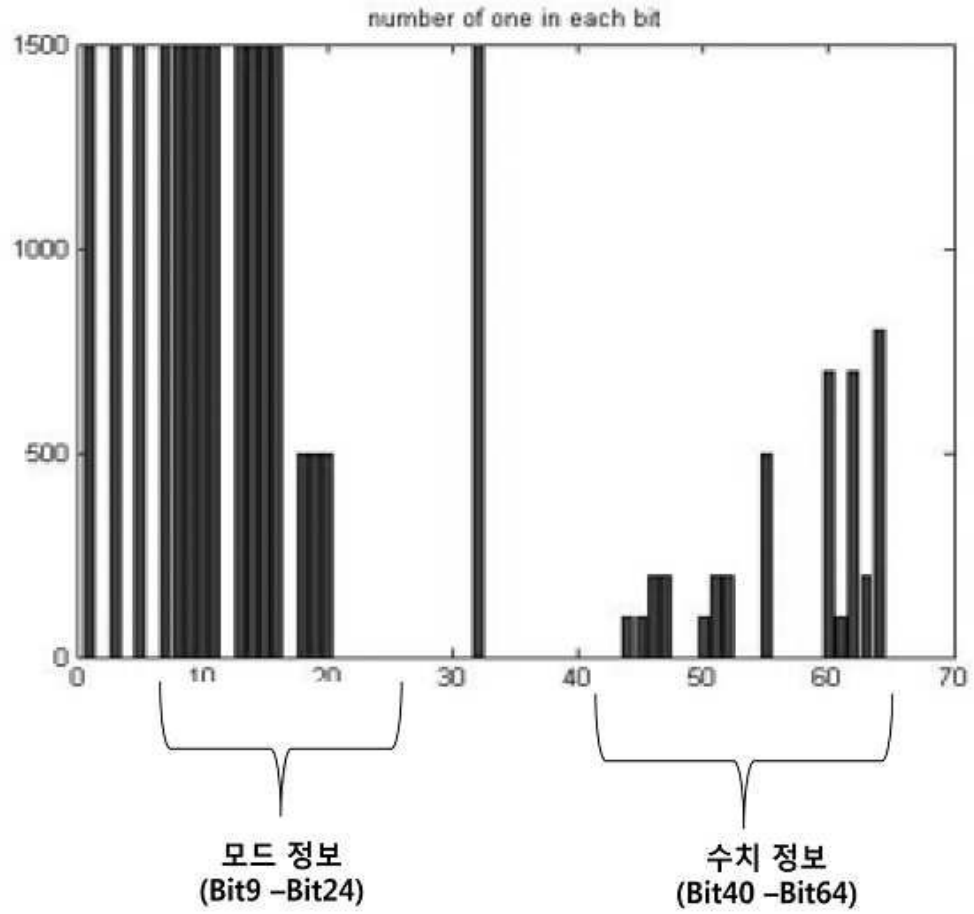
도면8


템플릿 영역 설정 및 설정된 템플릿 영역에 포함된 필드에 대응되는 비트 열에 대해 비트 단위 표현의 학습 통계 정보를 생성하여 프로파일링 모듈에 저장
S810
CAN 패킷 입력 시, 설정된 템플릿 영역에 상응하는 필드의 비트 열을 입력된 CAN 패킷으로부터 추출
S820
추출된 비트 열에 상응하는 학습 통계 정보를 프로파일링 모듈로부터 추출
S830
추출된 학습 통계 정보를 이용하여 신경망을 구성하고, 추출된 비트 열을 신경망에 입력
S840
신경망의 출력 값에 기반하여 공격 여부 판단
S850

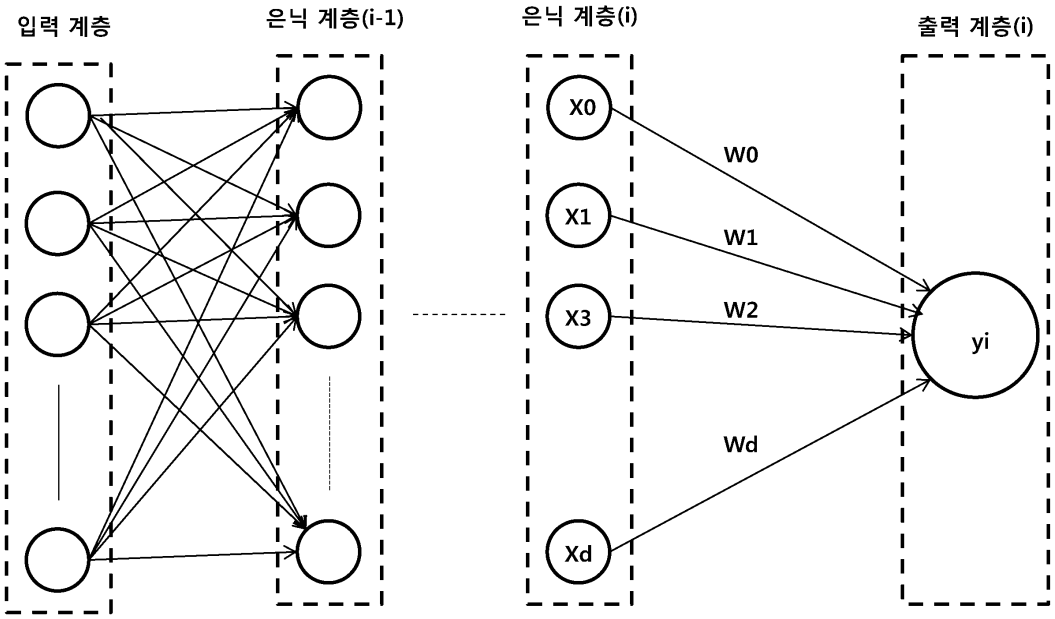
도면9


number of one in each bit
1500
1000
500
0
0
10
20
30
40
50
60
70
데이터 필드의 비트열
(64bits)

도면10


number of one in each bit
1500
1000
500
0
0
10
20
30
40
50
60
70
모드 정보
(Bit9 –Bit24)
수치 정보
(Bit40 –Bit64)

도면11


입력 계층
은닉 계층(i-1)
은닉 계층(i)
출력 계층(i)
X0
X1
X3
Xd
W0
W1
W2
Wd
yi