Deep learning-based malicious smart contract detection scheme for internet of things environment : 논문 리뷰 및 구현

https://youtu.be/pJie6VfwT6w

스마트 컨트랙트

시나리오

모델 학습 과정

실습

서론

- **스마트 컨트랙트**
  - 블록체인 기반의 디지털 계약
  - 서면을 통해 이루어지던 계약을 코드로 구현
  - 두 당사자가 스마트 컨트랙트를 통해 계약을 체결
    → 제 3자인 인증기관 개입 X
  - EVM을 통해 스마트 컨트랙트 실행
    → EVM (Ethereum Virtual Machine) : 이더리움 가상 환경

서론

- **악성 스마트 컨트랙트 종류**
  - 자기 파괴 컨트랙트 (Suicidal contract)

    → 임의의 사용자가 삭제할 수 있는 컨트랙트
  - 방탕한 컨트랙트 (Prodigal contract)

    → 이더를 다른 아무 주소로 보낼 수 있는 컨트랙트
  - 탐욕 컨트랙트 (Greedy contract)

    → 이더를 인출할 수 없도록 무기한으로 잠금이 걸릴 수도 있는 컨트랙트

본론

- **기존 스마트 컨트랙트 흐름**
  - 솔리디티를 통해 스마트 컨트랙트 작성
  - EVM을 통해 컴파일하여 바이트 코드로 변환
  - 블록체인에 해당 컨트랙트 배포

- **제안된 기법**
  - 솔리디티를 통해 스마트 컨트랙트 작성
  - EVM을 통해 컴파일하여 바이트 코드로 변환
  - **바이트 코드를 전처리한 후, 분류기를 통해 탐지**

    → 안전한 스마트 컨트랙트로 분류될 경우 배포

    → 악성 스마트 컨트랙트로 분류될 경우 배포 중단 + 배포자에게 패널티 부과

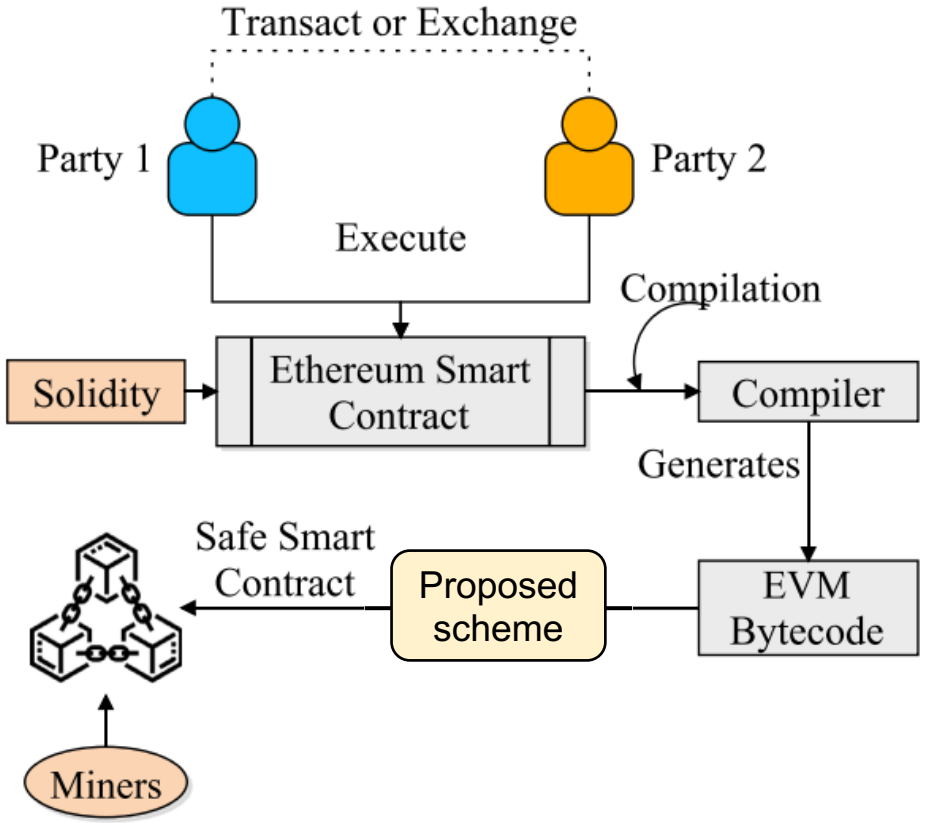


Fig. 1. Traditional flow of smart contract execution over Ethereum [4].

본론

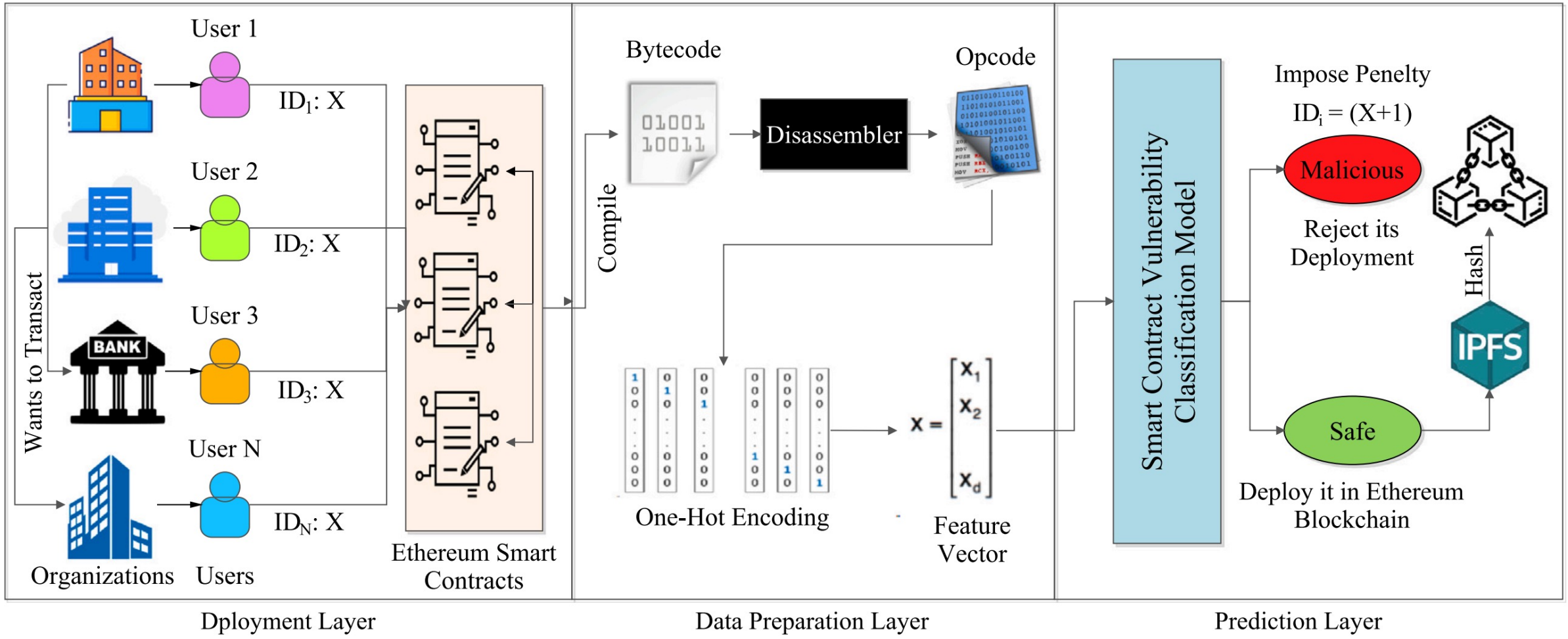
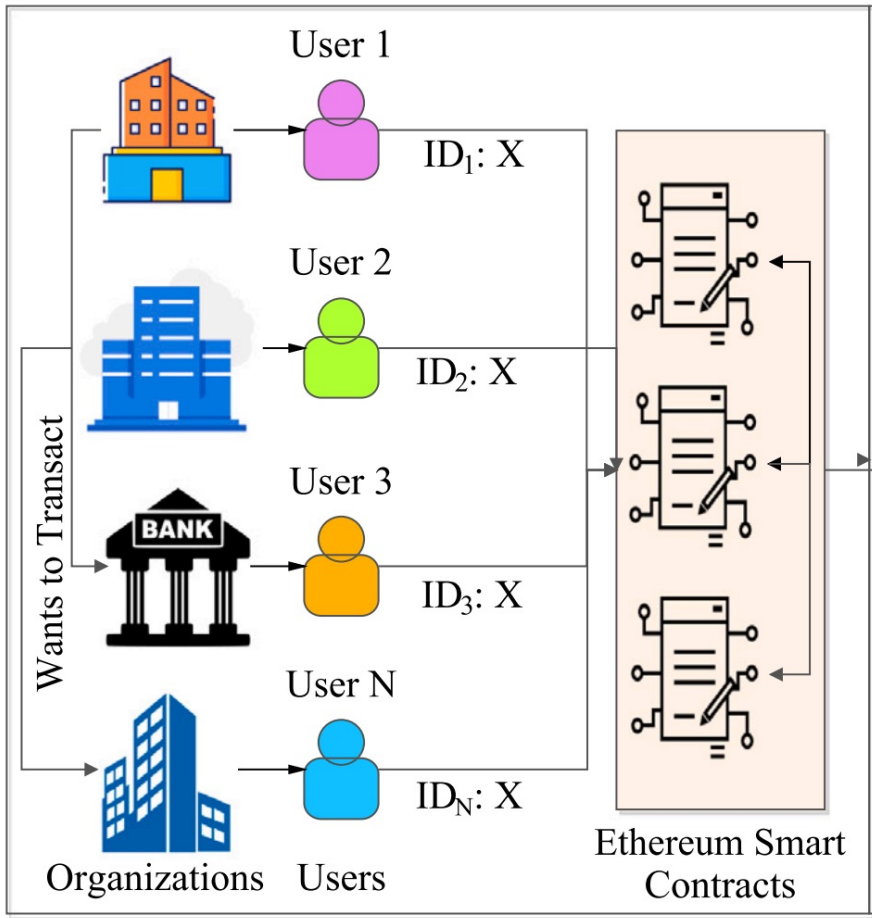


Fig. 3. The proposed malicious smart contract detection scheme.

본론

- **배포 계층**
  - User : 개발자
  - ID : 누가 스마트 컨트랙트를 배포했는지 구분하기 위함
  - X : 악성 스마트 컨트랙트를 작성한 횟수
  - 각 개발자들이 솔리디티, 고 , 자바 언어를 통해 스마트 컨트랙트 작성

$$ID_i : X = \begin{cases} X < 3, & u_i \ can \ deploy \ more \ S \\ X >= 3, & u_i \ suspended \ from \ blockchain \end{cases}$$



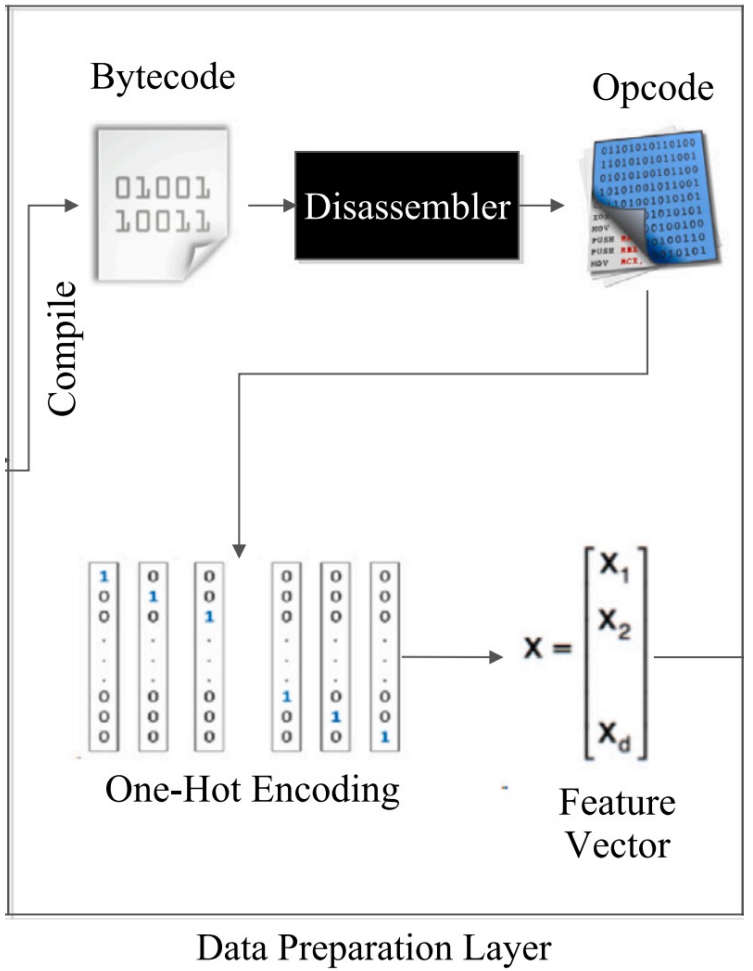
Dployment Layer

본론

- **데이터 전처리**

1. 스마트 컨트랙트를 Bytecode로 컴파일
2. Bytecode → Opcode로 변환
3. Opcode → One-Hot Encoding

- Bytecode on “3x5”: 6003600502
- Opcode: 0x60,0x03,0x60,0x05,0x02
  - 0x60: PUSH1
  - 0x03/0x05: Value
  - 0x02: MUL




```
6080604052348015600f57600080fd5b5060878061001e6000396000f3f
e6080604052348015600f57600080fd5b506004361060285760003560e0
1c8063037a417c14602d575b600080fd5b60336049565b6040518082815
260200191505060405180910390f35b6000600190509056fe84ba30ec42
dadbdeb8edf5cd8b261e89b8d42730ec42080fd592646033600050a0032
```

Fig. 4. Sample bytecode.

```
60 60 52 36 15 61 57 60 35 7c 90 04 63 16 80 63 14 61 57 5b
34 15 61 57 fe 5b 61 5b 61 60 60 90 54 90 61 0a 90 04 73 16
61 56 5b 5b 56 5b 00 5b 34 15 61 57 fe 5b 61 60 80 80 35 73
16 90 60 01 90 91 90 50 50 61 56 5b 00 5b 60 81 90 50 60 60
90 54 90 61 0a 90 04 73 16 73 16 33 73 16 14 80 15 61 57 50
```

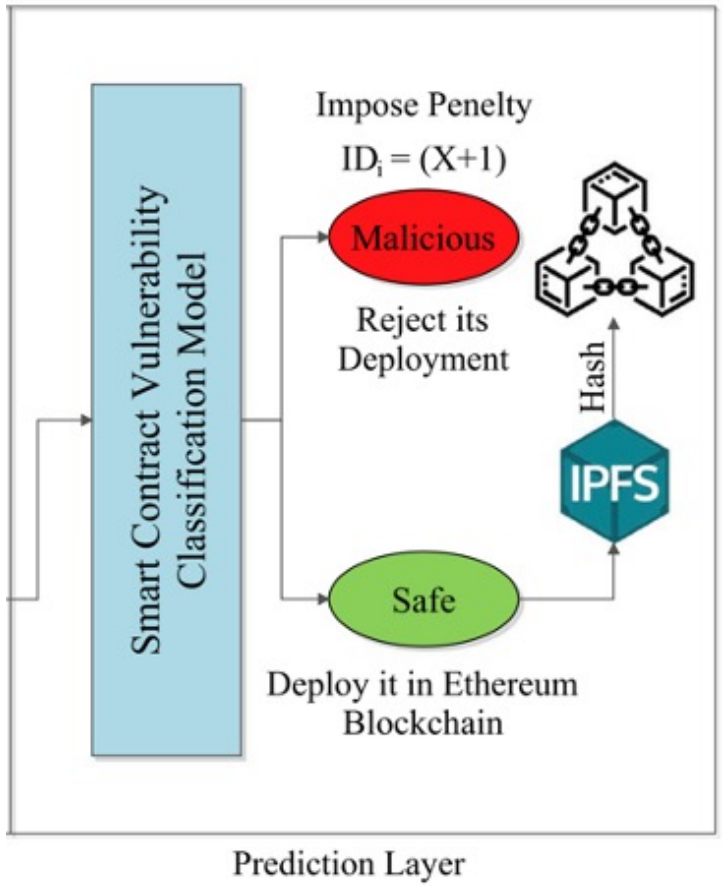
Fig. 5. Sample opcode.

본론

- **악성 컨트랙트 탐지**
  - 전처리과정을 거친 특징 벡터를 모델에 입력하여 탐지

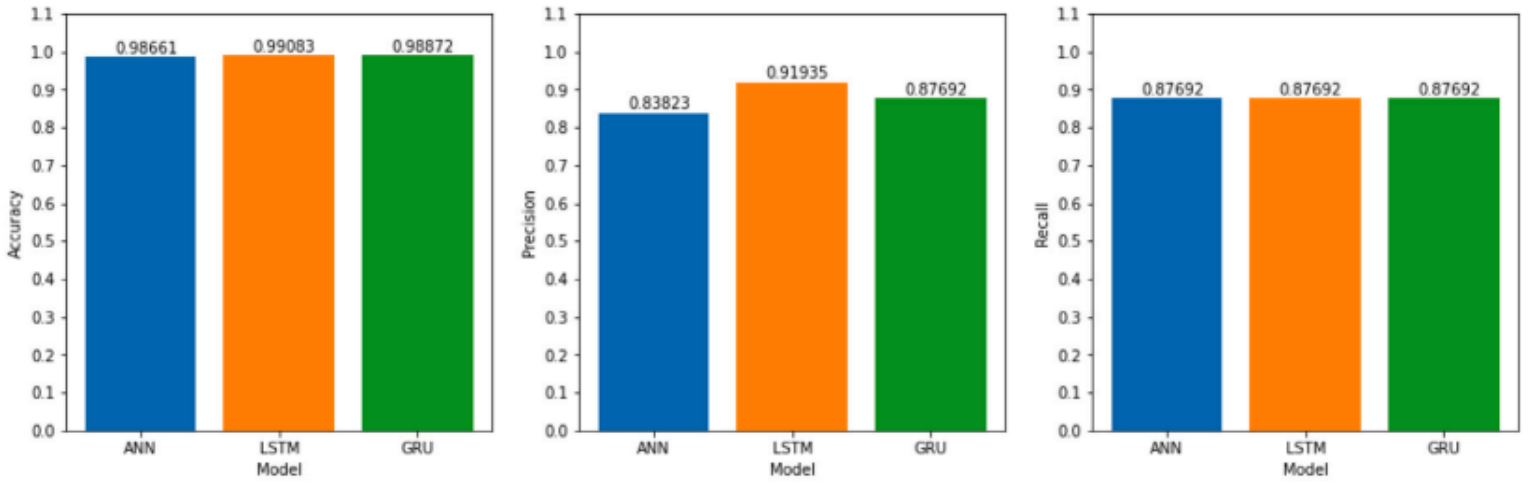
    악성 → 배포 중단 + 패널티

    안전 → IPFS에 배포 후 해시한 후 블록체인에 기록



성능평가

- **데이터셋**
  - 구글 Bigquery으로부터 스마트 컨트랙트 7000개 추출
  - MAIAN 도구를 통해 7000개의 스마트 컨트랙트에 대해 레이블링
- **모델**
  - ANN, LSTM, GRU에 대한 성능



(a) Accuracy (b) Precision (c) Recall

Fig. 11. Comparative analysis of different models based on accuracy, precision and recall.

실습

Q & A