DEOBFUSCATOR

# 문제점 개요서

2025.03.22

컴퓨터융합학부 202002514 안상준
인공지능학과 202202487 박혜연
컴퓨터융합학부 202202602 손예진

# Contents

**논문 별 분석 내용**

- 기존 연구 소개
- 연구 주요 내용 요약
- 연구의 한계점
- 연구 필요성
- 본 연구의 기여

## 기존 연구 소개

# Can LLMs Obfuscate Code? A Systematic Analysis of Large Language Models into Assembly Code Obfuscation

**Seyedreza Mohseni[1*], Seyedali Mohammadi[1*], Deepa Tilwani[2], Yash Saxena[1†], Gerald Ketu Ndawula[1], Sriram Vema[1], Edward Raff[3], Manas Gaur[1]**

[1]University of Maryland, Baltimore County, MD, USA
[2]University of South Carolina, SC, USA
[3]Booz Allen Hamilton, NY, USA
{mohseni1, m294, ysaxena1, geraldn1, sriramv1, edraff1, manas}@umbc.edu, dtilwani@mailbox.sc.edu

저널/컨퍼런스(연도) : AAAI(2025)

## 기존 연구 소개

- LLM의 prompting과 in-context-learning을 위한 **어셈블리 코드 난독화 데이터셋 MAD**를 제공
- MAD는 LLM의 코드 난독화 성능을 테스트하기 위해 만들어진 첫 어셈블리 코드 난독화 데이터셋
- 이 데이터셋을 GPT, CodeLLAMA, CodeGemma 등의 모델로 pre-training 하거나 prompting 함으로써 **데이터셋의 학습가능성과 신뢰성을 입증**

- Control Flow Change, Dead Code Insertion, Register Substitution 으로 세가지 난독화 기법에 대해 테스트. 결과를 Delta Entropy와 Cosine Similarity를 사용해 평가, 사람이 평가하는 부분에서는 20년 이상 경력의 전문가들을 통해 진행
- 결과적으로, **GPT 계열이 다른 Coder 모델보다도 난독화를 잘 수행**했으며, 이 논문에 사용된 모델 중 가장 좋은 성능을 보임.

# 연구의 한계점 및 필요성

## 한계점

**1개 fine-tuning vs. n개 in-context-learning**
**학습 모델의 편향, 모델간 비교의 부정확함**

**모델 평가시의 객관적인 지표의 부족**

## 필요성

다양한 LLM 사용한 실험에
fine-tuning 방법의 학습 모델을 2개 이상 확대

모델의 역난독화 성능을 평가하기 위한
지표에 대한 고민의 필요

## 본 연구의 기여

아직 LLM을 가상화 난독화에 사용한 연구가 없어,
보안 분야에서의 새로운 접근이라고 할 수 있음

---

LLM에 바이너리 코드를 fine-tuning 하는 것과
in-context-learning 만을 하는 것의 차이 비교를 통한
바이너리 코드 분석에 어떤 방법이 적합한지 분석할 수 있음

---

## 기존 연구 소개

# PalmTree: Learning an Assembly Language Model for Instruction Embedding


Xuezixiang Li
University of California Riverside
Riverside, CA 92521, USA
xli287@ucr.edu

Yu Qu
University of California Riverside
Riverside, CA 92521, USA
yuq@ucr.edu

Heng Yin
University of California Riverside
Riverside, CA 92521, USA
heng@cs.ucr.edu


저널/컨퍼런스(연도) : CCS (2021)

## 기존 연구 소개

- **PalmTree는 bert 기반의 모델**
- 일부를 마스킹 하는 **Masked Language Model,** 명령어 간의 문맥을 고려하는 **Context Window Prediction,** 데이터 흐름을 통해 명령어 간 데이터 흐름 관계를 학습하는 **Def-Use Prediction**을 사용해 **사전 학습**

- 내부 평가로는 이상치 탐지, 기본 블록 유사성 검색을, 외부 평가로는 바이너리 코드 유사성 탐지, 함수 원형 추론, 메모리 영역 분석을 활용
- **기존 임베딩 방법에 비해 높은 정확도**를 보임
- 따라서 PalmTree는 **기존 임베딩 방법의 한계를 극복**하고, **바이너리 분석에 유용한 명령어 임베딩 방법을 제안,** 소스 코드와 사전 학습 모델을 공개하여 다른 연구에 사용될 수 있도록 함

# 연구의 한계점 및 필요성

## 한계점

**Bert 기반의 transformer 네트워크의 사용**
**기존의 다른 모델보다 계산량이 많아**
**속도가 느림**

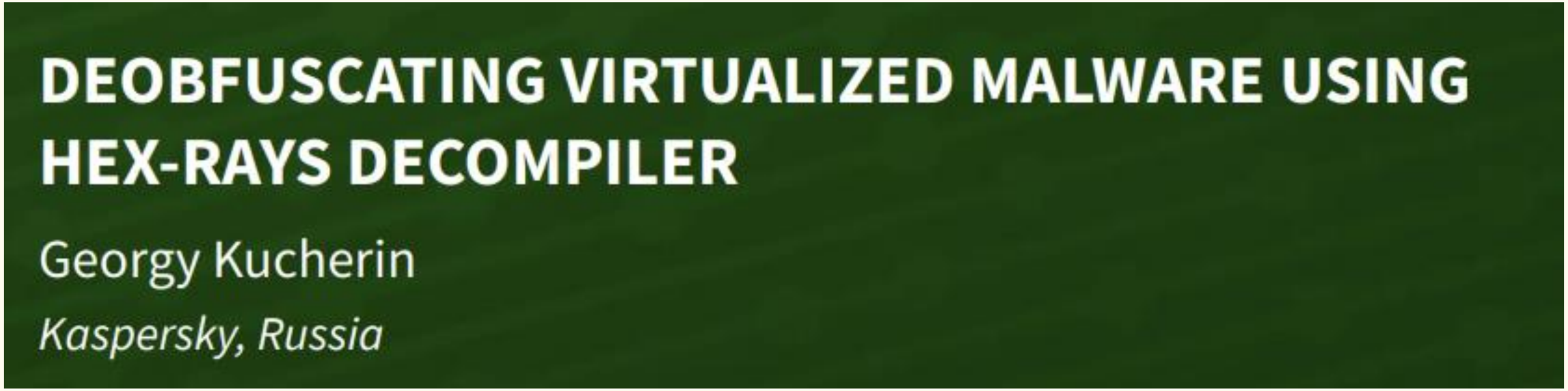
## 필요성

모델의 레이어 수를 줄이거나,
더 효율적인 transformer 아키텍쳐 사용
성능 유지, 계산 비용 감소

## 본 연구의 기여

재 바이너리 코드 벡터화 방법과 바이너리 코드 분석에 LLM을 사용하는
각 연구를 따로 진행하고 있으나,
본 연구에서는 두 가지를 함께 연구하여 바이너리 코드의 특성에 맞게 벡터화한
뒤 LLM을 학습시켜 기존보다 훨씬 좋은 성능을 낼 수 있을 것으로 예측

## 기존 연구 소개




저널/컨퍼런스(연도) : Virus Bulletin Conference(2023)

## 기존 연구 소개

- 독립적인 도구 대신 **IDA Pro, Hex_rays Decompiler**를 사용하여 **가상화된 코드를 역난독화하는 새로운 접근 방식**을 제시

- FinSpy VM을 통한 난독화된 코드에 대하여 IDA SDK의 기능을 사용하여 **역난독화를 자동화**
- 가상화된 코드를 **x86아키텍처로 변환하는 방법**을 설명

- **Hex-Rays 마이크로 코드**를 사용하여 변환된 어셈블리 코드를 C로 디컴파일하여 **역난독화**된 코드를 얻는 방법을 제시

# 연구의 한계점 및 필요성

## 한계점

Hex-Rays Decompiler를 사용한 코드 최적화 수행 시 모든 경우에서 최적화가 효과적이지 않을 수 있음

**다른 가상머신에서 역난독화 기법이 적용될 지 알 수 없음**

## 필요성

다양한 가상머신에 적용할 수 있는 역난독화 방법의 필요

IDA Pro을 사용한 코드 해체, x86 아키텍쳐 변환 후 분석하는 방식 등

## 본 연구의 기여

본 연구는 IDA Pro와 Hex-Rays Dcompiler를 활용하여 가상화된 코드를 역난독화하는
효율적인 접근 방식을 제시,
가상화된 코드의 복잡성을 줄이고 모델이 효율적으로 역난독화 수행

## 기존 연구 소개

# Loki: Hardening Code Obfuscation Against Automated Attacks

Moritz Schloegel, Tim Blazytko, Moritz Contag, Cornelius Aschermann, and Julius Basler, *Ruhr-Universität Bochum;* Thorsten Holz, *CISPA Helmholtz Center for Information Security;* Ali Abbasi, *Ruhr-Universität Bochum*

https://www.usenix.org/conference/usenixsecurity22/presentation/schloegel

저널/컨퍼런스(연도) : USENIX Security Symposium(2022)

## 기존 연구 소개

- LOKI는 기존의 가상 머신 기반 난독화 기법을 강화하여 **자동화된 역난독화 공격에 대응하는 방법**을 제시

- Mixed Boolean-Arithmetic(MBA) 표현식을 사용하여 코드를 난독화
- 기존의 공격 벡터인 Symbolic Execution, 테인트 분석, 프로그램 합성 등에 강력한 보호를 제공

- **프로그램 합성 공격의 성공률**을 **19%**로 줄이는데 성공

# 연구의 한계점 및 필요성

## 한계점

**LOKI의 복잡한 난독화 기법**
**상당한 기술적 지식의 요구**

**특정한 난독화 기법에 최적화된 LOKI**
**다른 유형의 난독화나 플랫폼에 적용하기 어려움**

## 필요성

LOKI의 복잡한 표현식을 최적화하여
오버헤드를 줄이는 연구가 필요

새로운 공격 벡터나 기술이 등장할 경우
대응하기 위한 지속적 연구 필요

## 본 연구의 기여

난독화 된 데이터셋을 통해 LLM을 학습시켜
난독화 된 코드를 역난독화 하는 능력 개선

## 기존 연구 소개

# VMProtect의 역공학 방해 기능 분석 및 Pin을 이용한 우회 방안

박 성 우† · 박 용 수††

저널/컨퍼런스(연도) : KTCCS(2021)

## 기존 연구 소개

- 상용 난독화 도구인 VMProtect3.5.0을 통해 **Debugger Detection, Virtualization Tools Detection을 적용시킨 실행 파일을 Pin Tool을 이용하여 우회하는 방안**을 제시

- VMProtect의 안티리버싱 기법의 위치를 예상하고 이를 바탕으로 알고리즘을 작성하여 **가상화 탐지, 디버거 탐지를 우회**

# 연구의 한계점 및 필요성

## 한계점

제시된 우회 방안이 VMProtect 3.5.0 버전에 특화되어 있음

VMProtect라는 난독화 도구 및 플랫폼에 특화

## 필요성

다양한 환경에서 활용할 수 있도록 확장 연구 필요

Pin을 사용하여 우회 코드를 자동화하 과정에서 성능 손실을 줄이는 연구 필요

제시한 우회 방안이 악의적으로 사용되지 않도록 하는 대책 마련 연구 필요

## 본 연구의 기여

Pin tool을 이용하여 trace를 뽑아 LLM 학습 데이터로 활용 가능
이를 통해 데이터셋의 크기를 줄이고 feature selection이 쉬워짐

---

동적 분석으로 역난독화를 좀 더 정확하게 수행

---

DEOBFUSCATOR

# 감사합니다

---


**컴퓨터융합학부 202002514 안상준**

**인공지능학과　202202487 박혜연**

**컴퓨터융합학부 202202602 손예진**