박찬혁
https://github.com/pbkpch3514

Github: pbkpch3514
Email : pbkpch@naver.com
Mobile : (+82) 010-9084-3768

About Me

KISA, 윈스 등 여러 국가기관 및 기업들과 협업하며, AI 기반 침입탐지 시스템과 문서 검색 연구를 통해
복잡한 문제를 해결해왔습니다.
이러한 경험들을 바탕으로 산업계에 있는 다양한 문제들을 해결해나가고 싶습니다.

Work Experience

- **Korea Institute of Science and Technology Information (KISTI)** Daejeon, South Korea
  Intern *Aug. 2023 - Aug. 2023*
  - DL/ML 기반 침입 탐지 모델 전처리 업무 담당
    전처리 개선으로 기존 모델의 테스트 f1-score에서 10%p 상승 효과
    LIME 알고리즘을 추가하여 전처리의 효과를 시각화함
  - Locality-Sensitive Hashsing(LSH) 기반 패킷 중복 제거 알고리즘 개발
    Cosine LSH 를 통하여 빠르게 비슷한 데이터를 찾는데 성공
    2048차원 벡터 230만개로 이루어진 데이터셋을 10분만에 56만개로 중복제거 하는데 성공함

Projects

Near Neighbor Searching

- **Remote Jaccard Filter (2024.08 - present)**
  원격 서버의 파일들을 요약한 경량 비트맵을 통해 직접 쿼리하지 않고도 비슷한 파일을 탐지할 수 있는 알고리즘 개발
  - 논문 제 1저자 및 프로젝트 리더
    Key achievements:
    1. 10만개의 파일 데이터셋을 단 64MB 비트맵에 요약 가능 (원본 파일 크기 무관)
    2. 10만개 파일 기준 쿼리 소요 시간 단 20초
    3. INFOCOM'25에 논문 심사중
- **Maximum Jaccard-Similarity Estimator (2023.06 - present)**
  파일 데이터셋과 쿼리 데이터 사이의 자카드 인덱스의 최대값을 빠르게 추정하는 비트맵 알고리즘 개발
  - 논문 제 2저자
    Key achievements:
    1. 최신 기술의 100배 이상의 메모리를 절약하여 문제 해결
    2. 원본 파일 크기와 무관하게 파일 당 약 80Bytes 정도로 요약 가능
    3. ICDE 2025에 논문 심사중
- **Source Code Plagiarism Detection Project (2022.03 - 2022.06)**
  학부 내 코딩 과제에서 다른 사람의 코드들 베껴오는 사람을 찾아내는 프로젝트
  - 프로젝트 메인 개발자
    Key achievements:
    1. Lexer기반 전처리로 변수 명과 주석으로 우회하는 경우를 막아냄
    2. 실제 학부 수업에서도 사용되어 표절 의심 학생들을 잡아냄
    3. MinHash+ElasticSearch 구조로 검색 속도가 빠르고 확장성이 좋음
- **Packet Near-Deduplicator (2023.06 - 2023.12)**
  대용량 패킷 데이터셋에서 내용이 비슷한 패킷을 제거하는 알고리즘

AI-based Intrusion Detection System

- **SHAP based Attack Classification Model Explainer (2023.06 - 2023.12)**
  SHAP 알고리즘을 통하여 머신러닝 기반 공격 분류 모델이 예측한 근거를 설명하기

- 알고리즘 구현
  Key achievements:
  1. 한국 인터넷 진흥원(KISA)과 (주)윈스와 진행한 프로젝트
  2. SHAP 알고리즘을 통하여 모델의 예측 경향성을 보안 전문가한테 설명할 수 있게 됨
  3. 보안 전문가는 SHAP 알고리즘의 설명을 기반으로 모델을 재가공 할 수 있게 됨

- **ASAP (Aggregate and Search for Alerts in Packets) (2022.02 - 2022.06)**
  보안관제 전문가의 과도한 업무 분석량을 감소시키는 인공지능 기반 패킷 분석 자동화 시스템
  - 프론트 엔드 개발 및 검색 엔진 구축 보조
    Key achievements:
    1. 보안관제 요원이 분석할 패킷의 양을 약 98% 감소시킴
    2. (주)윈스와 협업하여 진행한 산학협력 프로젝트
    3. MinHash+ElasticSearch 검색엔진과 결합하여 빠른 속도의 패킷 검색 가능.

- **Filtering and Machine Learning for Malware Detection in Edge Computing (2021.07 - 2021.12)**
  딥러닝 기반 악성코드 분류 semi-automization
  - 논문 제 2저자
    Key achievements:
    1. Sensors 2022에 논문 등재 (JCR 상위 25% 이내)
    2. low confidence를 가진 예측값을 unpredictable로 분류함으로써
       모델의 신뢰성을 높임.

Skills

- **Language** - Python, C++, Java(Basic), JavaScript(Basic)
- **Database** - MySQL, ElasticSearch
- **Web** - Flask, FastAPI, Django(Basic)
- **Deep Learning/Machine Learning** - Tensorflow, Keras, Pytorch
- **ETC** - Git, Github, Windows, Linux, Jupyter

Education

- **Kookmin University** — Seoul, South Korea
  Department of Computer Engineering, Master — *Mar. 2023 - Feb. 2025*
  - GPA : 4.19 / 4.5
- **Kookmin University** — Seoul, South Korea
  Department of Software, Bachelor — *Mar. 2019 - Feb. 2023*
  - GPA : 4.04 / 4.5

Research Experiences

- **Information Security Lab, Kookmin Universitsy** — Seoul, South Korea
  Graduate Student Reasearcher (Advisor: MyungKeun Yoon) — *Mar. 2023 - Feb. 2025*
- **Information Security Lab, Kookmin Universitsy** — Seoul, South Korea
  Undergraduate Research Intern (Advisor: MyungKeun Yoon) — *Jul. 2021 - Feb. 2023*

Papers

- **Remote Jaccard Filter for Finding Similar Files with Minimal False Negatives**
  ChanHyeok Park, JungHyeok Im, and MyungKeun Yoon, IEEE INFOCOM 2025 (under review)
- **Maximum Jaccard-Similarity Estimator**
  HyeongBin Seo, ChanHyeok Park, and MyungKeun Yoon, IEEE ICDE 2025 (under review)
- **FILM : Filtering and Machine Learning for Malware Detection in Edge Computing**
  YeongJae Kim, ChanHyeok Park, and MyungKeun Yoon, MDPI Sensors 2022, no.6:2150

- **원격주소 프로파일링과 딥러닝을 이용한 네트워크 이상탐지 연구**
  Haeun Jeon, MinSong Kim, ChanHyeok Park, and MyungKeun Yoon, 2022 제6회 금융보안원 논문공모전 대상
- **MaaD: 머신러닝을 이용한 보안데이터 라벨 디버거**
  HyeonGy Shon, SeokGyu Hong, HyeongBin Seo, and ChanHyeok Park, 2021 AI+Security 우수논문 아이디어 공모전 최우수상
- **보안관제 인공지능 모델의 신뢰성 향상 연구**
  HyeonGy Shon, SeokGyu Hong, HyeongBin Seo, and ChanHyeok Park, 2021 제5회 금융보안원 논문공모전 장려상

Awards

- **제6회 금융보안원 논문공모전 (대상)** Nov. 2022
  원격주소 프로파일링과 딥러닝을 이용한 네트워크 이상탐지 연구
- **AI+Security 우수논문 · 아이디어 공모전 (최우수상, 과학기술정보통신부 장관상)** Dec. 2021
  논문 제목 : MaaD: 머신러닝을 이용한 보안데이터 라벨 디버거
- **제5회 금융보안원 논문공모전 장려상** Oct. 2021
  논문 제목 : 보안관제 인공지능 모델의 신뢰성 향상 연구