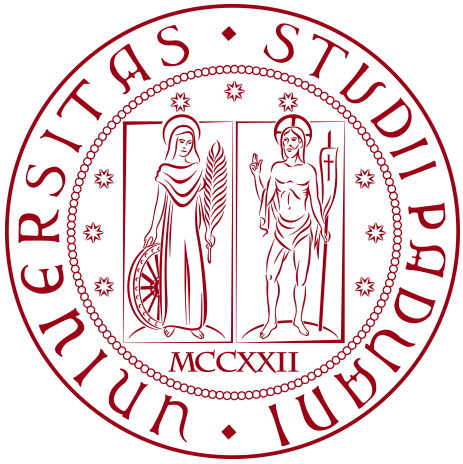
Università degli Studi di Padova

Dipartimento di Matematica "Tullio Levi-Civita"

Corso di Laurea in Informatica



Decoding GAN-Generated Malware using Explainable AI Techniques

Tesi di Laurea Triennale

Relatore

Prof. Galeazzi Alessandro

Laureando

Tiozzo Matteo

Matricola 2042882

Anno Accademico 2023-2024

"Colui il quale ha inseguito e sconfitto i demoni Sem, che ora vagano per il mondo, domandandosi: «ma nü, chi sëm?»"

— Il grande Pdor, figlio di Kmer, della tribù di Ishtar, della terra desolata dei Kfnir, uno degli ultimi sette saggi: Pfulur, Galér, Astaparigna, Sùsar, Param, Fusus e Tarìm.

Ringraziamenti

Desidero esprimere la mia gratitudine al professor Galeazzi Alessandro, mio relatore, per l'aiuto e il sostegno che mi ha dato durante la stesura dell'elaborato.

Vorrei anche ringraziare, con affetto, i miei genitori per il loro sostegno, il grande aiuto e la loro presenza in ogni momento durante gli anni di studio.

Desidero poi ringraziare i miei amici per i bellissimi anni trascorsi insieme e le mille avventure vissute.

Padova, Dicembre 2024 *Tiozzo Matteo*

Sommario

Il presente lavoro di tesi si inserisce nell'ambito dell'analisi e rilevazione di malware, con un particolare focus sull'utilizzo delle Reti Generative Avversarie (GAN) come strumento per migliorare la capacità di rilevamento e comprensione di queste minacce. Durante lo stage curricolare, sono stati affrontati diversi aspetti chiave della ricerca, a partire dalla revisione della letteratura esistente sulle tecniche che combinano le GAN con l'analisi del malware. L'obiettivo primario è stato quello di progettare e sviluppare un sistema di rilevamento del malware basato su tecniche di deep learning, come le Convolutional Neural Networks (CNN), supportate da reti avanzate quali InceptionNet e XceptionNet.

Per addestrare e valutare i modelli, è stato raccolto un dataset curato di eseguibili malware da fonti affidabili come MalwareBazaar e VirusShare, etichettato tramite servizi come VirusTotal e AVClass2. I binari di malware sono stati successivamente convertiti in un formato idoneo per essere utilizzati come input per le GAN. Il modello di rete generativa avversaria sviluppato, ispirato a architetture come DCGAN e WGAN, è stato monitorato tramite metriche chiave come la Fréchet Inception Distance (FID) e la qualità visiva dei campioni generati.

Un ulteriore focus del progetto è stato l'impiego di tecniche di Explainable AI (XAI), come Grad-CAM e Lime, per migliorare la trasparenza e l'interpretabilità del modello di rilevamento. Le analisi effettuate includono una valutazione quantitativa dell'interpretabilità dei modelli e l'esecuzione di un'analisi ablation per determinare l'importanza delle caratteristiche chiave.

Il progetto fornisce contributi significativi sia in termini di innovazione nel campo della rilevazione del malware, sia nell'applicazione delle GAN come strumento per la generazione di malware sintetico e la valutazione della loro efficacia. I risultati ottenuti verranno discussi con i ricercatori coinvolti e saranno oggetto di approfondimento nella relazione finale.

Indice

Elenco delle figure

Elenco delle tabelle

Elenco dei codici sorgenti

Capitolo 1

Introduzione

Capitolo 2

Processi e metodologie

2.1 Processo sviluppo prodotto

Capitolo 3

Descrizione dello stage

3.1 Introduzione al progetto

3.2 Analisi preventiva dei rischi

3.3 Requisiti e obiettivi

3.4 Pianificazione

3.4.1 subsection

3.4.1.1 subsubsection

3.4.1.1.1 paragraph

Capitolo 4

Analisi dei requisiti

Capitolo 5

Progettazione e codifica

5.1 Tecnologie e strumenti

5.2 Ciclo di vita del software

5.3 Progettazione

5.3.1 Namespace 1

5.4 Design Pattern utilizzati

Capitolo 6

Verifica e validazione

Capitolo 7

Conclusioni

7.1 Consuntivo finale

7.2 Raggiungimento degli obiettivi

7.3 Conoscenze acquisite

7.4 Valutazione personale

Bibliografia

Sitografia