Aplikovaná AI pomáhá při odhalování kybernetických hrozeb

Filip Marvan

Obsah

- Aktuální hrozby a trendy, doporučení
- Jak AI pomáhá útočníkům
- Jak AI pomáhá se bránit
- Ukázka vyšetřování kybernetického útoku

Aktuální hrozby

Nové Ransomware kampaně, profesionalizace, nové cíle útoků

Gameradon

2020 - present

- Targets government organizations in Ukraine
- Is delivered by spear-phishing emails, typically by attaching Word files that trigger remote template injection and execute a malicious macro
- Is possibly a general intelligence operation against Ukraine

Earth Longhzhi

December 2022 - March 2023

- Targets government, healthcare, technology, and manufacturing organizations
- Abuses a Windows Defender executable to perform DLL sideloading
- Disables security products by "stack rumbling" via Image File Execution Options

NOBELIUM's Spear-Phishing Attack

January 2022 - present

- Targets diplomatic agencies
- Sent spear-phishing emails to diplomats and impersonated embassies in European countries
- Attached a PDF file or had an email body that contained an embedded link to the next stage payload source
- Possibly aimed at finding information on the diplomatic policies of each target country

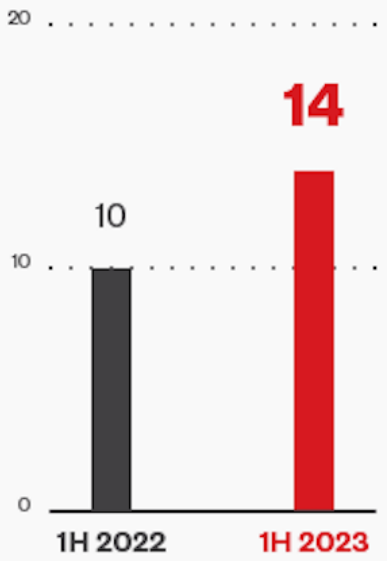
SharpPanda

May 28, 2023

- Targets government agencies in Europe, United States, and Asia
- Is distributed by a decoy file that is a malware executable via the RoyalRoad exploit
- Is possibly a general intelligence operation against the Central Asian region

New Ransomware Families

found in the first half of 2023



Newly detected ransomware families

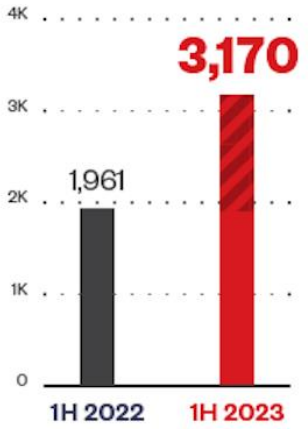
Mimic ransomware

- Targets Russian- and English-speaking users.
- Abuses APIs of Everything, a Windows file name search engine, to query target files to encrypt.
- Has code similar to the 2022 Conti ransomware.

DarkBit ransomware

- Targets educational institutions in Israel.
- Is written in Go programming language, which simplifies the process of supporting various operating systems.
- Employs AES-256 encryption, which can affect a wide range of file types.
- Accepts command-line arguments and can be run autonomously.

Most Increased Detection by Operating System



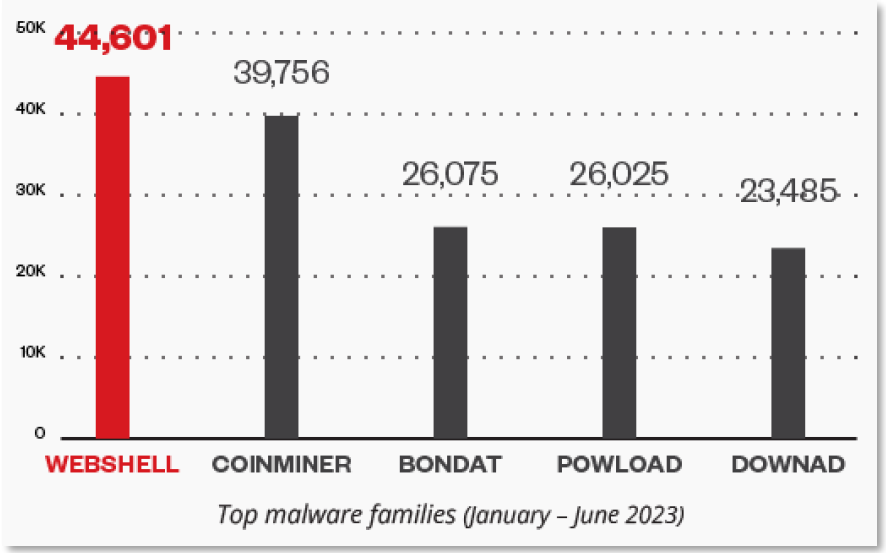
Ransomware detections on Linux

Linux variants remain a staple for new ransomware families with over 62% more ransomware detections in the first half of 2023 compared to the first half of 2022.

Ransomware families that target Linux include RTM Locker, BlackSuit, Akira, and Trigona

Aktuální trendy

- Roste počet útoků na Linux, IOT, NAS, routery, hypervizory
- Automatizace útoků, především fáze Initial access a lateral movement
- Státem podporované kybernetické útoky maskující se za Ransomware
- Profesionalizace, vyhlášen bug bounty na ransomware platformu (LockBit)
- Nejčastějším typem malware je webshell



Top malware families (January – June 2023)

Zajímavé útoky a techniky z roku 2023

- Útoky na Zero-Day zranitelnosti GoAnywhere, 3CX, PaperCut a MOVEit (Využívají skupiny BlackBasta a Nokiyawa)
- AI pomáhá útočníkům s cílením útoků a přípravou scam kampaní
- Využití programovacího jazyka NIM pomáhá útočníkům vyhnout se detekci
- Útoky Bring-your-own-vulnerable-driver (BYOVD) skupiny AuKill, SpyBot a BlackCat

Hrozby které přináší AI/ML

Zajímavé směry zneužití ML technologie, nejzajímavější PoC

Příklady z minulosti zneužívající ML

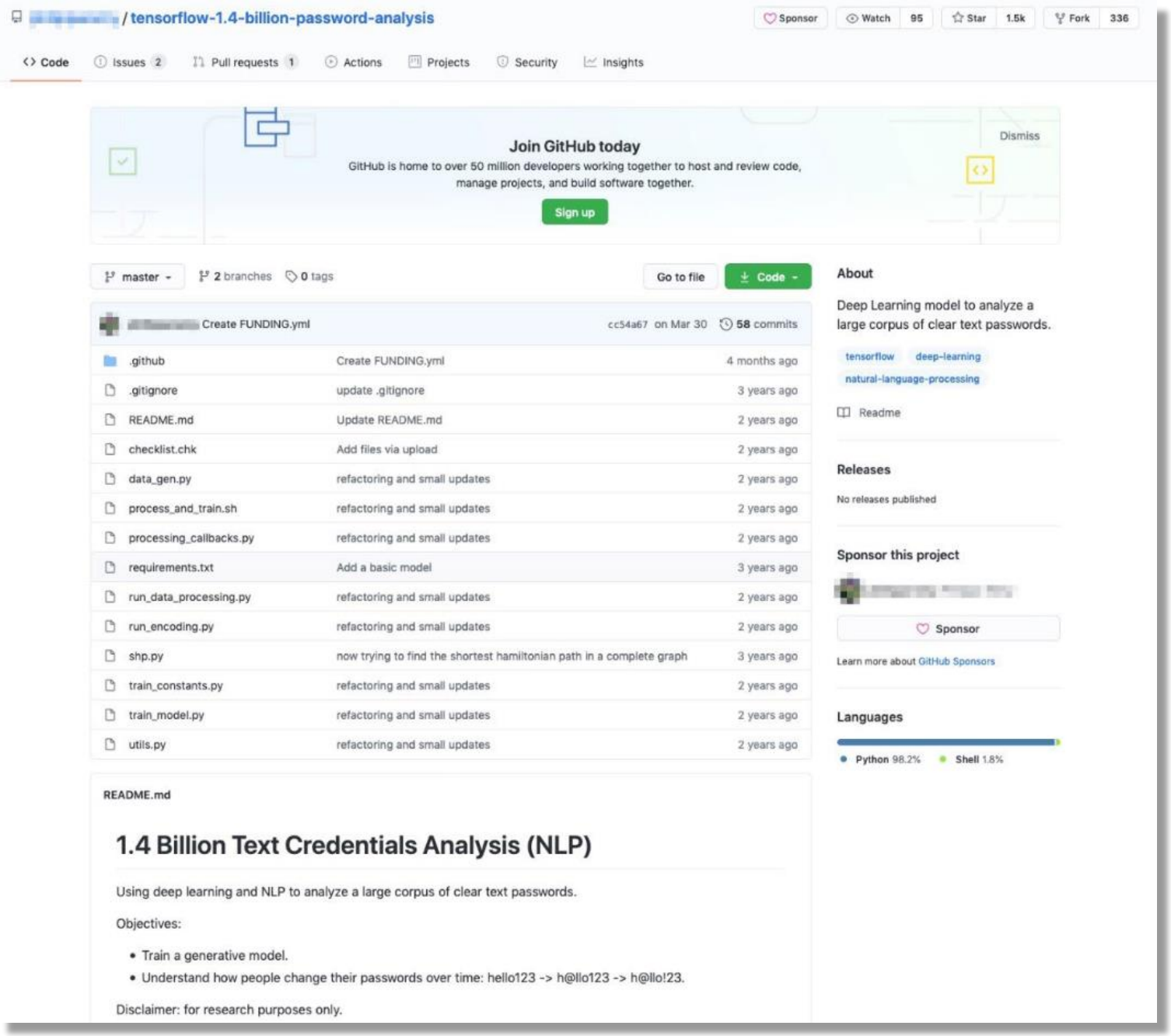
- Využití ML pro generování e-mailu, který projde filteringem (2015)
- Analýza dat pro využití útoků Business Email Compromise (2017 Black Hat USA)
- Využití ML pro analýzu detekčních schopností libovolného antimalware a na základě výsledků vytvoření nedetekovaného malware (2017 Black Hat USA, AVPASS, 0% detekcí VirusTotal z 5000 vytvořených vzorků).

Příběh DeepLocker & DeepExploit

- DeepLocker PoC Malware využívající AI přímo v rámci svého kódu
- Používá DNN (Distributed Neural Network) k určení, zda se nachází na cílovém zařízení a následně dešifruje vlastní kód a obchází prevenční systémy
- DeepExploit využívá ML k automatizaci penetračních testů a automatickém nasazení vhodného exploitu na základě detekovaného prostředí (s využitím Metasploitu)

Využití Machine Learning při útocích na hesla

- Využití ML modelů pro analýzu uniklých hesel ke zpřesnění slovníků
- Generování vzorců, podle kterých si uživatelé vytvářejí hesla
- Například PassGAN systém
  - o 50-70% úspěšnější než HashCat


/ tensorflow-1.4-billion-password-analysis
Sponsor
Watch 95
Star 1.5k
Fork 336
Code
Issues 2
Pull requests 1
Actions
Projects
Security
Insights
Join GitHub today
GitHub is home to over 50 million developers working together to host and review code, manage projects, and build software together.
Sign up
Dismiss
master
2 branches
0 tags
Go to file
Code
Create FUNDING.yml
cc54a67 on Mar 30
58 commits
.github Create FUNDING.yml 4 months ago
.gitignore update .gitignore 3 years ago
README.md Update README.md 2 years ago
checklist.chk Add files via upload 2 years ago
data_gen.py refactoring and small updates 2 years ago
process_and_train.sh refactoring and small updates 2 years ago
processing_callbacks.py refactoring and small updates 2 years ago
requirements.txt Add a basic model 3 years ago
run_data_processing.py refactoring and small updates 2 years ago
run_encoding.py refactoring and small updates 2 years ago
shp.py now trying to find the shortest hamiltonian path in a complete graph 3 years ago
train_constants.py refactoring and small updates 2 years ago
train_model.py refactoring and small updates 2 years ago
utils.py refactoring and small updates 2 years ago
About
Deep Learning model to analyze a large corpus of clear text passwords.
tensorflow deep-learning natural-language-processing
Readme
Releases
No releases published
Sponsor this project
Sponsor
Learn more about GitHub Sponsors
Languages
Python 98.2% Shell 1.8%
README.md
1.4 Billion Text Credentials Analysis (NLP)
Using deep learning and NLP to analyze a large corpus of clear text passwords.
Objectives:
Train a generative model.
Understand how people change their passwords over time: hello123 -> h@llo123 -> h@llo!23.
Disclaimer: for research purposes only.

Virtuální únosy

- Využití AI k vytvoření deepfake audio z volně dostupných zdrojů
- Nástroje typu VoiceLab dokáží generovat hlas z velmi malého vzorku dat
- Stále více zdokumentovaných případů (duben 2023, Jennifer DeStefano)
- Nebezpečné například v kombinaci se SIM swap útokem (známým cílem SIM swap útoky byl Jack Dorsey nebo Vilaik Buterin)
- SNAP - social network analysis and propensities' modeling

Trendy do budoucna

- Obcházení procesů verifikace (například při vytváření bankovního účtu)
- Vytváření obsahu, například na základě analýzy textu při BEC, DeepFake
- Extrakce zajímavého obsahu z nestrukturovaných dokumentů při exfiltraci dat (Ransomware)
- Manipulace s telemetrickými daty a exploitace samotného ML modelu
- Automatizace cílených útoků (Harpoon Whaling)

Doporučení

- Včasná **aktualizace**, instalace softwarových nebo virtálních patchů
- Nespoláhat na prevenci, extrémě důležitá je především **detekce** (XDR)
- Ochrana ve **více úrovních** (perimetr, endpoint/server, síť, web, e-mail)
- **Visibilita** (nelze chránit něco o čem nevíte)
- Monitoring **rizikových faktorů** (zranitelnosti, veřejně dostupné služby, aplikace, chování uživatelů… atd.)
- **Školení uživatelů** (ideálně zábavnou formou s pozitivní motivací)
- Investice do lidí zodpovědných za bezpečnost, **managed služby**

Přínosy AI/ML v kybernetické bezpečnosti

AI/ML v kybernetické bezpečnosti

- Detekce neznámého malware jak na základě metadat, tak na základě chování s využitím ML
- BEC pokročilá analýza komunikace, například e-mailu
- Detekce anomálií na koncových zařízeních i v síťové komunikaci
- Automatická detekce a vyšetřování nad velkým množstvím logů
- Jazykové modely zaměřené na kyberbezpečnost

Reálná ukázka využití AI při vyšetřování

Filip Marvan

Filip_Marvan@trendmicro.com