Inteligência Artificial na Cibersegurança: Deteção de Malware

Tomás Antunes Nº48511

Junho 2022

Resumo

Software malicioso, ou malware, é uma das maiores ameaças à cibersegurança. Os ataques informáticos são cada vez mais comuns, todos os dias são criados novos programas maliciosos, com o intuito de causar dano a alguém, a empresas ou até mesmo a governos. Este software malicioso pode danificar a máquina em que está hospedado, pode danificar ficheiros, roubar informações, remover permissões do utilizador ou até executar operações no dispositivo sem a autorização do utilizador. A utilização de várias técnicas de inteligência artificial, como por exemplo Deep Learning e Machine Learning, tem vindo a aumentar, sendo que estas diminuem o trabalho necessário do humano para defender a máquina, fazendo assim uma mais rápida e mais precisa deteção destes softwares maliciosos, podendo assim poupar tempo e dinheiro às entidades atacadas, e evitar danos aos sistemas informáticos.

O objetivo deste trabalho é explorar a utilização da inteligência artificial na cibersegurança, na área da deteção de malware, e como esta permite que a internet seja um local mais seguro.

Introdução

A tecnologia tem vindo a evoluir rapidamente ao longo dos anos, criando a imagem de que vivemos numa sociedade de informação. Estamos constantemente rodeados por tecnologia, que usamos no nosso dia-a-dia. Esta rápida evolução da tecnologia apresenta um grande desafio, que é defender esta nova tecnologia de cibercriminosos. A inteligência artificial pode ajudar neste desafio ao ajudar os operadores de segurança humanos a detetar software malicioso, e assim facilitar o trabalho dos mesmos, poupando tempo e

poupando dinheiro. Assim, também a defesa destes sistemas será mais precisa e mais eficaz, pois os operadores sabem exatamente o que "atacar".[3]

O que é Inteligência Artificial

Inteligência Artificial é uma tecnologia que mostra inteligência através de computadores ou outro tipo de máquinas que a use. Esta usa o estudo de algoritmos como forma de aprender e melhorar a sua capacidade de processar o seu ambiente ou de realizar uma determinada tarefa dada. Esta é bastante usada no reconhecimento facial, pois é uma tecnologia bastante eficiente e eficaz no reconhecimento de imagens. Outra das suas muitas aplicações são os motores de busca, onde é utilizada para fornecer os melhores resultados possíveis ao utilizador, normalmente com base nas informações e dados recolhidos sobre o mesmo.

No âmbito da cibersegurança a inteligência artificial é utilizada em grande parte para deteção de malware, ou seja, reconhecimento de ameaças.[10]

O Nascimento da Cibersegurança

Em 1970 o investigador Bob Thomas criou um programa chamado "Creeper", que conseguia movimentar-se pela rede ARPANET, esta era uma rede de computadores construída em 1969 para a transmissão de dados militares secretos e ligar os departamentos de pesquisa dos Estados Unidos da América, deixando um rasto por onde passava. O inventor do email, Ray Tomlinson, criou um programa chamado "Reaper"que perseguia e apagava o rasto deixado pelo programa de Bob Thomas, tornando-se assim o primeiro anti-vírus a existir, e o nascimento da cibersegurança.[2][9]

Deteção de Malware com Inteligência Artificial

Deteção de Assinaturas: Humanos vs Inteligência Artificial

Antes da inteligência artificial entrar para a área da segurança, uma das técnicas usadas para detetar software malicioso, era a deteção com base em assinaturas. Este é um método lento, e pouco eficaz, pois é realizado por operadores humanos, que podem cometer erros. O processo de desenvolver uma assinatura de deteção muitas vezes demora semanas, ou seja, o software malicioso podia já ter cumprido a sua missão, e sistemas vitais poderiam

estar comprometidos. Mesmo com uma assinatura desenvolvida, o atacante apenas precisa de fazer pequenas alterações no software para que o malware ultrapasse a mesma.

A inteligência artificial, consegue usar a mesma técnica mas de maneira muito mais eficaz. Sendo sendo fornecida uma base de dados, com várias assinaturas/padrões vistos em outros malwares, esta usa algoritmos sofisticados e eficazes para percorrer todos os ficheiros que possam ser maliciosos, à procura destes padrões. Se existir uma correspondência dentro de um ficheiro este é classificado como malicioso e é movido para quarentena, onde um operador humano, ou até mesmo o utilizador da própria máquina, pode decidir o que vai fazer com este ficheiro.[3][4][5]

Deteção de Malware baseada em Anomalias: A deteção de malware com base em Anomalias é feita com uma abordagem baseada na classificação para identificar comportamentos normais ou comportamentos estranhos. Esta aborda as limitações da técnica de Deteção de Malware por Assinaturas, na capacidade em que é capaz de avaliar aquilo que é conhecido, tal como, aquilo que não é conhecido.[6][3]

Deteção Heurística de Malware: A deteção Heurística é a técnica proposto para ultrapassar as desvantagens das técnicas anteriores. Este usa mineração de dados e algoritmos de machine learning para decifrar o comportamento de ficheiros executáveis. A deteção Heurística pode ser feita através de por exemplo, API Calls e OpCodes.

API Calls: Quase todos os programas que existem usam API Calls para fazer pedidos ao sistema operativo ou mesmo para pedir algo da internet, como por exemplo a metereologia. Estes podem ser usados para infetar máquinas com malware.

OpCodes: OpCode é uma subdivisão da linguagem de máquina que identifica a operação a ser executada. Os programas são sequências de instruções assembly, e cada instrução tem o seu próprio OpCode. Através da análise destes num programa, é possível perceber se tal programa é maligno ou benigno, tendo em conta os padrões de OpCodes presentes no programa.
[8][1]

Limitações

Como qualquer outra tecnologia, a inteligência artificial também tem limitações. A deteção de malware é dividida em duas partes, estática e dinâmica, onde ocorre o mapeamento do código do software detetado, sendo a análise dinâmica melhor, comparada com a estática. A deteção de malware estática pode falhar, tendo em conta as ferramentas que existem hoje em dia, como por exemplo, ferramentas com a capacidade de alterar ou reescrever binário, permitindo assim mascarar programas para que estes passem por benignos, ou até apenas porque falha a detetar uma anomalia que nunca foi detetada anteriormente. Uma das limitações do processo dinâmico é por exemplo, algumas operações realizadas pelo utilizador podem parar este processo, fazendo com que este possa falhar algum ou alguns vírus que estejam ativos. Este processo também assume que o estado inicial do ficheiro nunca é alterado o que nem sempre é verdade, podendo assim deixar passar alterações feitas pelo próprio vírus, fazendo com que este tenha maior probabilidade de cumprir a sua missão sem ser detetado.[3][7]

Conclusão

Concluindo, a cibersegurança é uma área que está cada vez mais a precisar de ferramentas mais rápidas e mais eficazes, devido à rápida evolução da tecnologia e cada vez haver mais ataques informáticos, e os mesmos sendo cada vez mais sofisticados. A inteligência artificial é uma tecnologia em constante desenvolvimento e melhoria, e uma tecnologia bastante útil na área da cibersegurança. Esta promete ser mais eficaz e mais rápida que o operador humano, e assim poupar recursos e dinheiro aos que usufruírem da mesma, assegurando assim uma internet mais segura.

Referências

[1] Zahra Bazrafshan, Hashem Hashemi, Seyed Mehdi Hazrati Fard, and Ali Hamzeh. A survey on heuristic malware detection techniques. In *The 5th Conference on Information and Knowledge Technology*, pages 113–120. IEEE, 2013.

[2] Cyber. The history of cybersecurity. `https://cybermagazine.com/cyber-security/history-cybersecurity`.

[3] Md Jobair Hossain Faruk, Hossain Shahriar, Maria Valero, Farhat Lamia Barsha, Shahriar Sobhan, Md Abdullah Khan, Michael Whitman, Alfredo Cuzzocrea, Dan Lo, Akond Rahman, et al. Malware detection and prevention using artificial intelligence techniques. In *2021 IEEE International Conference on Big Data (Big Data)*, pages 5369–5377. IEEE, 2021.

[4] Jannatul Ferdaos, Chandani Vaya, Anchal Bhalla, Ami Tharayil, and May El Barachi. Smart malware detection: From signatures to artificial intelligence. In *2020 5th International Conference on Smart and Sustainable Technologies (SpliTech)*, pages 1–6. IEEE, 2020.

[5] Li-Chin Huang, Chun-Hsien Chang, and Min-Shiang Hwang. Research on malware detection and classification based on artificial intelligence. volume 22, pages 717–727. , 2020.

[6] Benoît Morel. Anomaly based intrusion detection and artificial intelligence. pages 19–38, 2011.

[7] Andreas Moser, Christopher Kruegel, and Engin Kirda. Limits of static analysis for malware detection. In *Twenty-Third Annual Computer Security Applications Conference (ACSAC 2007)*, pages 421–430. IEEE, 2007.

[8] Imtithal A Saeed, Ali Selamat, and Ali MA Abuagoub. A survey on malware and malware detection systems. volume 67. Citeseer, 2013.

[9] Wikipedia. Arpanet. `https://pt.wikipedia.org/wiki/ARPANET`.

[10] Wikipedia. Artificial intelligence. `https://en.wikipedia.org/wiki/Artificial_intelligence`.