Inteligência Artificial e Malwares

A análise dos artigos "An Early Detection of Android Malware Using System Calls Based Machine Learning Model" e "Unleashing Malware Analysis and Understanding With Generative AI" revela abordagens distintas e complementares para a detecção e compreensão de malwares. Ambos os estudos buscam melhorar a capacidade de identificação e interpretação de comportamentos maliciosos, mas cada um adota um caminho metodológico diferente, refletindo tendências emergentes na segurança cibernética.

O primeiro artigo enfoca a detecção precoce de malwares em dispositivos Android por meio da análise de chamadas de sistema, aplicando técnicas de Machine Learning (ML). A proposta consiste em capturar chamadas de sistema nos estágios iniciais da execução de aplicativos e, através de modelos como o Multi-Layer Perceptron (MLP), alcançar uma acurácia de 99,34% na classificação de aplicativos benignos e maliciosos. A abordagem destaca-se pela eficiência e precisão, utilizando n-gramas e TF-IDF para identificar padrões comportamentais que diferenciam malwares de aplicativos legítimos. Essa técnica de detecção precoce visa uma implementação prática, permitindo que a segurança seja aplicada em tempo real, o que é crucial em sistemas operacionais móveis.

Em contraste, o segundo artigo investiga a aplicação de modelos de linguagem natural (LLMs) na análise e geração de relatórios de inteligência sobre ameaças. O estudo propõe a criação de gráficos de cenários de ataque (ASG) para sintetizar longos traços de syscalls em representações visuais e acessíveis. A metodologia transforma chamadas de sistema em descrições em linguagem natural (NLDs), permitindo que o modelo ChatGPT produza relatórios narrativos que descrevem a operação do malware. Essa abordagem inovadora de tradução de traços técnicos para uma linguagem compreensível facilita a interpretação para analistas e reforça a documentação sobre comportamentos de malware. Um diferencial importante é o potencial para que LLMs forneçam insights adicionais e interpretem as intenções subjacentes do malware, contribuindo para um entendimento mais profundo e contextualizado das ameaças.

Uma semelhança entre os dois estudos é o uso de chamadas de sistema como base para a análise de comportamento de malware. Ambos reconhecem a importância de capturar as interações do malware com o sistema operacional, mas o fazem de maneiras distintas: o primeiro se concentra na classificação precoce usando um número limitado de chamadas de sistema, enquanto o segundo foca em traduzir grandes volumes de traços em relatórios descritivos e interpretativos. Outra diferença significativa está nos objetivos principais: enquanto o primeiro artigo visa a

detecção rápida e eficiente de ameaças, o segundo busca enriquecer a compreensão e a análise detalhada das atividades maliciosas.

Ao avaliar as abordagens, observa-se que ambas têm vantagens e desafios. A abordagem baseada em ML é prática e escalável, mas pode ser limitada na interpretação de intenções complexas. Já o uso de LLMs para geração de relatórios oferece profundidade interpretativa, embora enfrente desafios técnicos, como a necessidade de ajuste fino para evitar a geração de informações imprecisas ou irrelevantes. Essas diferenças refletem a diversidade de estratégias que podem ser aplicadas em Sistemas Operacionais, evidenciando que a combinação de técnicas, como ML para detecção rápida e LLMs para interpretação, poderia formar um sistema de segurança robusto e completo.