AUDITORIA PREDITIVA DE LICENCIAMENTO DE SOFTWARE E INTEGRIDADE DE KERNEL: MAPEAMENTO DE RISCOS DE CONFORMIDADE E MOVIMENTAÇÃO LATERAL VIA REDES NEURAIS

Ricardo Pereira C. Santos[1] - januariaricardo@gmail.com
Nilton Alves Maia[2] - nilton.maia@unimontes.br
[1]Mestrando em Modelagem Computacional e Sistemas – Unimontes.
[2]Docente do Programa de Pós-Graduação em Modelagem Computacional e Sistemas – Unimontes.

Abstract. *Este trabalho apresenta uma evolução do Framework MDR-SLA (Managed Detection and Response - Software Licensing Auditor), voltado para auditoria contínua e detecção de anomalias estruturais em sistemas operacionais corporativos. A pesquisa aborda a correlação estatística entre o desvio do subsistema de licenciamento do Windows (sppsvc) e a introdução de vetores de movimentação lateral. Utilizando uma Rede Neural Artificial Perceptron Multicamadas (MLPClassifier) acoplada a um pipeline de dados dinâmicos de vulnerabilidades (API CVE/CVSS) e inteligência de ameaças via API do VirusTotal, o modelo segregou inconformidades administrativas de riscos críticos à segurança empresarial. Testes em laboratório validaram a eficácia do classificador estatístico de 4 dimensões em inferir o risco de colapso de integridade local, fornecendo insumos em tempo real para Centros de Operações de Segurança (SOC).*

Keywords: *RNA; MDR Corporativo; Vulnerabilidade; VirusTotal; Compliance.*

1. INTRODUÇÃO

A superfície de ataque em redes corporativas expandiu-se com a descentralização de ativos, trabalho remoto e códigos maliciosos persistentes. Na governança de TI, o compliance digital constitui a primeira linha de defesa contra incidentes cibernéticos graves (FONTES, 2022). Contudo, uma das brechas mais negligenciadas reside na integridade interna dos sistemas operacionais (SO) das estações de trabalho, onde o uso de ferramentas de bypass de licenciamento (ativadores piratas ou cracks) representa uma vulnerabilidade crônica que quebra o princípio de custódia da segurança (NAKAMURA; GEUS, 2007).

Ferramentas como KMSPico, KMSAuto e scripts baseados em Direito Digital (HWID) violam o Kernel do sistema ao estabelecer servidores proxy locais em loopback (127.0.0.1) ou realizar interceptações na memória RAM (Hooking) para silenciar o subsistema de

licenciamento do Windows (`sppsvc.exe`). Para garantir persistência, exigem a desativação ou inserção na lista de exceções em soluções de EDR e antivírus.

Em redes empresariais, o comprometimento desses endpoints cria uma ponte para a infecção e movimentação lateral (pivoteamento). Conforme a matriz MITRE ATT&CK (2024), a movimentação lateral abrange técnicas para estender o acesso via sistemas remotos, permitindo o roubo de credenciais corporativas em cache e o mapeamento de ativos críticos como os controladores de domínio (Active Directory).

Este trabalho propõe um modelo preditivo baseado em agentes de aprendizado supervisionado (RUSSELL; NORVIG, 2013), estruturado via Framework MDR-SLA. O sistema automatiza a coleta heurística por scripts de inventário, extrai assinaturas criptográficas dos binários em disco e consulta dinamicamente gateways de segurança (VirusTotal e APIs CVE). Esses fatores são unificados em uma rede neural MLPClassifier para classificar a severidade do ativo no ecossistema empresarial e disparar playbooks automatizados de contenção, reduzindo o MTTR (Mean Time to Respond).

2. REFERENCIAL TEÓRICO E METODOLOGIA

2.1 Coleta, Enriquecimento e Modelagem de Dados

O Framework opera sob um modelo descentralizado de coleta e processamento centralizado. A telemetria baseia-se em um script nativo em PowerShell nos hosts. Adotou-se o padrão JSON (JavaScript Object Notation) para tráfego leve e estruturado em chave-valor (CROCKFORD, 2018), serializando caminhos de registro, hashes e estados de licença. Em cada estação (VM), o agente varre o status da licença via WMI (Windows Management Instrumentation) e calcula o hash SHA-256 de binários ou scripts alteradores.

Para anular falsos negativos, inseriu-se um pipeline de enriquecimento de dados via Inteligência de Ameaças (Threat Intelligence) conectada à API do VirusTotal, submetendo os hashes ao consenso de mais de 70 motores globais de antivírus. Para alimentar a Rede Neural, os dados foram estruturados em um vetor de recursos de quatro dimensões (X):

$$X = [X_{\text{ativador}}, X_{\text{licenca}}, X_{\text{cvss}}, X_{\text{vt}}] \tag{1}$$

Onde $X_{\text{ativador}} \in \{0, 1\}$ indica a presença de ativadores; $X_{\text{licenca}} \in \{0, 1\}$ denota o não-licenciamento; $X_{\text{cvss}} \in [0.0, 10.0]$ é a maior nota de vulnerabilidade ativa extraída via API do NVD/NIST; e $X_{\text{vt}} \in [0.0, 70.0]$ reflete a quantidade de motores do VirusTotal que classificaram o binário como malicioso.

2.2 Pré-Processamento e Normalização Estatística

Devido à disparidade escalar entre as variáveis binárias (0 ou 1), o Score CVSS (0 a 10) e a escala do VirusTotal (0 a 70), o uso direto dos dados causaria distorções nos gradientes na retropropagação (backpropagation). Implementou-se a padronização via StandardScaler, aplicando o cálculo da pontuação Z (Z-Score):

$$z = \frac{x - \mu}{\sigma} \tag{2}$$

Onde μ representa a média da amostra de treinamento e σ denota o desvio padrão de cada recurso do vetor de entrada.

2.3 Topologia e Treinamento da Rede Neural (MLPClassifier)

O núcleo preditivo baseia-se em uma rede Perceptron Multicamadas (feedforward), cuja arquitetura hierárquica e capacidade de aproximação universal de funções são fundamentadas no mapeamento de padrões não-lineares (HAYKIN, 2001). A topologia possui 4 neurônios na Camada de Entrada; Oculta 1 com 8 neurônios; Oculta 2 com 4 neurônios; e Saída com 2 neurônios de classificação probabilística para o SOC (STALLINGS, 2014):

- **Classe 0 (Risco Mitigável / Aceitável):** Estações em conformidade total ou com falhas meramente burocráticas de licença vencida, sem sinais de violação de código.
- **Classe 1 (Ameaça Crítica):** Sistemas com comportamento de evasão, manipulação de Kernel e assinaturas maliciosas confirmadas por múltiplos motores de Threat Intel.

Para a camada de saída, utilizou-se a função de ativação Sigmóide:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

A base exponencial $e(\approx\ 2{,}71828)$ realiza a compressão assintótica no intervalo $[0, 1]$, convertendo os outputs internos em probabilidade direta de risco (0% a 100%) e otimizando o cálculo dos gradientes no backpropagation. O otimizador de pesos sinápticos escolhido foi o L-BFGS por sua convergência rápida em dados estruturados.

3. RESULTADOS E DISCUSSÃO

O framework foi testado em um ambiente com 10 estações virtuais corporativas configuradas com diferentes estados de integridade para simular um cenário real.

3.1 Automação da Ingestão e Dashboard Streamlit

Desenvolveu-se um script em PowerShell nativo executado localmente. O script extrai chaves de registro de evasão, o mapeamento do Software Protection Platform, o identificador do Kernel, flags de scripts maliciosos e hashes criptográficos. O payload consolida essas variáveis em arquivos JSON e os envia ao diretório de ingestão do MDR.

A camada visual foi implementada via Streamlit e hospedada em nuvem. A aplicação processa assincronamente os arquivos JSON das VMs, expondo métricas, matriz de conformidade e o modo de inspeção aprofundada da IA (Verbose), onde inspecionam o vetor de bias (b), pesos (W) e propagação direta (Forward Propagation). A pipeline de pré-processamento, normalização de dados e treinamento da rede neural Perceptron Multicamadas foi construída utilizando os módulos da biblioteca Scikit-learn em linguagem Python (PEDREGOSA et al., 2011). O portal público está disponível em: `https://mdr-sla-licence.streamlit.app`.

3.2 Comportamento e Inferência da Rede MLP no SOC

O comportamento preditivo da rede MLP demonstrou estabilidade absoluta. As amostragens dos testes e inferências em tempo real estão sumarizadas na Tabela 1.

Table 1: Matriz de Auditoria, Heurística Expandida com VirusTotal e Predição da Rede MLP.

Hostname	**Status / Licenciamento**	**CVSS**	**Motores VT**	**Prob. Risco**	**Predição**
DESKTOP1	Não Licenciado / Suspeito	5.00	0/70	0.1%	ATENÇÃO
DESKTOP2	Licenciado Script Suspeito	9.80	53/70	100.0%	CRÍTICO
DESKTOP3	Original Licenciado	0.00	0/70	3.9%	ACEITÁVEL
DESKTOP4	Licenciado Script Suspeito	9.80	53/70	100.0%	CRÍTICO
DESKTOP5	Não Licenciado / Suspeito	5.00	0/70	0.1%	ATENÇÃO
DESKTOP6	Não Licenciado / Suspeito	5.00	0/70	0.1%	ATENÇÃO
DESKTOP7	Original Licenciado	0.00	0/70	3.9%	ACEITÁVEL
DESKTOP8	Original Licenciado	0.00	0/70	3.9%	ACEITÁVEL
DESKTOP9	Original Licenciado	0.00	0/70	3.9%	ACEITÁVEL
DESKTOP10	Original / Licenciado	0.00	0/70	3.9%	ACEITÁVEL

Ativos conformes foram mapeados como ACEITÁVEL (risco marginal de 3.9%). Inconformidades estritamente administrativas (DESKTOP1, 5 e 6) obtiveram CVSS 5.00, mas com zero positivações no VirusTotal (0/70), o modelo mitigou a probabilidade para 0.1% (ATENÇÃO). Já os hosts com injetores de memória (DESKTOP2 e 4) acionaram 53 dos 70 motores globais; a MLP elevou a probabilidade de ameaça para 100.0% (CRÍTICO, Score 9.80).

3.3 Estudo de Caso de Ameaça Real: Vetor de Persistência e Mascaramento

Para validar a eficácia prática, o ambiente foi submetido a uma infecção real por vetor de persistência de malware. A análise forense revelou persistência em duas chaves anômalas no registro: `HKLM\Software\Microsoft\Windows\CurrentVersion\Run`, conforme demonstrado na Figura 1.

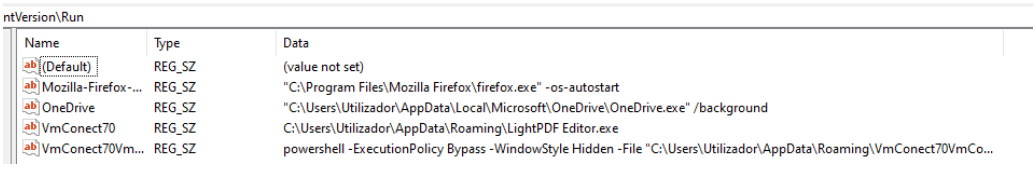

Figure 1: Evidência forense das chaves de persistência anômalas e comandos PowerShell mascarados no Registro do Windows.

A primeira entrada (VmConect70) executava o binário disfarçado `LightPDF Editor.exe` em `AppData\Roaming`. A segunda disparava uma instância do PowerShell (`-ExecutionPolicy Bypass -WindowStyle Hidden`) para evadir políticas locais. Essa técnica de mascaramento (Masquerading) foi correlacionada pelo framework que, ao cruzar a ausência de assinatura com a modificação de Run, atribuiu o Score Crítico (100%).

4. CONCLUSÃO

O Framework MDR-SLA demonstrou a viabilidade de aplicar inteligência artificial estruturada na governança e defesa cibernética corporativa. A transição para uma auditoria preditiva enriquecida por dados de vulnerabilidades (CVE) e inteligência multimotores (VirusTotal) permitiu segregar inconformidades administrativas de ameaças graves à integridade do sistema operacional.

Trabalhos futuros focarão na expansão multi-plataforma e ingestão de métricas de saúde de TI (CPU, RAM e disco). Planeja-se também a integração nativa com barramentos de microrsegmentação Zero Trust para bloqueio físico instantâneo de portas de switch de ativos classificados como críticos.

Agradecimentos

Os autores agradecem à Universidade Estadual de Montes Claros (Unimontes) e ao Programa de Pós-Graduação em Modelagem Computacional e Sistemas pelo suporte institucional.

REFERENCES

Crockford, Douglas (2018), *How JavaScript Works*, Virginia: Virman Press, 2018.

Fontes, Edison (2022), *Segurança da Informação: Governança, Compliance e Gestão de Riscos*, 2. ed. São Paulo: Saraiva, 2022.

Haykin, Simon (2001), *Redes Neurais: Princípios e Prática*, 2. ed. Porto Alegre: Bookman, 2001.

MITRE ATT&CK (2024), "Lateral Movement: Tactics and Techniques Matrix", *MITRE*, 2024. Disponível em: `https://attack.mitre.org/tactics/TA0008/`. Acesso em: 6 jun. 2026.

Nakamura, Emilio T.; Geus, Paulo L. (2007), *Segurança de Redes em Ambientes Corporativos*, São Paulo: Novatec, 2007.

Pedregosa, Fabian et al. (2011), "Scikit-learn: Machine Learning in Python", *Journal of Machine Learning Research*, v. 12, p. 2825-2830, 2011.

Russell, Stuart; Norvig, Peter (2013), *Inteligência Artificial: uma abordagem moderna*, 3. ed. Rio de Janeiro: Elsevier, 2013.

Stallings, William (2014), *Criptografia e Segurança de Redes: Princípios e Práticas*, 6. ed. São Paulo: Pearson, 2014.