Priorização de Alertas de Intrusão

Wellinton Piassa[1], **Bruno Bogaz Zarpelão**[1]

[1]Departamento de Computação – Universidade Estadual de Londrina (UEL)
Caixa Postal 10.011 – CEP 86057-970 – Londrina – PR – Brasil

wellinton.piassa@uel.br, brunozarpelao@uel.br

Abstract. *Together with the exponential growth of the Internet, it has been emerging a need for strengthening network security against attacks and intrusions and keeping sensitive data safe. An Intrusion Detection System (IDS) is a tool for preserving network security, as it alerts network administrators of the occurrence of malicious events, and even performs pre-defined actions for countering them. A common problem that may arise when using IDSs is the generation of a high volume of alerts, of which many might be just false positive instances. That issue hinders network administrators from effectively dealing with real alerts and preserving network security. Thus, this work aims to study methods and techniques for prioritizing alerts to help administrators to focus only on the most important ones and identify the potentially false ones.*

Resumo. *Com o grande crescimento da Internet, vem-se criando uma necessidade de fortalecer a segurança das redes contra ataques e invasões de atores mal intencionados, e proteger dados pessoais importantes. O Sistema de Detecção de Intrusão é uma opção para ajudar a combater os ataques, informando o administrador da rede sobre a ocorrência de eventos de segurança, e até mesmo executando ações pré-definidas. Um problema que aparece ao utilizar o Sistema de Detecção de Intrusão é, primeiramente, a geração de grandes volumes de alertas. Soma-se a isso o fato de que, dentre os alertas que são gerados, existem uma parte deles que são falsos-positivos, atrapalhando o administrador de redes a dar um foco aos alertas mais importantes. Visando abordar essa questão, este trabalho tem a proposta de estudar técnicas para priorizar alertas, com intuito de ajudar administradores de redes a dar ênfase aos alertas que devem ser analisados primeiro, e até mesmo identificar quais desses alertas são potencialmente falsos.*

1. Introdução

Com a grande evolução da tecnologia, surgiram ferramentas que se tornaram essenciais no mundo atual e dentre elas está a Internet que, de forma bem simples, é uma enorme quantidade de máquinas interligadas em todo o mundo, conectadas e trocando informações entre si [5]. A partir disso, serviços puderam ser ofertados através da internet, tais como Lojas Virtuais, Banco de Dados, serviços financeiros, serviços governamentais, dentre outros.

Sendo assim, dados importantes e confidenciais são transmitidos e armazenados nessas máquinas que estão conectadas à Internet, despertando atenção de usuários mal intencionados que aproveitam vulnerabilidades das redes para realizar ataques prejudiciais. Portanto, medidas de segurança são fundamentais para combater todo tipo de ameaça,

e proteger dados sensíveis que podem ter valores inestimáveis para uma instituição [8], além do próprio ambiente de funcionamento das máquinas.

Visando isso, surgiram vários desafios para a área de Segurança de Redes, pois o rápido desenvolvimento da tecnologia da informação dificulta a construção de redes totalmente confiáveis. Ou seja, detectar e combater todo tipo de ataque é uma tarefa bastante difícil, pois existem vários tipos que ameaçam a integridade, disponibilidade e confiabilidade dos computadores. Dentre esses ataques, podemos citar como exemplo aquele que se refere ao de negação de serviço (DoS), considerado um dos ataques prejudiciais mais comuns. [3].

Para auxiliar no combate e na prevenção de ataques, algumas ferramentas estão sendo desenvolvidas e aprimoradas, e uma dessas ferramentas são os IDSs (*Intrusion Detection System* - Sistema de Detecção de Intrusão). O IDS foi implementado pela primeira vez em 1987 e desde então vem se tornando um tema forte de pesquisa, já que é uma ferramenta importante para a segurança de computadores [1]. Um IDS monitora o tráfego que flui na rede, verifica violações a políticas e atividades maliciosas, e, caso algum tipo de atividade prejudicial for descoberta, é relatado ao administrador da rede [21].

Porém, os IDSs geralmente produzem muitos alertas classificados como falsos-positivos, que são dados ou comportamentos normais na rede, julgados como como uma ameaça, fazendo um alerta ser gerado. [9]. Se uma quantidade muito grande dos alertas forem falsos-positivos, implicará na perda de desempenho do IDS, e tornará o processo do administrador de redes muito mais complexo ao lidar com esse grande volume de alertas.

Pensando nesse problema, a proposta deste trabalho é estudar e aplicar técnicas para priorizar os alertas. A proposta será baseada em alguns métodos que irão analisar grandes volumes de alertas, e estabelecer alguns critérios para priorizá-los. Dentre os critérios, uma das hipóteses é utilizar fontes de dados externas ao IDS, como *blacklists* e informações sobre os *hosts* atacados. Portanto, será possível ajudar administradores de redes a lidar mais facilmente com o grande volume de alertas, ou até mesmo identificar quais alertas são potencialmente falsos.

2. Fundamentação Teórico-Metodológica e Estado da Arte

2.1. Sistema de Detecção de Instrusão

Um Sistema de Detecção de Intrusão (Intrusion Detection System - IDS) é uma ferramenta de segurança que tem como objetivo monitorar o comportamento e o tráfego da rede em busca de encontrar atividades maliciosas e violações políticas [21, 12]. O IDS fica analisando a rede em busca de uma atividade maliciosa, e, no momento que é detectada, um alerta é enviado ao administrador de rede para que sejam tomadas as medidas necessárias. [14, 21, 13, 4, 17].

O IDS pode ser baseado em dois grupos dependendo da topologia de rede e quanto ao seu posicionamento nela, sendo esses grupos Sistema baseado em redes (NIDS) e Sistema baseado em *host* (HIDS) [1, 16, 13]. Um NIDS geralmente se encontra no ponto de entrada de uma rede, e analisa os pacotes que são trafegados nela junto com seu cabeçalho na busca de detectar alguma ameaça [1]. Já o HIDS é instalado em um único sistema, realizando o monitoramento do tráfego de rede daquele dispositivo, e além disso, analisa

logs/eventos do sistema, rastrear processo e ter acessos a mudanças de arquivos do sistema [6, 16, 13].

Existem dois métodos de detecção em IDS, por assinatura e por anomalias [1, 4, 2, 5], cada um tendo suas vantagens e desvantagens. No método por anomalias, a detecção é baseada no comportamento da rede, ou seja, é estabelecido um comportamento padrão usando aprendizado de máquina, métodos baseados em estatística ou baseados em conhecimento [12]. Uma vez que ocorre o desvio do comportamento padrão, ele será considerado como uma anomalia ou um ataque [1]. A vantagem de utilizar a detecção por anomalia é conseguir detectar ataques tanto conhecidos como ataques desconhecidos, possibilitando que o IDS consiga lidar com a constante mudança da natureza de ataques [2]. Sua desvantagem seria o potencial de gerar um grande número de alertas falsos-positivos, podendo caracterizar comportamentos legítimos, como uma anomalia [6, 3, 14].

No método de detecção por assinatura, o sistema possui armazenado padrões de vários ataques já conhecidos, e caso a situação que está sendo analisada se encaixe em um desses padrões, ele será identificado como uma atividade maliciosa [5, 21]. A vantagem dessa abordagem é ter um volume bem menor de alertas falsos-positivos em comparação com a abordagem de detecção por anomalia e sendo mais eficiente com ataques conhecidos [14]. Já sua desvantagem seria a impossibilidade de detectar ataques desconhecidos, pois não se enquadram nas regras que definidas em armazenamento [1, 12].

2.2. Processamento de Alertas

A utilização do IDS (Intrusion Detection System) vem ajudando fortemente os administradores de redes a minimizar os danos causados por diferentes ataques [7]. No entanto, lidar com grandes volumes de alertas desorganizados e, muitas vezes irrelevantes, se torna uma tarefa bastante difícil, dificultando a vida do administrador de redes [18]. Portanto algumas técnicas vêm sendo estudadas e aplicadas na área e, dentre essas técnicas estão a correlação, agregação, redução e priorização dos alertas.

2.2.1. Correlação de Alertas

A correlação de alerta é um método para encontrar relação entre alertas, onde é levado em consideração as suas similaridades [15]. O objetivo das correlações é formar um grupo de alertas para apresentar a visão de um cenário de ataque que está acontecendo, ou prever um que possivelmente pode acontecer [7].

A correlação de alertas pode ser classificada em três métodos, sendo eles o método baseado em regras, método baseado em causa e consequência e o método baseado em similaridade [19].

2.2.2. Agregação de Alertas

A agregação de alertas tem como intuito diminuir o volume de alertas que são similares, fazendo a redução de redundâncias e tornando mais fácil a análise posterior dos alertas. Para realizar a agregação, são utilizados métodos de Clusterização e Comparação de Atributos, onde cada um deles possui suas técnicas [11].

No método de Clusterização, é feita a separação dos alertas similares dos dissimilares [11], sendo assim, um só alerta vai representar todo o *cluster*, ou seja, todos os outros alertas similares. O método de Comparação de Atributos pode ser considerado mais simples do que a de Clusterização. O método realiza o agrupamento de alertas baseado nos valores do seus atributos, como por exemplo endereços IP de origem e destino.

2.2.3. Redução de Alertas Falsos-Positivos

Alertas Falsos-Positivos são alertas que os IDS's geram sobre uma atividade normal, porém, considerada pelo IDS como maliciosa [10]. Os alertas falsos, além de tornar mais complexa sua manutenção, ainda reduz o desempenho do IDS, portanto reduzir o número dessas ocorrências é de grande importância [9].

Para realizar as reduções necessárias, técnicas de Aprendizado de Máquina como Árvore de Decisão e Máquina de Vetor de Suporte, são utilizados [11]. Ambas técnicas utilizam o Aprendizado de Máquina Supervisionado, portanto necessitam de dados rotulados para treinar os algoritmos, e rotular manualmente um conjunto de dados é uma tarefa bastante árdua.

2.2.4. Priorização de Alertas

Dentre os vários alertas que o administrador da rede tem que lidar, alguns deles têm maiores prioridades e necessitam de mais atenção. Definir qual, ou quais alertas devem ser priorizados varia de acordo com o contexto, e essa importância pode ser definida pelo administrador de rede [20].

Exemplos de alertas de alta prioridade são, os alertas que podem estar relacionados a ataques contra elementos importantes de um sistema, ou ataques com alta taxa de sucesso [11]. Uma das técnicas usadas para realizar a priorização de alertas é a técnica de Árvore de Decisão [11].

3. Objetivos

O trabalho tem como objetivo geral estudar e aplicar métodos para a priorização de alertas de intrusão gerados por um Sistema de Detecção de Intrusão. Para alcançar tal objetivo geral, os seguintes objetivos específicos foram definidos:

1. Estudar dados dos eventos que possam contribuir no processo de priorização dos alertas.
2. Estudar técnicas de análise e mineração de dados que possam realizar a priorização de forma automatizada para subsequentemente aplicá-las.
3. Ao final, definir um processo completo na qual se inicia coletando os alertas e finalize apresentando esses alertas de uma maneira fácil para se analisar, além da automatização dos alertas de maior interesse.
4. Realizar a implementação dos algoritmos de priorização baseado nas técnicas estudadas e definidas.

4. Procedimentos metodológicos/Métodos e técnicas

Em um primeiro momento, será feito um levantamento dos dados que são necessários para realizar a análise dos alertas, definindo quais dados são mais relevantes e devem ser enfatizados. Logo em seguida, será feito um levantamento dos possíveis critérios que devem ser levados em conta ao analisar os alertas.

Após o levantamento dos dados e dos critérios que serão utilizados para analisar os eventos, será realizada a obtenção dos conjuntos de alertas e das demais informações denominadas necessárias. O conjunto de alertas a ser trabalhado será obtido através de fontes públicas na internet, *datasets* com um grande volume de alertas com diferentes tipos de particularidades de cada ataque, além de conter alertas Falsos-Positivos.

Posteriormente, será realizada uma revisão bibliográfica de trabalhos similares na área, com o intuito de estudar as técnicas existentes. Após o levantamento das técnicas existentes, será definido qual será utilizada para analisar e aplicar a priorização dos alertas, levando em consideração critérios e dados levantados anteriormente.

Com a revisão realizada e com a(s) técnica(s) definida(s), inicia-se o processo de implementação da ferramenta que irá fazer a análise e a priorização dos conjuntos de alertas. Tendo a implementação feita, será possível realizar os testes sobre os dados obtidos, buscando visualizar quais pontos devem ser ajustados e melhorados.

Por fim, uma avaliação será feita com o propósito de observar a capacidade de redução do conjunto de alertas, e se há possibilidade de perda de eventos de interesse. Além do mais, apurar os alertas que foram priorizados, no intuito de aferir se a priorização realmente ressalta os alertas com maior chance de ser um ataque.

5. Cronograma de Execução

Atividades:

1. Definir critérios e dados necessários para a análise dos alertas.
2. Obter um conjunto de dados.
3. Fazer Revisão Bibliográfica das técnicas existentes e definir a que melhor fará a priorização.
4. Fase de implementação.
5. Análise da capacidade de redução do conjunto de alertas, realizando ajustes na implementação caso necessário.
6. Escrita do Trabalho de Conclusão de Curso.

Tabela 1. Cronograma de Execução

	ago	set	out	nov	dez	jan	fev	mar	abr	mai
Atividade 1	x	x								
Atividade 2		x	x							
Atividade 3		x	x	x	x					
Atividade 4					x	x	x			
Atividade 5								x	x	x
Atividade 6						x	x	x	x	x

6. Contribuições e/ou Resultados esperados

O projeto em questão almeja fazer a priorização de alertas gerados por Sistemas de Detecção de Intrusão, com intuito de ajudar o(s) administrador(es) da rede a lidar com os alertas que realmente são importantes e que geram de fato um risco para o sistema.

Além disso, espera-se estudar e utilizar de uma técnica para aplicar nesse contexto de detecção e filtragem de alertas. Também implementar uma ferramenta que possa trazer confiabilidade nas priorizações, que pode ser aplicado em um ambiente de mercado, ajudando a equipe que trabalha com a segurança de rede a lidar com grandes volumes de alertas diários.

7. Espaço para assinaturas

Londrina, *12 de Setembro de 2022*.

Aluno

Orientador

Referências

[1] Abdulla Amin Aburomman and Mamun Bin Ibne Reaz. Survey of learning methods in intrusion detection systems. In *2016 international conference on advances in electrical, electronic and systems engineering (ICAEES)*, pages 362–365. IEEE, 2016.

[2] Lubna Ali Hassan Ahmed and Yahia Abdalla Mohamed Hamad. Machine learning techniques for network-based intrusion detection system: A survey paper. In *2021 National Computing Colleges Conference (NCCC)*, pages 1–7. IEEE, 2021.

[3] Mohammad Almseidin, Maen Alzubi, Szilveszter Kovacs, and Mouhammd Alkasassbeh. Evaluation of machine learning algorithms for intrusion detection system. In *2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY)*, pages 000277–000282. IEEE, 2017.

[4] Khalid Alsubhi, Ehab Al-Shaer, and Raouf Boutaba. Alert prioritization in intrusion detection systems. In *NOMS 2008-2008 IEEE Network Operations and Management Symposium*, pages 33–40. IEEE, 2008.

[5] Nitesh Singh Bhati and Manju Khari. Comparative analysis of classification based intrusion detection techniques. In *2021 5th International Conference on Information Systems and Computer Networks (ISCON)*, pages 1–6. IEEE, 2021.

[6] Anna L Buczak and Erhan Guven. A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications surveys & tutorials*, 18(2):1153–1176, 2015.

[7] Huwaida Tagelsir Ibrahim Elshoush. An innovative framework for collaborative intrusion alert correlation. In *2014 Science and Information Conference*, pages 607–614. IEEE, 2014.

[8] Kathleen Goeschel. Reducing false positives in intrusion detection systems using data-mining techniques utilizing support vector machines, decision trees, and naive bayes for off-line analysis. In *SoutheastCon 2016*, pages 1–6. IEEE, 2016.

[9] Liang Hu, Taihui Li, Nannan Xie, and Jiejun Hu. False positive elimination in intrusion detection based on clustering. In *2015 12th International conference on fuzzy systems and knowledge discovery (FSKD)*, pages 519–523. IEEE, 2015.

[10] Neminath Hubballi and Vinoth Suryanarayanan. False alarm minimization techniques in signature-based intrusion detection systems: A survey. *Computer Communications*, 49:1–17, 2014.

[11] Cláudio Toshio Kawakani. Geração online de hiperalertas com base no histório de estratégias de ataque. Master's thesis, Universidade Estadual de Londrina, 2017.

[12] Ansam Khraisat, Iqbal Gondal, Peter Vamplew, and Joarder Kamruzzaman. Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, 2(1):1–22, 2019.

[13] Wenjuan Li, Weizhi Meng, and Lam For Kwok. Surveying trust-based collaborative intrusion detection: State-of-the-art, challenges and future directions. *IEEE Communications Surveys & Tutorials*, 24(1):280–305, 2021.

[14] Zakiyabanu S. Malek, Bhushan Trivedi, and Axita Shah. User behavior pattern -signature based intrusion detection. In *2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, pages 549–552, 2020.

[15] Eduardo Alves Moraes. Aplicação de Aprendizado de Máquina Supervisionado e técnicas de correlação na análise de alertas de instrusão. Master's thesis, Universidade Estadual de Londrina, 2018.

[16] Dipika Narsingyani and Ompriya Kale. Optimizing false positive in anomaly based intrusion detection using genetic algorithm. In *2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE)*, pages 72–77. IEEE, 2015.

[17] Tadeusz Pietraszek. Using adaptive alert classification to reduce false positives in intrusion detection. In *International workshop on recent advances in intrusion detection*, pages 102–124. Springer, 2004.

[18] Phillip A Porras, Martin W Fong, and Alfonso Valdes. A mission-impact-based approach to infosec alarm correlation. In *International Workshop on Recent Advances in Intrusion Detection*, pages 95–114. Springer, 2002.

[19] Saeed Salah, Gabriel Maciá-Fernández, and Jesús E Díaz-Verdejo. A model-based survey of alert correlation techniques. *Computer Networks*, 57(5):1289–1317, 2013.

[20] Kristijan Vidović, Ivan Tomičić, Karlo Slovenec, Miljenko Mikuc, and Ivona Brajdić. Ranking network devices for alarm prioritisation: Intrusion detection case study. In *2021 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pages 1–5. IEEE, 2021.

[21] Charvi Vij and Hemraj Saini. Intrusion detection systems: Conceptual study and review. pages 694–700, 2021.