Proposta de Método de Detecção de Intrusão utilizando kNN Clusterizado

Natan Morya Paredes, Renato Bobsin Machado, Gustavo dos Santos Vieira
Laboratório de Pesquisa em Segurança Computacional (LaPSeC)
Universidade Estadual do Oeste do Paraná (UNIOESTE)
Foz do Iguaçu, Brasil

***Abstract*— The increase number of security incidents has brought the need to use methods to ensure the security of information exchanged over the Internet. Computational attacks can be identified and prevented by intrusion detection systems. These systems can monitor computer networks or isolated systems in order to identify events that occur and prevent them. This work proposes a method for intrusion detection using k-Nearest Neighbors with clustering techniques in order to reduce the computational cost.**

Keywords — segurança computacional; ataques computacionais; metodologias de detecção; k-Nearest Neighbor.

***Resumo*— O aumento dos índices de segurança computacional trouxe a necessidade de utilizar métodos para garantir a segurança de informações trocadas pela Internet. Ataques computacionais podem ser identificados e impedidos por sistemas de detecção de intrusão. Esses sistemas podem monitorar redes de computadores ou sistemas isolados a fim de identificar eventos que ocorram e impedi-los. Este trabalho propõe um método para detecção de intrusão utilizando k-*Nearest Neighbor* com técnicas de clusterização a fim de diminuir seu custo computacional.**

I. INTRODUÇÃO

A utilização de métodos para garantir a segurança dos dados se mostra de fundamental importância num ambiente onde a Internet se tornou intrínseca à organização da sociedade. Técnicas em segurança computacional são propostas, desenvolvidas e avaliadas continuamente. Essas técnicas são utilizadas em diferentes etapas do processo de transmissão de dados entre dispositivos conectados em uma rede, também sendo utilizadas para garantir a integridade de dados armazenados em dispositivos locais.

Segundo dados do Centro de Estudos, Respostas e Tratamento de Incidentes no Brasil (CERT.br), os índices de incidentes de segurança reportados no decorrer dos anos tem aumentado, reforçando a necessidade da utilização de métodos para garantir a segurança de dados.

Este trabalho apresenta a proposta de um método de detecção de intrusão utilizando métodos de inteligência computacional. Para tal, o algoritmo k-*Nearest Neighbor* (kNN) foi escolhido, dada a sua eficiência na classificação de dados. É conhecido que este algoritmo possui um custo computacional elevado quando lida com grandes quantidades de dados. Portanto, técnicas de clusterização serão utilizadas para contornar esta característica.

II. SEGURANÇA COMPUTACIONAL

A. Confidencialidade, Integridade e Disponibilidade

Segundo [1], a segurança computacional é definida em cima de três pilares básicos: confidencialidade, integridade e disponibilidade.

A confidencialidade é a propriedade de garantir que as informações não serão obtidas sem autorização. Isto é, envolve a proteção dos dados, fornecendo acesso apenas às pessoas autorizadas e não permitindo que outros saibam sobre seu conteúdo.

Integridade é a propriedade de que uma informação não foi alterada por um usuário não autorizado, ou seja, os dados devem chegar ao seu destino exatamente da forma que foram enviados, sem que nenhuma modificação ocorra durante o processo de transmissão, sendo de forma acidental ou mal-intencionada.

A disponibilidade é a propriedade de a informação poder ser acessível e modificável no momento oportuno por aqueles que estejam autorizados a fazer isso.

B. Criptografia

Técnicas criptográficas permitem que dados sejam codificados de modo que, caso sejam interceptados, não posam ser entendidos. A reconstrução da mensagem ocorre apenas no destinatário, onde este deve possuir as habilidades adequadas para reconstruir a mensagem original [2]. A criptografia pode utilizar chaves simétricas e chaves assimétricas.

Criptografia de chaves simétricas é uma técnica que utiliza a mesma chave para criptografar e descriptografar uma mensagem. O emissor e o destinatário devem possuir a mesma chave para realizar tal operação e essa chave é distribuída através da criptografia de chaves assimétricas [3].

A criptografia de chaves assimétricas exige que cada usuário possua duas chaves: uma pública, utilizada pelo mundo inteiro para criptografar as mensagens que serão enviadas para este usuário, e uma chave privada, que é utilizada para descriptografar as mensagens [4].

O processo de criptografia apresenta a necessidade de autenticar a chave pública, sendo necessária a utilização de um método de autenticação para garantir a identidade do usuário da chave [4].

C. *Autenticação e Assinatura Digital*

Autenticação é uma propriedade que garante a identidade de um usuário [4] e, para garantir isso, um sistema deve ser capaz de enviar uma mensagem assinada para outra parte de modo que o receptor possa confirmar a identidade do emissor, o emissor não possa repudiar o conteúdo da mensagem e o receptor não possa criar ele mesmo a mensagem [5]. A forma utilizada para garantir essa propriedade é a utilização de assinatura digital.

Estratégias para as assinaturas digitais é a existência de uma autoridade central (AC), onde cada usuário possui uma chave secreta e a informa para essa AC, de modo que a chave torna-se conhecida apenas pelos dois. Quando um usuário quer enviar uma mensagem para outro usuário, essa mensagem é descriptografada pela AC e enviada para o outro receptor, de modo que garante a autenticidade das partes [5].

Um modo de contornar a necessidade de uma AC é a utilização de chaves públicas, onde os usuários apenas precisam conhecer as chaves públicas uns dos outros. A chave privada é mantida em segredo, de modo que as mensagens são criptografadas usando as chaves públicas e descriptografadas com as chaves privadas [5].

III. Ataques Computacionais

Segundo [6], um evento é uma mudança discreta de estado de um sistema ou dispositivo. Dentro do contexto de segurança computacional, essas mudanças de estados resultam em ações que são direcionadas a alvos específicos. Essas ações podem ser ou não autorizadas.

Um ataque é uma sequência de etapas de um evento que visam obter acesso não autorizado sobre um sistema. [7] apresenta duas formas de classificação de ataques: ativos e passivos. Ataques passivos ocorrem quando um atacante tenta descobrir ou utilizar informações do sistema sem afetar seus recursos. Em ataques ativos, o atacante modifica o fluxo de dados existentes ou cria um fluxo de dados falsos.

Ataques podem ser agrupados com base em características em comum. Para isso, [7] apresenta uma taxonomia útil para classificar ameaças à segurança, representada pela figura 1.

- Interrupção: um componente do sistema é destruído, torna-se indisponível ou inutilizável, impossibilitando a troca de informações entre sistemas;
- Interceptação: uma parte não autorizada ganha acesso a um fluxo de dados, possuindo acesso às informações trocadas;
- Modificação: uma parte não autorizada não apenas ganha acesso ao fluxo de dados do sistema, mas também passa a alterar as informações;

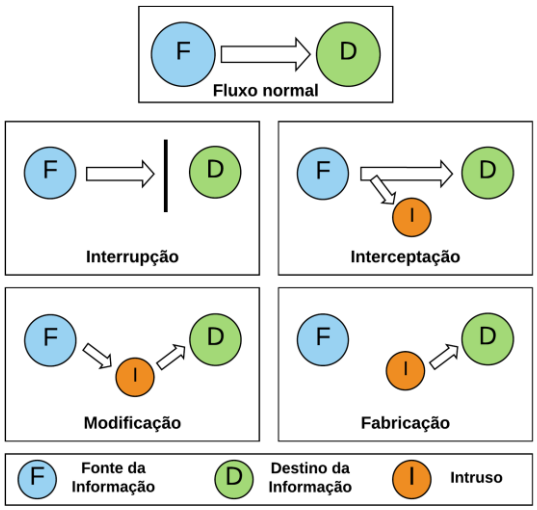

Fig 1: Taxonomia de ataques.

- Fabricação: uma parte não autorizada passa a criar novos dados no sistema.

IV. Sistemas de Detecção de Intrusão

Segundo [8], um Sistema de Detecção de intrusão (SDI) são sistemas de software ou hardwares que automatizam o processo de detecção e, além disso, possuem capacidades para impedir ataques.

Um SDI pode ser implementado na própria máquina (*Host Based*), onde monitora a atividade local, como processos, conexões de rede, chamadas de sistemas e níveis de uso de recursos locais com o objetivo de identificar ataques e acessos indevidos à própria máquina ou pode ser instalado em uma máquina que tenha visibilidade sobre o tráfego da rede a ser monitorada (*Network Based*), onde se baseia em capturar e analisar o tráfego da rede, buscando detectar assinaturas de ataques conhecidos ou padrões anormais de atividades nas máquinas da rede monitorada [9].

[8] classifica SDI em três tipos conforme sua metodologia de detecção:

- Baseados em assinatura: uma assinatura é um padrão que corresponde a um ataque ou ameaça já conhecidos. O SDI processa o evento e o compara com a base de assinaturas a fim de determinar a ocorrência de um ataque;
- Baseados em anomalias: uma anomalia é um desvio de comportamento definido como padrão para o monitoramento de atividades regulares numa redes. O SDI verifica as atividades de conexões na rede e os compara com o tráfego considerado normal para aquela rede;
- Híbrido: SDI utilizam ambas as abordagens acima para detectar padrões de ataques já conhecidos com base em assinaturas e anomalias.

Para resultados possíveis de detecção, [1] apresentam quatro tipos:

- Verdadeiro positivo: um evento ou atividade que é um ataque é classificado de forma correta;
- Verdadeiro negativo: um evento ou atividade que não é um ataque é classificado de forma correta;
- Falso positivo: um evento ou atividade que não é um ataque é detectado e classificado como ataque;
- Falso negativo: um evento ou atividade que é um ataque é detectado, porém classificado como não ataque.

V. Bases de Dados de Eventos Computacionais

Uma das bases de dados de eventos computacionais mais utilizada é a KDD Cup 99, criada pela DARPA em 1999 [10]. Essa base contém mais de 5 milhões de instâncias que representam uma conexão TCP/IP, composta por 41. Ela ainda possui 39 tipos de ataques, agrupados em quatro classes:

- *Denial of servisse* (DoS): ataque que torna alguns recursos de computação ou de memória muito ocupados para lidarem com solicitações legítimas, impedindo o acesso de outros usuários;
- *Probe*: invasor varre uma rede para coletar informações ou encontrar vulnerabilidades;
- *Remote-to-Local* (R2L): invasor envia pacotes a uma máquina pela rede e explora vulnerabilidades para obter acesso não autorizado como usuário;
- *User-to-Root* (U2R): invasor inicia o acesso a uma conta de usuário normal no sistema e é capaz de explorar vulnerabilidades para obter acesso *root* do sistema.

VI. k-*Nearest Neighbors*

O k-*Nearest Neighbor* (kNN), ou k-Vizinhos mais próximos, é um método simples de classificação baseado em instâncias, pertencentes a família de algoritmos de aprendizagem baseada em exemplos [11].

No algoritmo kNN, uma nova instância é comparada com outros exemplos já classificados presentes na base de dados e o outro exemplo mais próximo é utilizado para determinar a quase à qual a nova instância pertence. Mais de uma instância pode ser utilizada e a classificação é determinada pela classe da maioria dos k vizinhos mais próximos [12].

Métricas de classificação são usadas no algoritmo kNN e, segundo [11], o classificador entre a nova instância e um exemplo já classificado é comumente baseado na Distância Euclidiana.

O parâmetro k do algoritmo representa o número de vizinhos mais próximos num conjunto de treinamento que serão considerados para classificar uma instância [11]. A figura 2 representa o de classificação do algoritmo com k = 3.

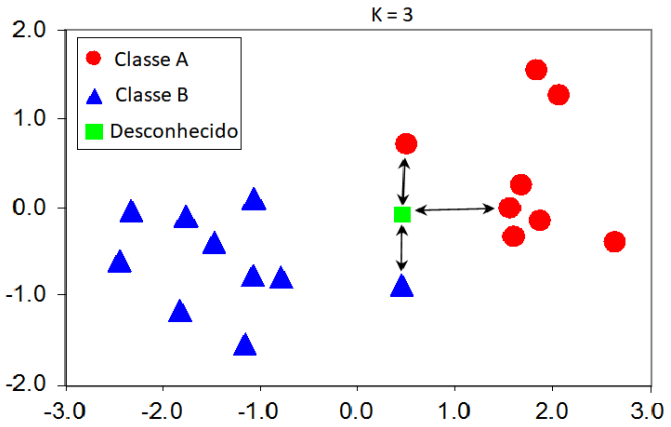

Fig. 2: Esquema de classificação do kNN com k = 3.

VII. Clustrização e k-Means

Métodos de clusterização, também conhecidos como aprendizagem por agrupamento, agem selecionando uma coleção inteira ou um agrupamento de hipóteses a combinar suas previsões baseadas em determinados atributos, visando maximizar a similaridades ou minimizar as diferenças mediante um critério escolhido [12].

Quando clusters, ao invés de apenas um classificador, são aprendidos, a saída assume a forma de um diagrama que apresenta como as instâncias se enquadram nos clusters. Cada instância é associada a um cluster [13].

O algoritmo k-Means é do tipo não supervisionado. Ele encontra similaridades entre os dados e os agrupa conforme o número de clusters passado pelo argumento k. Ele utiliza métricas de distâncias, podendo ser escolhido a Distância Euclidiana como no kNN. O processo executado pelo algoritmo é composto de quatro etapas:

- Inicialização: o algoritmo gera de forma aleatória k centroides, que serão utilizados como referências para calcular a distância entre os dados e gerar os clusters;
- Atribuição ao cluster: cálculo da distância entre todas as instâncias e os centroides. Cada instância é atribuída ao centroide mais próximo;
- Movimentação dos centroides: é calculada a média dos valores dos pontos de dados de cada cluster e o valor médio será a nova posição do centroide;
- Otimização do k-Means: as fases de Atribuição ao cluster e Movimentação dos centroides são repetidas até o cluster se tornar estático ou algum critério de parada tenha sido atingido, como por exemplo, um número de iterações máxima. O cluster se torna estático quando nenhum dos pontos de dados alteram o cluster.

Por fim, o algoritmo k-Means chega ao fim da sua execução dividindo os dados no número de clusters especificado pelo argumento k.

VIII. METODOLOGIA

Este trabalho consiste no desenvolvimento de um método de detecção de intrusão utilizando kNN clusterizado, a fim de diminuir seu custo computacional, que será comparado com outras ferramentas de detecção de intrusão utilizadas no mercado. Para tal, testes estatísticos serão feitos para determinar a eficiência do método.

Para desenvolvimento do método, serão utilizadas a linguagem Python 3.7.0, com a biblioteca SciKit-Learn 0.19.2. A biblioteca SciKit-Learn possui diversos métodos de *machine learnig* já implementados.

Um computador pessoal com especificações: processador Intel i7-7700HQ, memória RAM de 8 GB, SSD 240 GB e Windows 10 64 bits.

O Weka 3.8.2 será utilizado na fase de pré-processamento dos dados, a fim de obter a melhor configuração de atributos da base de dados.

Os testes do método utilizarão a base de dados de eventos computacionais NSL-KDD executando também com o Snort 2.9.11.1, a fim de identificar ataques e não ataques e comparar os resultados entre ambos.

Os resultados obtidos pelo método proposto serão então comparados quantitativamente com os obtidos pela ferramenta Snort por meio de testes estatísticos. Os resultados serão avaliados nos critérios de acurácia, especificidade e sensibilidade. A figura 3 ilustra a arquitetura do projeto.

IX. PRÓXIMAS ETAPAS

[1] GOODRICH, M. T.; TAMASSIA, R. Introdução à segurança de computadores. [S.l.]: Bookman, 2013.

[2] ROSS, K. W.; KUROSE, J. F. Redes de computadores e a internet. São Paulo, 2010.

[3] CARISSIMI, A. D. S.; ROCHOL, J.; GRANVILLE, L. Z. Redes de computadores. Bookman, 2009.

[4] FOROUZAN, B. A. Comunicação de dados e redes de computadores. [S.l.]: AMGH Editora, 2009.

[5] TANENBAUM, A. S.; WETHERALL, J. Redes de computadores. 5ª edição. Rio de Janeiro: Editora Campus, 2011.

[6] HOWARD, J. D.; LONGSTAFF, T. A. A common language for computer security incidents. Sandia National Laboratories, v. 10, p. 751004, 1998.

[7] STALLINGS, W. Criptografia e segurança de redes: princípios e práticas. [S.l.]: Pearson, 2014.

[8] LIAO, H.-J. et al. Intrusion detection system: A comprehensive review. Journal of Network and Computer Application, Elsevier, v. 36, p. 16-24, 2012.

[9] LAUREANO, M. A. P.; MAZIERO, C. A.; JAMHOUR, E. Detecção de intrusão em máquinas virtuais. 5º Simpósio de Segurança em Informática-SSI. São José dos Campos, p. 1-7, 2003.

[10] KENDALL, K. K. R. A database of computer attacks for the evaluation of intrusion detection systems. Tese (Doutorado) – Massachusetts Institute of Technology, 1999.

[11] PETERSON, L. E. K-nearest neighbor. Scholarpedia, v. 4, n. 2, p. 1883, 2009.

[12] NORVIG, P.; RUSSELL, S. Inteligência Artificial: Tradução da 3a Edição. [S.l.]: Elsevier Brasil, 2014. V. 1.

[13] WITTEN, I. H. et al. Data Minig: Pratical machine learning tools and techniques. [S.l.]: Morgan Kaufmann, 2016.

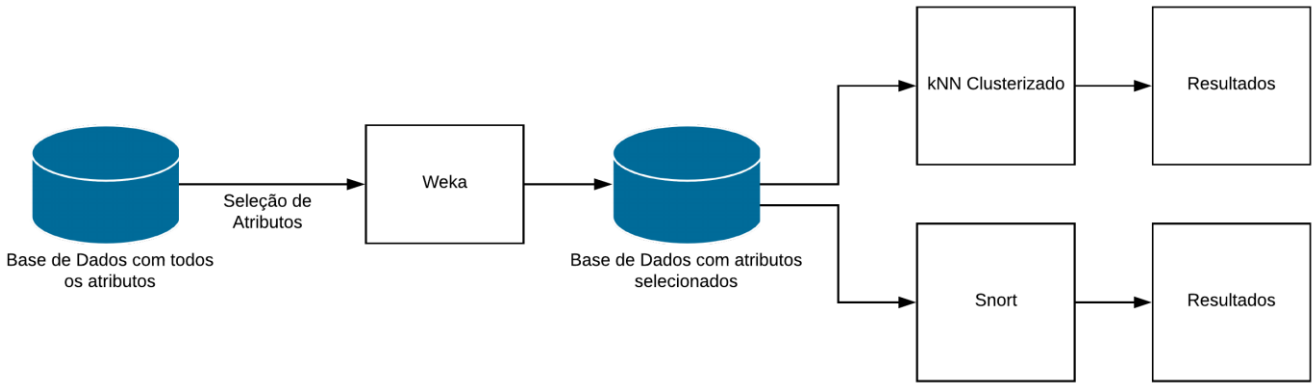

Fig. 3: Arquitetura do projeto.