Detecção de Intrusos usando Conjunto de k-NN gerado por Subespaços Aleatórios

Márcia Henke, Celso Costa, Eulanda M. dos Santos, Eduardo Souto

Instituto de Computação
Universidade Federal do Amazonas (UFAM) – Amazonas, AM – Brasil
{henke,ccosta,esouto,emsantos}@dcc.ufam.edu.br

Abstract. *Several studies have been proposed in the literature to deal with Internet anomaly detection by using machine learning techniques. Most of these works use individual classifiers such as k-NN (k-Nearest Neighbor), SVM (Support Vector Machines), Artificial Neural Networks, Decision Tree, Naive Bayes, k-means, among others. However, the literature has recently focused on applying classifier combination in order to increase detection rate. In this paper, a set of classifiers, more precisely, a set of k-NN generated through Random Subspaces Method is employed. Such an ensemble of classifiers method is compared to the hybrid technique TANN (Triangle Area based Nearest Neighbor), published recently in the literature. Results obtained using ensemble of k-NNs were superior to those obtained with TANN in terms of classification accuracy as well as false alarm reduction rate.*

Resumo. *Diversos estudos apresentam propostas para detecção de anomalia na Internet empregando técnicas de aprendizagem de máquina. A maioria desses trabalhos utiliza classificadores individuais como: k-Nearest Neighbor (k-NN), Support Vector Machines (SVM), redes neurais artificiais, árvores de decisão, Naive Bayes, k-means, dentre outros. Recentemente, porém, observa-se um interesse na literatura em aumentar a taxa de detecção de anomalia através do uso de combinação de classificadores. Este trabalho propõe o uso de conjunto de classificadores, mais especificamente conjunto de k-NNs gerados através do método subespaços aleatórios (RSM), para aumentar a taxa de detecção de anomalias em redes de computadores. O método é comparado à técnica híbrida Triangle Area based Nearest Neighbor (TANN), publicada recentemente na literatura. Os resultados alcançados pelo conjunto de k-NNs foram superiores aos obtidos com TANN, tanto em termos de aumento da precisão de classificação, quanto em termos de redução de falsos alarmes.*

1. Introdução

Atualmente a grande preocupação quando se fala em Internet é a questão segurança. Segurança na Internet tem sido alvo de muitas pesquisas no âmbito mundial, visto que a rede mundial de computadores passou, em um curto espaço de tempo, de apenas um meio de comunicação para uma poderosa ferramenta de negócios. Infelizmente, os sistemas atuais para segurança na Internet não conseguem atender a total demanda de

novos ataques, atividades maliciosas e intrusas, que se propagam na rede mundial de computadores de maneira agressiva e progressiva.

São inúmeras as vítimas de ataques originados através de atividades fraudulentas, como, vírus, *worms*, *trojan horses*, *bad applets*, *botnets*, *phishing*, *pharmimg*, mensagens eletrônicas não desejadas (*spam*), entre outras. Tais atividades podem ser consideradas uma pandemia cujas conseqüências são refletidas no crescimento dos prejuízos financeiros dos usuários da Internet [Feitosa, Souto e Sadok 2008].

Para tentar minimizar tais ameaças, diferentes abordagens de detecção de intrusão em redes de computadores foram propostas, as quais podem ser classificadas em duas categorias [Anderson 1995] [Rhodes, Mahaffey e Cannady 2000]: 1) detecção de abuso ("*misuse detection*"); e 2) detecção de anomalias ("*anomaly detection*").

Um exemplo da primeira classe de abordagens de detecção são as ferramentas antivirais baseadas em uma lista contendo "assinaturas" de vírus e *worms* conhecidos. Desta forma, ataques conhecidos são detectados com bastante rapidez e com baixa taxa erro. Por outro lado, a principal limitação dessas ferramentas é que elas não podem detectar novas formas de códigos maliciosos que não sejam compatíveis com as assinaturas existentes.

A segunda classe de detecção de intrusão, detecção de anomalias, é baseada na construção de perfis de comportamento para padrões considerados como atividade normal. Desvios da normalidade são então tratados como ameaças. Entretanto, é difícil saber o que procurar quando atividades não autorizadas sob um sistema assumem diferentes formas ou mesmo imitam atividades legítimas. Na tentativa de evitar que atividades com potencial malicioso sejam autorizadas, muitos sistemas emitem uma taxa elevada de alarmes falsos, reduzindo substancialmente sua efetividade.

A detecção de anomalias tem sido tratada na literatura através de diversas propostas de sistemas para detecção de intrusão que utilizam técnicas de aprendizagem de máquina, tais como: redes neurais artificiais [Souza e Monteiro 2009] [Xia, Yang e Li 2010], k-*means* [Tian e Jianwen 2009], k-NN [Tsai e Lin 2010] e SVM (*Support Vector Machines*) [Xiao *et al.* 2007], entre outras. Essas técnicas têm sido usadas como classificadores individuais cuja função é detectar eventos inesperados que podem indicar possíveis ataques em redes de computadores.

Além da aplicação de classificadores individuais, técnicas híbridas e combinação de classificadores têm recentemente atraído a atenção dos pesquisadores em diversas áreas de aplicação, inclusive em segurança de redes para detecção de intrusão. Técnicas híbridas são cooperações de dois ou mais classificadores, como por exemplo, a abordagem TANN (*Triangle Area based Nearest Neighbor*) [Tsai e Lin 2010]. TANN é um método para detecção de anomalias que utiliza em um primeiro nível a técnica de agrupamento k-*means* para transformar dados primários, ou seja, o espaço de características original, em novos dados que servem de entrada para outro classificador. No segundo nível, o classificador k-NN é utilizado para definir a classificação final das amostras.

A combinação de classificadores (*classifier ensembles*) é baseada na hipótese de que combinar a decisão de um conjunto de classificadores pode aumentar a taxa de detecção correta, superando o desempenho de classificadores individuais. Os métodos

mais comuns para a geração de conjuntos de classificadores são *bagging* [Breiman 1996] e subespaços aleatórios [Ho 1998]. Uma vez criados, os membros do conjunto devem ter suas opiniões combinadas em uma única decisão. A regra do voto majoritário é a função mais popular de combinação de conjuntos de classificadores.

Em [Xiao *et al.* 2007] é proposta a utilização de um conjunto de SVMs para detecção de anomalias. Os membros do conjunto não são gerados nem por *bagging* e nem por subespaço aleatórios, e sim, através de uma estratégia que envolve uma adaptação de ambos os métodos. Essa estratégia envolve processos de seleção de características e de amostras, aplicados aos dados de entrada para proporcionar diversidade aos membros do conjunto.

Este artigo propõe o emprego da combinação de classificadores para melhorar o processo de detecção de anomalias em redes de computadores. Trata-se de um conjunto de k-NNs gerados por subespaços aleatórios. Além disso, este trabalho também fornece um estudo comparativo de métodos de classificação com o objetivo de mostrar a superioridade do conjunto de classificadores sobre outras técnicas mais clássicas de classificação. Os métodos comparados são: o conjunto de k-NNs gerados com o método de subespaços aleatórios (*Random Subspace Method - RSM*) e o método híbrido proposto em [Tsai e Lin 2010].

Fora isso, este trabalho também apresenta uma alteração no método de classificação proposto por [Tsai e Lin 2010]. Esses autores propõem o uso de classificadores híbridos. Eles utilizam k-means e k-NN, conforme mencionado anteriormente. Porém, a literatura mostra que SVM apresenta normalmente desempenho superior ao k-NN. Devido a esse fato, neste trabalho k-NN será substituído pelo classificador SVM. Os resultados obtidos demonstram que a troca de classificador ocasiona um aumento na taxa de classificação do método TANN. Porém, não supera a combinação de classificadores proposta neste trabalho. Os experimentos foram realizados na base de dados KDD Cup 99 [KDD Cup 1999], que é uma base de dados muito utilizada para a experimentação de soluções de detecção de intrusão de rede.

O restante deste artigo está organizado como segue. Na Seção 2 é apresentada uma visão geral dos trabalhos mais recentes na área de detecção de intrusos em redes de computadores que utilizam técnicas de aprendizagem de máquina. Na Seção 3 são descritas as técnicas de aprendizagem de máquina comparadas neste trabalho. Na Seção 4, o protocolo experimental e os resultados obtidos são apresentados. Finalizando na Seção 5, com conclusões e trabalhos futuros.

2. Trabalhos Relacionados

Diversas abordagens têm sido propostas para sistemas de detecção de intrusão. Destaca-se na literatura da área, o fato de muitos desses sistemas terem sido desenvolvidos com base na utilização de diferentes técnicas de aprendizagem de máquina e mineração de dados. [Tsai *et al.*, 2009] apresentam uma revisão de literatura que investiga a aplicação de técnicas de aprendizagem de máquina em problemas de detecção de intrusão em trabalhos publicados entre os anos 2000 e 2007. De acordo com os autores, os métodos mais bem sucedidos são os classificadores híbridos, seguidos de métodos do tipo conjunto de classificadores e por último, classificadores individuais.

Ainda segundo [Tsai *et al.* 2009], dentre os classificadores simples, os que têm sido referenciados e testados de forma mais ampla são os métodos SVM e k-NN, ambos apresentando elevado desempenho de classificação com relação à precisão, falso alarme e taxa de detecção. Com relação a bases de dados investigadas, os referidos autores destacam a base KDD Cup 99 [KDD Cup 1999], utilizada nos experimentos deste artigo, como a base mais usada dentre as três bases mais citadas, incluindo DARPA 1998 e DARPA 1999 [Darpa 1999]. Os trabalhos relacionados nesta seção são organizados a partir da divisão definida em [Tsai *et al.* 2009].

2.1. Classificadores Híbridos

[Tsai e Lin, 2010] propõem o método TANN que transforma o espaço de características original em um novo espaço de características. Inicialmente, *k-means* é utilizado para agrupar as amostras da base de dados. O resultado dessa etapa são os centróides (ponto central) de cada grupo formado pelas amostras. Em seguida, os centróides, juntamente com cada amostra da base de dados, são projetados no espaço de característica. A área do triângulo formado entre a amostra e os centróides, combinados em pares, é calculada e então, é usada como entrada para o classificador k-NN que definirá a classe da amostra. A base de dados investigada foi a KDD Cup 1999.

[Mafra *et al.* 2008] propõem um sistema multicamadas, chamado POLVO – IIDS, que utiliza redes neurais de *Kohonen* e SVM. A primeira camada analisa o tráfego da rede e a classifica em quatro categorias: DoS (*Denial of Service*), *Worm*, *Scan* e R2L (*Remote to Local*)/Normal. A segunda camada é responsável pela detecção de intrusão propriamente dita.

[Lee *et al.* 2007] propõem uma abordagem híbrida para sistemas de detecção de intrusos em redes de tempo real. É feita uma seleção de características usando o método Florestas Aleatórias (*Random Forest)*, que elimina as características irrelevantes, enquanto que *Manimax Probability Machine* (MPM) é aplicado como classificador. A base de dados usada também foi a KDD Cup 1999.

2.2. Conjunto de Classificadores

[Xiao *et al*., 2007] utilizam um conjunto de três SVMs. O primeiro classificador é treinado com os dados originais, sem alterações, como normalmente é feito com classificadores individuais. O segundo classificador é treinado com os dados originais submetidos a um processo de seleção de características, ou seja, apenas um grupo escolhido das características originais é utilizado durante o treinamento. Por fim, o último SVM é treinado com apenas uma parte dos dados de entrada, isto é, é feita uma seleção de amostras. A combinação da decisão do conjunto é obtida através de uma função de voto ponderado. A base de dados utilizada nos experimentos foi a DARPA.

2.3. Classificadores Individuais

[Chimphlee *et al.* 2006] aplicam a idéia de *Fuzzy Rough C-means* (FRCM), método individual baseado em agrupamento. FRCM integra a vantagem do conjunto de teoria *fuzzy* e a técnica *k-means* para melhorar a tarefa de detecção de intrusão. Os resultados experimentais obtidos na base de dados KDD Cup 99, mostraram que o método supera os resultados obtidos com *k-means* unicamente.

[Lião e Vemuri, 2002] usaram em sua abordagem k-NN, para classificar comportamento de programas como normal ou intrusivo. Com o classificador k-NN, as

freqüências de chamadas ao sistema são usadas para descrever o comportamento do programa. Técnicas de categorização de texto são adotadas para converter cada execução de programa para um vetor e calcular a similaridade entre duas atividades de programas. Utilizando base de dados DARPA 1998 foi demonstrado que o classificador k-NN pode efetivamente detectar ataques intrusivos e atingir as mais baixas taxas de falso positivo.

[Chen et al., 2005] realizaram um estudo comparativo entre SVM e redes neurais artificiais. Os autores concluíram que SVM atinge um melhor desempenho em relação às redes neurais. O objetivo de usar redes neurais e SVM para detecção de ataques foi desenvolver uma capacidade de generalização de dados de treino limitado. Esses experimentos foram baseados na base DARPA 1998.

Todos os trabalhos mencionados nesta seção influenciaram de alguma forma os experimentos apresentados neste artigo, pois se configuram como um guia indicando aspectos promissores e relevantes na área de detecção de intrusos utilizando aprendizagem de máquina. Porém dentre estes os dois principais artigos são os artigos de [Ho 1998] e [Tsai e Lin, 2010], que tratam a dimensionalidade de espaço de características de uma forma diferenciada com problemas de bases muito grandes. Como mencionado anteriormente [Ho 1998] reduz essa dimensionalidade com um subespaço aleatório de características (de 41 características para 20) procurando combinar suas diversas visões do problema e finalizando com o voto majoritário. Já [Tsai e Lin, 2010] se utiliza da redução desse espaço de características reduzindo as 41 características, com a proposta de uma área triangular, para 10 características. Desta forma o presente artigo apresenta as seguintes contribuições:

- Substituição do classificador k-NN, no conjunto de classificadores híbridos proposto por [Tsai e Lin, 2010], pelo classificador SVM, conforme indicado na literatura como um classificador de ótimo desempenho para detecção de intrusão [Tsai et al., 2009];

- Aplicação de um método proposto para redução de dimensionalidade de características para o problema de reconhecimento de dígitos [Ho 1998], em um problema de detecção de intrusão.

Na próxima seção, os métodos comparados neste artigo são descritos com mais detalhes.

3. Abordagens Aplicadas

Conforme mencionado na seção anterior, a literatura indica que a combinação de classificadores produz resultados superiores aos resultados obtidos por classificadores individuais. Por essa razão, duas abordagens que usam combinação de classificadores são comparadas neste trabalho. Esta seção apresenta essas abordagens. A primeira propõe a utilização de conjunto de classificadores k-NN treinados em diferentes subespaços aleatórios, enquanto que a segunda propõe uma alteração no classificador híbrido proposto por [Tsai e Lin, 2010].

3.1. Conjunto de KNNs gerado por subespaços aleatórios

O método de subespaços aleatórios (RSM) para classificação k-NN foi proposto em [Ho, 1998]. É considerada uma abordagem baseada em seleção de características, pois trabalha da seguinte forma.

Dada uma base de dados *D*, cada amostra x que compõe a base é representada por características que são organizadas em um vetor com l dimensões, $x = [x_1, x_2, ..., x_l] \in R^l$, em que R^l é chamado espaço de características e cada eixo corresponde a uma característica original dos dados. RSM escolhe aleatoriamente n subespaços diferentes, com dimensão j cada, a partir do espaço de características original R^l, em que $j < l$, e representa toda a base de dados D a partir de cada subespaço escolhido. Então, cada nova representação da base D_i é utilizada para treinar um classificador individual c_i. A Figura 1 apresenta uma visão geral do funcionamento de RSM.

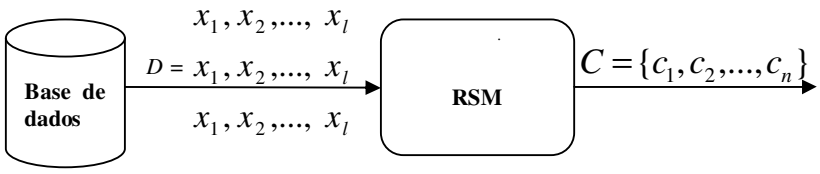

Figura 1. Visão geral do método RSM. São escolhidos aleatoriamente *j* características dentre as *l* características originais, sendo que *j<l*. A base de dados *D* é representada por cada subespaço escolhido, sendo que cada representação é utilizada para o treinamento de um classificador c_i.

[Ho,1995] propôs o método RSM para criar diversidade de opiniões entre os classificadores treinados com os diferentes subespaços e também como forma de minimizar os problemas ocasionados pela alta dimensão dos dados de entrada. Inicialmente o método foi testado usando-se classificadores do tipo Árvores de Decisão. Posteriormente, o método foi aplicado com classificadores do tipo k-NN [Ho, 1998]. Segundo a autora, um conjunto de k-NNs gerado por RSM tem a vantagem de prover elevada taxa de classificação ao mesmo tempo em que reduz a dimensão de entrada dos dados. Este último é um dos maiores problemas com o método de classificação k-NN.

Depois de treinados, as decisões dos n classificadores gerados por RSM são combinadas pela regra do voto majoritário. Essa função de combinação é a mais simples e a mais popular na literatura de conjunto de classificadores. A definição apresentada na equação 1 é também chamada de *Plurality vote* [Kuncheva, 2004]. Considerando o conjunto de n classificadores, y_i como o rótulo da classe de saída do i-ézimo classificador, e o problema da classificação com o seguinte conjunto de rótulos $\Omega = \{w_1, w_2, ...w_c\}$, voto majoritário para a amostra x é calculada como:

$$vm(x) = \max_{k=1}^{c} \sum_{i=1}^{n} y_{i,k} \qquad (1)$$

Quando ocorre um empate em números de votos, a classe vencedora é escolhida aleatoriamente ou uma estratégia de rejeição deve ser aplicada.

Diante das razões e definições descritas acima, o método de aprendizagem de máquina baseado em conjunto de classificadores é aplicado neste artigo ao problema de detecção de intrusão através do RSM. Os classificadores membros foram k-NNs, sendo utilizado o voto majoritário para combinar as decisões dos k-NNs.

3.2. Classificadores Híbridos

Esta abordagem proposta por [Tsai e Lin 2010] é dividida em três estágios: (1) extração de centróides das amostras; (2) formação de uma nova base de dados; e (3) treino e teste do classificador. No primeiro estágio, todas as amostras da base de dados são projetadas no espaço de características original para que o algoritmo não-supervisionado k-*means* seja aplicado a fim de agrupar as amostras de acordo com o número de classes do problema, e calcular os centróides de cada grupo. No segundo estágio, um novo espaço de características é gerado através do cálculo de área triangular (descrito com detalhes a seguir). Por fim, k-NN é treinado e usado para classificar as amostras desconhecidas.

Para calcular a área triangular, são considerados três pontos de dados: dois centróides obtidos por k-*means* e uma amostra da base de dados. Os autores utilizaram a mesma base de dados investigada neste artigo, isto é KDD Cup 99, que é composta por amostras de cinco classes, das quais uma é do tipo tráfego normal e as quatro restantes do tipo ataque. Por serem cinco classes, k-*means* é direcionado a encontrar cinco centróides. A Figura 2 mostra um exemplo dos cinco grupos e seus respectivos centróides (A, B, C, D, E) e um ponto de dado (X_i).

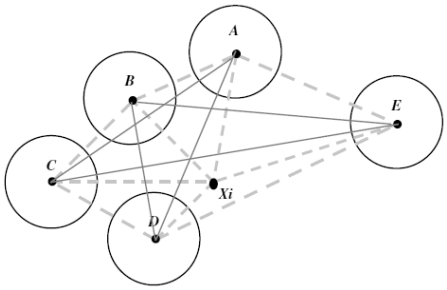

Figura 2. Formação da área triangular. Fonte: [Tsai e Lin 2010].

Portanto, para a base KDD-Cup, dez áreas são obtidas para formar um novo vetor de caracteríticas para o ponto de dado (X_i):

$\Delta X_i AB$, $\Delta X_i AC$, $\Delta X_i AD$, $\Delta X_i AE$, $\Delta X_i BC$, $\Delta X_i BD$, $\Delta X_i BE$, $\Delta X_i CD$, $\Delta X_i CE$ e $\Delta X_i DE$.

Em seguida é calculado o perímetro de cada triângulo através da distância Euclidiana, determinando o ponto de dado X_i ($i = 1,..., m$, onde m é o total do número de amostras). Sendo a distância Euclidiana entre dois pontos $A = (a_1, a_2, ..., a_n)$ e $B = (b_1, b_2, ..., b_n)$ no espaço de n características, definida pela equação 2.

$$\overline{AB} = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + ... + (a_n - b_n)^2} = \sqrt{\sum (a_i - b_i)^2} \quad (2)$$

O perímetro do triângulo $\Delta X_i AB$ é definido como $G = a + b + c$, onde $a = \overline{AB}$, $b = \overline{BX_i}$ e $c = \overline{AX_i}$, isto é, a distância entre A e B, B e X_i, e A e X_i, respectivamente.

Depois de obter o perímetro dos triângulos para cada amostra, a fórmula de Heron é calculada. A equação 3 exibe a fórmula de Heron.

$$T = \sqrt{S(S-a)(S-b)(S-c)} \tag{3}$$

Onde o $S = \frac{G}{2}$ é o semi perímetro de $\Delta X_i AB$.

Portanto, 10 triângulos, T_1-T_{10} são identificados para cada X_i e são usados para formar os novos dados. Esses novos dados são então usados para treinar e testar o classificador k-NN.

Porém, a literatura de detecção de intrusão mostra que SVM é um classificador que alcança taxas de detecção melhores quando comparado a k-NN [Tsai *et al.* 2009]. Devido a esse fato, este artigo propõe uma modificação no método TANN. A modificação sugerida ocorre no terceiro estágio do método, isto é, na classificação final. A proposta envolve a troca de k-NN por SVM. Portanto, a detecção dos centróides e o cálculo da área triangular permanecem inalterados. A nova representação dos dados será usada para o treinamento de um classificador do tipo SVM.

Na próxima seção são apresentados os resultados obtidos com o conjunto de k-NN gerado por RSM, TANN original (com o classificador k-NN) e TANN modificado (com classificador SVM).

4. Experimentos

Esta seção descreve os experimentos realizados para avaliar as abordagens investigadas. Essa descrição apresenta a composição da base de dados, as métricas utilizadas e sua relevância ao problema de detecção de intrusão.

4.1. Base de dados

Os métodos investigados usam "10% do KDD-Cup 99" e "*Corrected KDD-Cup* (*Test*)" [KDD 1999], como bases de treino/teste e validação, respectivamente, exatamente como na proposta de [Tsai e Lin, 2010] . Esses dois conjuntos de dados descrevem a conexão em uma rede de trabalho representada por um vetor de 41 características, distribuídas da seguinte forma: 9 características do tipo intrínsecas, 13 do tipo conteúdo e as restantes do tipo de tráfego. Cada padrão do conjunto de dados é rotulado como pertencente a uma de cinco classes, das quais uma é do tipo tráfego normal e as quatro restantes do tipo ataque como segue:

i) Probing – vigilância e sondagem de sistema;

ii) DoS (*Denial of Service*) – ataques de negação de serviço;

*iii)*R2L (*Remote to Local*) – acesso não autorizado de uma máquina remota;

*iv)*U2L (*User to Root*) – acesso não autorizado a privilégio de super usuário;

Em todos os experimentos (classificadores híbridos e conjunto de classificadores) a base de dados foi dividida em 40% para treino e 60% para teste. Essa estratégia é

conhecida na literatura de aprendizagem de máquina como *holdout validation* [Kohavi 1995]. A base foi dividida em bases de treino e de teste por ser uma base com muitas amostras, conforme Tabela1. Segundo a literatura, bases que contêm muitas amostras são bastante representativas, não havendo necessidade de serem tratadas através de validação cruzada.

A tabela 1 demonstra a distribuição das bases usadas nos experimentos em treino/teste e validação.

Tabela 1. Distribuição do Conjunto de Dados para Treino/Teste e Validação

Classes	Conjunto de Dados Treino/Teste		Conjunto de Dados para Validação	
	Qta. Amostras por Classe	(%)	Qta. Amostras por Classe	(%)
Normal	92.277	19,68%	60.593	19,40%
Probe	4.107	0,83%	4.166	1,33%
DoS	391.458	79,24%	231.455	74,40%
U2R	52	0,01%	88	0,03%
R2L	1.126	0,23%	14.727	4,73%
Total	494.020	100%	311.029	100%

4.2. Medidas de desempenho

As medidas de desempenho adotadas neste artigo seguem o padrão de métricas encontrado na literatura para detecção de anomalia. São medidas que podem ser calculadas pela matriz de confusão mostrada na Tabela 2, considerando as seguintes legendas: TP (Verdadeiro Positivo) - número de ataques devidamente classificados como ataque; FP (Falso Positivo) - número de tráfego normal classificado como ataque; FN (Falso Negativo) - número de ataques classificados como normal e TN (Verdadeiro Negativo) - número de tráfego normal devidamente classificado como normal.

Tabela 2. Matriz de confusão utilizada como base para o cálculo das medidas de desempenho.

Atual	Previsto	
	Normal	Intrusão
Normal	TN	FP
Intrusão	FN	TP

As métricas utilizadas para comparar os métodos de classificação investigados são taxa de precisão, de detecção e de falso alarme. Essas medidas podem ser obtidas por:

$$\text{Precisão} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Taxa de detecção} = \frac{TP}{TP + FN}$$

$$\text{Alarme Falso} = \frac{FP}{FP + TN}$$

4.3. Resultados

Os resultados obtidos por cada técnica são apresentados separadamente, através de um quadro comparativo e de gráficos, para que o desempenho geral dos métodos investigados seja comparado, sobre a base de teste.

4.3.1. Classificadores Híbridos

Conforme mencionado na seção 3.2, o método TANN foi replicado neste artigo de acordo com a descrição apresentada em [Tsai e Lin 2010]. *K-means* foi aplicado para agrupar os dados originais, posteriormente, as áreas triangulares foram calculadas e utilizadas como entrada para o treinamento do classificador k-NN. O único parâmetro que deve ser definido para k-NN é o número *k* de vizinhos mais próximos. O valor desse parâmetro foi o mesmo atribuído pelos autores do método, isto é, *k=17*. A Tabela 3 exibe a matriz de confusão obtida com este método.

Tabela 3. Matriz de confusão produzida pelo método TANN original (kNN)

	Normal	Ataque
Normal	58.038	328
Ataque	618	237.428

As taxas de detecção, falso alarme e precisão obtidos com TANN original foram: 99,74%, 0,56% e 99,66%, respectivamente.

Conforme destacado na seção 3.2, neste artigo ocorre a troca do classificador k-NN pelo classificador SVM, que é considerado um método que apresenta desempenho superior à k-NN. Essa alteração na abordagem TANN é chamada aqui de TANN modificado.

Dois parâmetros iniciais precisam ser definidos pelo usuário para SVM, o tipo de função *kernel* e o parâmetro de regularização *C*. Dependendo da função *kernel* escolhida, outros parâmetros devem ser definidos, como por exemplo, o grau do polinômio para o *kernel* polinomial. Neste artigo, a base de validação foi utilizada para a definição desses parâmetros. Os melhores resultados foram obtidos com o *kernel* RBF (*Radial Basis Function*), e fator de penalização $C=1$ e $\gamma=0,5$. A tabela 4 exibe a matriz de confusão obtida com aplicação de TANN modificado.

Tabela 4. Matriz de confusão produzida pelo método TANN modificado (SVM)

	Normal	Ataque
Normal	58.078	288
Ataque	321	237.725

Com relação à taxa de detecção, falso alarme e precisão para TANN modificado, os valores obtidos foram: 99,86%, 0,49% e 99,79%, respectivamente.

4.3.2. k-NNs gerados por RSM

Dois parâmetros devem ser definidos para que RSM seja aplicado: dimensão dos subespaços aleatórios e quantidade de membros do conjunto de classificadores. Segundo a autora do método RSM [Ho 1998], o tamanho recomendado para os subespaços aleatórios deve ser aproximadamente igual à metade da quantidade de características originais. Portanto, considerando que a base investigada neste artigo contém

originalmente 41 características, 20 características foram escolhidas aleatoriamente para compor cada subespaço. Em termos de quantidade de classificadores, o conjunto criado para os experimentos é composto por 15 k-NNs. Essa escolha foi baseada nos resultados obtidos por [Ho 1998] que mostraram que normalmente não há necessidade de utilização de muitos classificadores membros do conjunto.

Como os classificadores membros do conjunto são k-NNs, o valor de *k* também precisou ser definido. O valor utilizado nos experimentos foi *k=1*, uma vez que a literatura [Ho 1998] mostra que conjuntos de k-NNs gerados por RSM apresentam normalmente elevada taxa de precisão quando *k=1*, embora a escolha desse parâmetro dependa muito do problema de aplicação. A Tabela 5 exibe a matriz de confusão obtida pelo conjunto de 15 kNNs gerados por RSM.

Tabela 5. Matriz de confusão produzida pelo método RSM/kNN

	Normal	Ataque
Normal	58.355	11
Ataque	69	237.977

A taxa de precisão, taxa de detecção e falso alarme da combinação de classificadores são 99,97%, 99,97% e 0,019%, respectivamente.

4.3.3. Resultados comparativos

A Tabela 6 compara os resultados obtidos pelos três métodos investigados neste artigo. É importante destacar que esses valores foram calculados na base de teste. Os resultados indicam que o método RSM/k-NN apresentou melhor desempenho no problema de detecção de intrusão na base KDD Cup. O RSM/k-NN produziu maior taxa de precisão e de detecção, e ao mesmo tempo menor taxa de falso alarme.

A Figura 3 compara os métodos RSM/k-NN, híbrido original e híbrido modificado em termos de taxa de falsos alarmes. A Figura 4 mostra a precisão média por classe, incluindo a classe normal e as quatro diferentes classes de ataque, obtida por cada método. Observando-se que as classes R2L e U2R, não foram generalizadas pelo classificador, devido a pouca representatividade dessas classes no conjunto de dados exibido na tabela 1.

Tabela 6. Comparação dos Resultados Obtidos

	Classificador Híbrido		Conjunto de Classificadores
	TANN modificado (SVM)	TANN original (k-NN)	RSM/k-NN
Taxa Detecção	99,865	99,740	99,971
Falso Alarme	0,493	0,561	0,0188
Precisão	99,794	99,68	99,973

Os resultados também mostram que a modificação proposta neste trabalho para a estratégia TANN produziu melhores resultados quando comparada à TANN original. Esses resultados confirmam a literatura na área de detecção de intrusão que mostra que SVM é um classificador que alcança taxas de detecção melhores quando comparado a k-NN [Tsai *et al.*, 2009]. Porém, o método TANN modificado não superou o conjunto de k-NNs gerados por RSM.

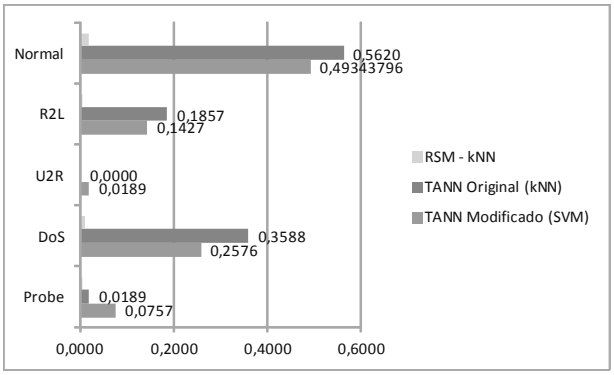

Figura 3. Comparação entre os métodos investigados em relação à taxa de falsos alarmes.

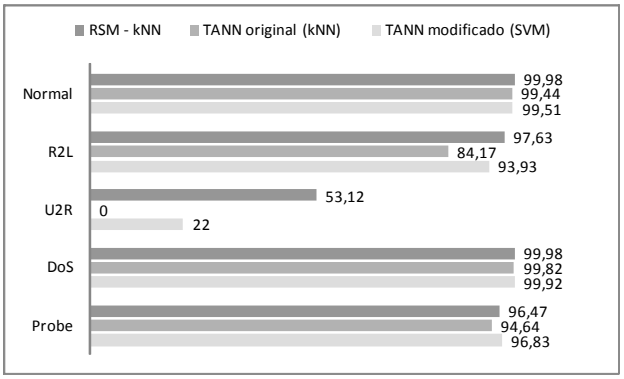

Figura 4. Comparação entre os métodos investigados em relação à precisão média por classe.

A principal tarefa de um sistema de classificação de intrusão é filtrar potenciais ataques e permitir acesso a uma conexão normal. Logo, a taxa de detecção correta de conexões, tanto como ataques quanto como normais, deve ser elevada. Como conseqüência, a taxa de falsos alarmes deve ser minimizada no intuito de aumentar a efetividade dos sistemas de detecção de anomalias. Portanto, os resultados obtidos neste artigo mostram a superioridade de RSM/k-NN sobre as demais técnicas investigadas.

Entretanto, é importante destacar que os resultados apresentados neste trabalho não têm o objetivo de sanar todas as lacunas existentes em um problema de detecção de intrusão. O objetivo é contribuir com pesquisas e experimentos que auxiliem a comunidade de pesquisadores na formulação de melhores soluções.

5. Conclusões e Trabalhos Futuros

Este artigo apresentou os resultados de um estudo experimental envolvendo a aplicação do método de subespaço aleatório na geração de um conjunto de classificadores do tipo k-NN aplicado ao problema de detecção de anomalias em redes de computadores. O método foi comparado à estratégia híbrida TANN, e à uma versão modificada de TANN, proposta neste trabalho para melhorar a taxa de classificação da estratégia original.

Embora o método híbrido modificado tenha superado o método original em termos de taxa de detecção correta e de falsos alarmes, a combinação de conjuntos de k-NNs (RSM/k-NN) atingiu um desempenho superior a ambos os métodos de classificação

híbrida. É também importante destacar a redução na taxa de falsos alarmes obtida por RSM/k-NN. Esse fato indica que conjuntos de classificadores podem ser ferramentas fundamentais no desenvolvimento de soluções efetivas para a detecção de anomalias na Internet.

Para trabalhos futuros algumas questões podem ser consideradas. Primeiro, seria interessante avaliar a proposta deste artigo em uma base de dados de ataques recentes dos tipos *phishing*, *cross-site scripting*, *spam*, entre outros. Segundo, demonstrar o desempenho de conjuntos de SVMs gerados por subespaços aleatórios. Além disso, outros métodos de geração de conjuntos de classificadores, como *bagging* e conjuntos heterogêneos poderiam ser investigados.

Referências

Anderson, J. (1995). An introduction to neural networks. Cambridge: MIT Press.

Breiman, L. (1996). Bagging Predictors. Machine Learning, 1996, volume 24 (2), 123-140.

Chen W., Hsu S., Shen H., (2005). Application of SVM and ANN for intrusion detection. In: Computer & Operations Research, Volume 32, Issue 10, October 2005, Pages 2617-2634

Chimphlee W., Abdullah A. H.,Sap M. N., Srinoy S., Chimphlee S., (2006). Anomaly-Based Intrusion Detection using Fuzzy Rough Clustering. In: ICHIT '06 Proceedings of the 2006 International Conference on Hybrid Information Technology - Volume 01

DARPA Intrusion Detection Data Sets 1999. Cyber Systems e Technology. http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/data/index.html

Feitosa, E. L. ; Souto, E. ; Sadok, D. (2008) . Tráfego Internet não Desejado: Conceitos, Caracterização e Soluções. In: SBC. (Org.). Livro-Texto de Minicurso do VIII Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais. Porto Alegre: UFRGS, 2008, v. 1, p. 17-30.

Ho T. K., (1995). Random Decision Forests. Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on

Ho T. K., (1998). Nearest Neighbors in Random Subspaces. Advances in Pattern Recognition. Lecture Notes in Computer Science, 1998, volume 1451/1998, 640-648.

Issariyapat C., Fukuda K., (2009). Anomaly detection in IP networks with principal component analysis. Proceedings of the 9th international conference on Communications and information technologies 1229-1234.

KDD Cup 1999 Dataset, UCI KDD repository, http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

Kleinberg, E.M., (1990). Stochastic discrimination. Annals of Mathematic and Artificial Intelligence, 1 (1990) 207-239.

Kleinberg, E.M., (1996). An overtraining-resistant stochastic modeling method for pattern recognition. Annals of Statistics, 4, 6 (1996) 2319-2349.

Kohavi R., (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. Appear in the International Joint Conference on Artificial Intelligence (IJCAI).

Kuncheva L.I., Combining Pattern Classifiers: Methods and Algorithms. John Wiley & Sons, LTD, USA, 2004.

Lee M. S., Kim S. D. e Park S. J. (2007), A Hybrid Approach for Real-Time Network Intrusion Detection Systems. International Conference on Computational Intelligence and Security.

Liao Y. and Vemuri V. R., (2002). Use of K-Nearest Neighbor classifier for intrusion detection. In: Computer & Security, Volume 21, Issue 5, 1 October 2002, Pages 439-448

Mafra M. P., Fraga S. J., Moll V., Santin O. A (2008), POLVO-IIDS: Um Sistema de Detecção de Intrusão Inteligente Baseado em Anomalias. VIII Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais.

Nguyen T.T.T. e Armitage G. (2007), A Survey of Techniques for Internet Traffic Classification using Machine Learning. Centre for Advanced Internet Architectures. Swinburne University of Technology, Melbourne, Australia. IEEE Communication Surveys and Tutorials.

Rhodes B., Mahaffey J. e Cannady J. (2000). Multiple self-organizing maps for intrusion detection. In Paper presented at the proceedings of the 23rd national information systems security conference. Baltimore, MD.

Souza E. P. e Monteiro J. A. S (2009), Estudo Sobre Sistema de Detecção de Intrusão por Anomalias, uma Abordagem Utilizando Redes Neurais. XIV Workshop de Gerência e Operação de Redes e Serviços - WGRS. Sociedade Brasileira de Redes de Computadores – SBRC.

Tian L. e Jianwen W., (2009). Research on Network Intrusion Detection System Based on Improved K-means Clustering Algorithm. Internacional Forum on Computer Science – Technology and Applications. IEEE Computer Science.

Tsai C., Hsu Y., Lin C., Lin W. (2009). Intrusion detection by machine learning: A review. Expert Systems with Applications 36 11004-12000.

Tsai C., Lin C. (2010). A triangle area based nearest neighbors approach to intrusion detection. Pattern Recognition 43 222-229.

Xia, D. X., Yang, S. H. e Li, C. G., (2010). Intrusion detection system based on principal component analysis and grey neural networks. The 2nd International Conference on Networks Security Wireless Communications and Trusted Computing 142-145.

Xiao H., Hong F., Zhang Z. e Liao J., (2007). Intrusion Detection Using Ensemble of SVM Classifier. Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FKSD 2007). IEEE Computer Society.