2 Revisão de Literatura

Um trabalho feito na Information Systems Security Bureau (**??**), fez um comparativo de técnicas publicas entre os anos de 2010 e 2014 para a detecção de intrusão.

A pesquisa indica que existe um pequeno crescimento das técnicas de aprendizado de maquina (ML) e inteligencia artificial (AI) comparadas com outras técnicas que não são informadas, mostrando as técnicas como uma sub-área importante e com uma forte tendencia, concluindo ML contribui como a principal área da AI utilizada para a detecção de intrusão.

Outro comparativo foi com o numero de publicação entre 2010 e 2014 das diferentes algoritmos de inteligência artificial e aprendizado de maquina como Redes Neurais Artificiais, Logica Fuzzy, Algoritmo Genético, Arvore de Decisões, etc. A pesquisa mostrar que algoritmos com base em redes neurais artificiais são os algoritmos de inteligência artificial mais popular entre as publicações, mesmo com uma pequena queda em seu uso no ultimo ano. Algoritmo Genético, K-vizinhos mais próximos e arvore de decisão ficam muito próximos, quase empatando no segundo lugar.

O artigo conclui que inteligência artificial desempenha um papel substancial no estudo de detecção de intrusão, redes neurais artificiais são as mais populares, e que a pesquisa nessa área ainda é muito necessária, ha muitos resultados promissores nesses algoritmos, especialmente em abordagens hibridas, nos quais se utiliza a combinação entre técnicas diferentes.

No trabalho de Marley (**??**), é dada uma visão geral sobre Redes Neurais Artificiais(RNA), explicando que a expressão "rede neural" é a tentativa de representar a capacidade que cérebro humano possui de reconhecer, associar e generalizar padrões, sendo as principais áreas de atuação para a classificação de padrões e previsão.

A modelagem de uma rede neural consiste em três etapas: (1) Treinamento e Aprendizado, obtido pelo ambiente gerador dos dados, (2) Associação, obtido pelo reconhecimento de padrões distintos e (3) Generalização, relacionado a capacidade da rede de reconhecer com sucesso o ambiente que origina os dados, não propriamente os dados utilizados no treinamento.

O conhecimento é passado por um algoritmo de treinamento e aprendizado, este é transformado e armazenado nas conexões. O aprendizado é resultado de muitas apresentações de conjuntos de exemplos de treinamento.

O treinamento pode ser dado de duas maneiras: (1)batelada ou ciclos, onde a atualização somente acontece depois da apresentação de todos os pesos e (2) padrão a

padrão ou incremental, onde a atualização é feita após a apresentação de cada novo padrão. O procedimento de treinamento pode ser classificado em dois tipos: supervisionado e não supervisionado.

Muitas abordagens de RNA tem sido implementadas e testadas para IDS, um sistema proposto por Ryan (**??**), analisou o comportamento de usuários individuais. O padrão de entrada foi então combinado com os perfis de usuário para identificar o utilizador (Um nó correspondente ao utilizador com um valor > 0,5 é atribuída a esse usuário). Um flag é gerado se nenhuma correspondência for encontrada, este é considerado como uma anomalia. No entanto, isto exigia grande quantidade de dados para treinar a rede para cada utilizador. Dados insuficientes para um usuário pode levar à falsos positivos para o comportamento desse usuário na rede. O sistema teve uma taxa de detecção de anomalia de 96% e 7% de falsos alarmes.

Cannady propôs um modelo chamado Multi-layered Perceptron (MLP) (**??**), utilizando o algorítimo de *backpropagation* ao invés de criar perfis individuais. Foi necessário 26.13 horas para completar com aproximadamente 98% de acerto no conjunto de dados de treinamento e 97,5% no conjunto de dados de teste.

Devaraju e Ramakrishnan implementaram a Rede Neural Probabilística (PNN)

(**??**) esta obteve desempenho melhor do que uma rede do tipo FeedForward e Radial Basis. No entanto, a PNN (precisão = 80,38%) conseguiu ser apenas 0,02% melhor do que a rede do tipo FeedForward (precisão = 80,4%) nao sendo uma diferença significativa. A precisão do Radial Basis é de 75,4%. Resultados mais baixos que as abordagens que vimos até agora.

No trabalho em conjunto do Departamento de Automação e Sistemas da Universidade Federal de Santa Catarina (**??**) e com a Pontifícia Universidade Católica do Paraná (**??**), foi desenvolvido um modelo multi-camadas chamado de POLVO-IIDS. Com o objetivo principal de prover um sistema de detecto de intrusão inteligente (IIDS) que seja preciso, flexível, tolerante a variações de ataques, adaptativo a variações de ambiente, modular e que atue em tempo real. O modelo consiste em duas camadas: A primeira camada é o classificador, composta de uma rede neural de Kohonen, fazendo uma pré-seleção do trafego de entrada, através da analise de características contidas nos pacotes em um determinado período de tempo e indicando como saída se é uma intrusão ou normal. A segunda camada é detector de anomalias, composto por quatro Support Vector Machines, cada um trata de uma categoria distinta de ataque, sendo especializado no tipo correspondente e com duas opções de saída: trafego normal ou atividade maliciosa. As categorias usadas na criação do detector de anomalias foram: DoS, Worm, Scan, R2L/Normal. O protótipo foi implementado utilizando a linguagem de programação Java e a rede neural de Kohonen e as SVMs foram implementadas no framework Joone (Java Object Oriented Neural Engine Versao 2.0). Os resultados obtidos com os testes, mostrou que o modelo é eficiente, usando a base de dados KDD Cup 1999 Data,foram efetuados quatro testes para medir as taxas de acerto. Os testes 1 e 2 contaram com 15000 e 30000 dados de entrada, respectivamente, para treinamento da rede com reforço de 100 vezes para cada entrada. Os testes 3 e 4 utilizando os mesmos dados de entrada, porem com reforço de 1000 vezes para cada entrada. O teste 4 se mostrou o mais eficiente utilizando 30000 entradas e reforço de 1000 vezes para cada entrada, teve uma taxa de acerto médio de 96,55%, embora tenha levado cerca de 40 minutos.

Shraddha Surana propôs um modelo (**??**) utilizando clusterização Fuzzy e Redes neurais artificiais, onde o dataset de treinamento é dividido em N subsets, que são distribuídos em N redes neurais intermediarias, com mais uma camada de RNA que agrega os resultados das redes intermediarias para a partir dai realizar a classificação de novos dados. Esta abordagem conseguiu uma taxa de detecção de 81,6%.

Um trabalho feito na Pontifícia Universidade Católica do Rio de Janeiro (**??**), demonstra um teste de desempenho para algumas diferentes entradas para o treinamento de Redes neurais artificiais.

Utilizando 4 redes e no máximo 4 saídas possíveis, é testado diferentes composições, de forma a encontrar uma ideal maneira de treinamento de uma rede neural, utilizando as

41 categorias disponíveis. As 4 são categorizas como Normal ou entre 3 tipos de ataque DoS Smurf, Nepture ou Back. A primeira rede utiliza 9 entradas e uma saídas, onde a saída pode ser -1 para normal e 1 para ataque. A segunda rede utiliza todas as 41 entradas e uma saída, onde a saída pode ser -1 para normal e 1 para ataque. A terceira utiliza nove entradas e uma saída, onde a saída pode ser -1 para normal,0 para Neptune e 1 Smurf. A quarta e ultima rede, usa 9 entradas e quatro saídas que são organizadas em (-1 1 1 1) para Normal,(1 -1 1 1) para Neptune,(1 1 -1 1) para Back e (1 1 1 -1) para Smurf.

Utilizando a base de dados KDD Cup 1999 para o treinamento das redes neurais, os resultados foram bons e com baixa taxa de falsos positivos, as taxas de acertos foram acima de 90% para todas as configurações testadas, a rede que teve um melhor resultado foi a terceira rede, tendo 97,5% na sua taxa de acertos contra.

No trabalho feito na universidade de Waterloo no canada(**??**), foi apresentado duas técnicas de redes neurais baseados em hierarquia para IDS, hierarquia em series e hierarquia paralela, onde o objeto para a detecção de ataques de abuso e anomalia em tempo real sem a interrupção humana.

Usando dois pré-requisitos para redes neurais hierarquias. O primeiro, cada classificação individual deve acertar seu desempenho, caso contrario,o erro é enviado para os níveis acima, acumulando e influenciados os níveis mais baixos. A taxa de detecção e de falso positivo são os principais indicadores de desempenho.

A segunda, as classificações,basicamente, podem ser dividas em grupos seguindo

alguns critérios, cada grupo pode ser associado para seu própria classificação, então as classificações ou sua saída pode ser combinada em conjuntos.

O artigo conclui em seus resultados, uma ótima habilidade para detectar instruções conhecidas e desconhecidas com um curte período de tempo para o treinamento, uma alta taxa de detecção e um baixo índice de falsos positivos. O IDS de hierarquia de series conseguiu monitorar em tempo real o trafego na rede, treinando automaticamente novas intrusos modificando sua estrutura para novas classificações. O IDS de hierarquia paralela, resolveu em partes, os problemas da hierarquia em series, sendo mais rápido para ser executado.