

Segurança de Computadores e Redes

Universidade Estadual Paulista Júlio de Mesquita Filho

Prof. Dr. Kelton Augusto Pontara da Costa

Aula 3

Sumário

- 1 Abordagem da aula 3
 - Técnicas inteligentes para detecção de anomalias em redes
 - Machine Learning para sistemas de detecção de anomalias

- A detecção de anomalias em redes de computadores é uma área de estudo bastante ativa e várias técnicas são usadas;
- A classificação das técnicas de detecção de anomalias de rede, presentes na literatura, é uma tarefa difícil devido a diversidade e ao desenvolvimento constante de novas técnicas. Os métodos são classificados como de detecção de anomalias de rede em métodos baseados no Conhecimento, na Aprendizagem de Máquina e Análise Estatística.

- Conhecimento: Máquina de estados finitos; Sistemas especialistas ou baseado em regras; Busca por padrões (Pattern Matching);
- Aprendizagem de Máquina: Redes bayesianas; Cadeias de Markov; Redes Neurais; Lógica difusa (Fuzzy); Algoritmos genéticos; Algoritmos de agrupamento (Clustering); Sistemas imunológicos artificiais;
- Análise de Sinais: Análise estatística; filtros de Kalman; CUSUM (CUmulative SUM); Séries Temporais; Wavelets.

- Em relação a classificação adotada, foi acrescentado a classificação dos métodos de detecção, as técnicas derivadas da análise de sinais, separando-se algumas das técnicas de análise estatística. Na análise de sinais são usadas técnicas mais elaboradas para a modelagem dos dados e criação de um perfil que as baseadas na análise estatística básica.

- Os métodos baseados em conhecimento, ou baseados em regras, fazem uso de um conjunto de regras e parâmetros elaborados e classificados por um especialista, usando algum formalismo, como máquina de estados finitos por exemplo. Tais métodos são muito robustos, apresentando poucos falsos positivos, e flexíveis. A principal desvantagem, no entanto, está na dificuldade e demora em se obter o conhecimento de qualidade necessário.

- A abordagem de aprendizagem de máquina baseia-se no estabelecimento de um modelo implícito ou explícito que permite que padrões sejam analisados e classificados. São usadas diversas técnicas, como redes neurais e algoritmos de agrupamento, com diferentes propriedades. Contudo, a principal característica da abordagem está na necessidade de uma fase de treinamento com dados rotulados para a diferenciação do comportamento aceitável do não aceitável pelo sistema. As principais vantagens destes métodos estão na flexibilidade, adaptabilidade e capacidade de capturar interdependências desconhecidas nos dados. Porém, esta abordagem depende da determinação (rotulagem) do comportamento aceitável pelo sistema e os métodos empregados demandam muito de recursos computacionais.

- Métodos derivados da análise de sinais também são propostos para a detecção de anomalias de rede. Nos métodos baseados na análise de sinais, um perfil é criado representando o comportamento passado da rede. O perfil usa métricas de tráfego, como número de pacotes por protocolo, número de conexões e outras. Um alerta de anomalia é disparado quando o comportamento atual da rede difere significativamente do encontrado no perfil, ultrapassando algum limite (threshold) estabelecido. A principal vantagem desses métodos está em não precisar de algum conhecimento predefinido do comportamento padrão da rede, pois são capazes de se adaptar ao comportamento da rede. A principal dificuldade, no entanto, está na definição dos parâmetros, o que influencia na taxa de detecções e de falsos positivos.

- Tendo como vantagem não necessitar de conhecimento predefinido ou de uma etapa de treinamento, as abordagens baseadas na análise de sinais tornam-se interessantes para uso na detecção de anomalias devido a variabilidade do tráfego de rede. A maioria dos métodos baseados na análise de sinais para detecção de anomalias de rede presentes na literatura apresentam ao menos três etapas diferentes: Seleção de Variáveis, Transformação dos dados e Geração de Alarmes.

- A detecção de anomalias é uma atividade complexa. A seleção do conjunto de variáveis usadas pelo processo de análise de dados influencia na capacidade de detecção do SDI e o número de variáveis usadas impacta no desempenho computacional da ferramenta. No entanto, a seleção de variáveis normalmente é guiada por critérios empíricos. As variáveis selecionadas dependem também do tipo de SDI usado e dos tipos de ataques ou anomalias de interesse. Por exemplo, para um SDIR normalmente se está interessado nos endereços de origem e destino, portas e protocolos dos pacotes de rede. Quanto aos dados coletados em uma rede, um SDIR pode utilizar os dados do payload do pacote ou apenas as informações do header.

- Machine Learning, do inglês, é um ramo de Inteligência Artificial que adquire conhecimento através de dados de treinamento baseados em fatos conhecidos. Também nomeado ML, seu objetivo é permitir que computadores aprendam, e foca especialmente em previsões. Sistemas de detecção de intrusão baseados em anomalia normalmente utilizam técnicas de ML em sua metodologia. Sendo assim, a seguir, tópicos deste ramo serão abordados, destacando sua relação com o SDI.

- Uma tarefa árdua do machine learning, reconhecimento de padrões e mineração de dados é construir bons modelos a partir de conjuntos de dados. Geralmente, um conjunto de dados, do inglês dataset, consiste de vetores de características, no qual cada vetor descreve um objeto através de um conjunto de características. O número de características de um conjunto de dados é chamado de dimensão. Veja na Figura um exemplo, onde cada objeto é representado por um ponto no gráfico e é descrito pelas suas características; no caso, coordenada-x, coordenada-y e sua forma; logo, um vetor de características deste exemplo se assemelha a algo do tipo (.3, .6, cruz) ou (.8, .7, círculo).

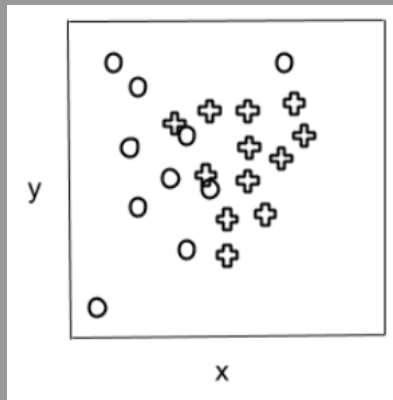


Figura 1: Dataset exemplo representado no gráfico

- Tendo isso em mente, um modelo é normalmente um modelo preditivo ou um modelo de uma estrutura dos dados dos quais deseja-se descobrir, provenientes do dataset. O processo de criar modelos a partir de dados denomina-se aprendizado ou treinamento, e é realizado através de um algoritmo de aprendizado; logo, o modelo treinado pode ser chamado de hipótese, ou aprendiz.

- Existem diferentes definições de aprendizagem, das super-visionadas e as não-supervisionadas. No aprendizado supervisionado, o objetivo é prever o valor de uma determinada característica em uma instância, ou vetor de características, desconhecida; este modelo treinado é chamado de pre-ditor. Tomando como base a Figura, caso desejasse prever a forma de algum dos objetos, define-se rótulos chamados círculo e cruz, e então o preditor deve ser capaz de prever o rótulo de uma instância da qual ele desconhece a informação do rótulo, baseando-se somente nas outras características, e.g., (.3, .6).

■

- Além disso, se este rótulo é categórico, como uma forma, a tarefa é chamada classificação, e seu aprendiz é nomeado classificador; no entanto, se o rótulo é numérico, como a coordenada-x, a tarefa é chamada regressão. Em ambos casos, o processo de treinamento é realizado em conjuntos de dados que possuem informação sobre o rótulo, e no contexto onde a classificação é binária, normalmente dividem-se os rótulos entre positivos e negativos para classificá-los.

- De maneira geral, depois do processo de aprendizagem, definir um modelo como bom é algo subjetivo, pois depende se ele consegue atender os requisitos do usuário ou não. Isso ocorre porque diferentes usuários podem ter diferentes expectativas perante os resultados da aprendizagem, e é difícil saber qual é a expectativa certa antes que a tarefa em questão seja abordada. Sendo assim, tendo em mente que o objetivo essencial da aprendizagem é generalização, ou seja, ser capaz de generalizar o conhecimento assimilado dos dados de treinamento para instâncias desconhecidas, um bom aprendiz deve ser capaz de generalizar corretamente, tendo poucos erros de generalização, que também podem ser chamados de erro de predição.

- Porém, é inviável estimar o erro de predição diretamente, pois ele depende que a informação verdadeira dos rótulos seja conhecida, o que é obscuro para instâncias desconhecidas. Logo, um processo típico é fazer com que o preditor faça predições em dados de teste nos quais as informações verdadeiras dos rótulos são conhecidas, e então observar os erros deste teste como uma estimativa dos erros de predição. Tal processo de expor um modelo aprendiz a dados desconhecidos é chamado teste. A seguir, serão abordados alguns algoritmos populares utilizados no processo de aprendizagem.

Classificador singular

- Classificador singular é uma técnica de machine learning que pode ser aplicada no desenvolvimento de um sistema de detecção de intrusão. Em trabalhos são discutidos a possibilidade de encontrar classificadores singulares sendo empregados com certa frequência. Neste estudo, Support Vector Machine (SVM) e Artificial Neural Network (ANN) são as abordagens mais populares para classificadores singulares. Enquanto isso, Lógica Fuzzy aparenta ser o classificador menos considerado dentre os observados em sua literatura, como pode ser visto na Figura. A seguir, serão detalhados alguns dos classificadores singulares mais comumente utilizados.

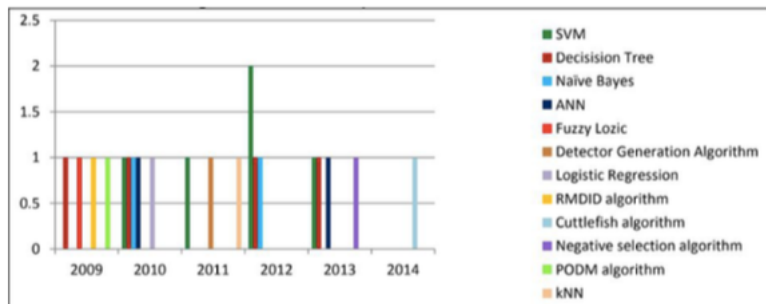


Figura 2: Distribuição dos classificadores singulares ao longo dos anos

Decision Tree

- Decision Tree é um modelo de classificação que utiliza uma estrutura semelhante a uma árvore para realizar uma decisão. Normalmente, são empregadas em operações de busca e detecção de intrusão, pois conseguem obter um bom desempenho. Uma das tarefas deste algoritmo é a criação de um classificador para a prever valor de uma classe a fim de testar instâncias desconhecidas. Este modelo é popular como um classificador singular por causa de sua simplicidade e fácil implementação.

Naive Bayes

- Naive Bayes é um dos métodos probabilísticos baseados no teorema de Bayes usado para determinar corretamente a classe de dados desconhecidos, calculando a probabilidade posterior para cada classe. Depois de calcular a probabilidade posterior para cada classe, ele determina um rótulo à classe que apresente a maior probabilidade para o dado desconhecido. Este algoritmo normalmente produz bons resultados em classificações que têm relações mais simples. Além disso, como ele requer apenas uma varredura dos dados de treinamento, ele pode facilitar o trabalho de classificação.

Artificial Neural Network

- Também conhecido como ANN, em português Redes Neurais Artificiais, é uma unidade de processamento de informações que teve seu desenvolvimento inspirado na funcionalidade do cérebro humano. Normalmente, as redes neurais estão organizadas em camadas, que são compostas por nós interconectados contendo uma função de ativação. Padrões são apresentados à rede através da camada de entrada, que se comunica com uma ou mais camadas ocultas onde, por meio de um sistema de conexões ponderadas, o processamento acontece. Essas camadas ocultas então unem-se a uma camada de saída para produzir o resultado da detecção.

MultiLayer Perceptron

- Também denominado MLP, ele é um tipo de rede neural artificial, mais específico, que realiza mapeamento linear do espaço de entrada para o espaço oculto e do espaço oculto para o espaço de saída. A estrutura da rede consiste de uma camada de entrada, n camadas ocultas e uma camada de saída. A camada de entrada é formada por nós que recebem os dados. A saída da camada anterior é a entrada da próxima camada. A camada entre as de entrada e saída é a oculta. A última camada é a de saída, utilizada para prever o rótulo do dado. Redes neurais artificiais MLP utilizam o algoritmo de aprendizagem chamado Back propagation para treinar a rede.

k-Nearest Neighbor

- Este método utiliza k registros no conjunto de treinamento que são semelhantes (vizinhança) à um novo registro para que consiga classificar este novo em relação às classes. O problema principal é como medir a similaridade entre os registros, sendo a maneira mais popular de medição a distância euclidiana entre dois registros.

Support Vector Machines

- É um método de machine learning usado para resolver problemas de classificação baseado em um hiperplano ideal em um espaço. O conceito por trás deste método para detecção de intrusão baseia-se em utilizar os dados de treinamento como uma representação somente do que são classes normais de objeto ou o que é reconhecido como uma atividade não maliciosa no SDI, considerando então todo o resto como anomalias. O classificador construído pela metodologia support vector machine distingue o espaço de entrada em uma região finita, onde estão os objetos normais, e assim considera que todo o resto do espaço contém as anomalias.

Avaliação e comparação

- Normalmente, existem múltiplas opções de algoritmos de aprendizagem a serem escolhidas, juntamente com diversos parametros a serem ajustados. A tarefa de escolher o melhor algoritmo e as configuracoes de seus parametros e conhecida como selecao de modelo, que depende da estimativa da performance do aprendiz (ZHOU, 2012). De maneira empirica, isto envolve o desenvolvimento de experimentos e testes para comparacao dos modelos. Tendo isso em mente, nao e aconselhavel estimar o erro de predicao de um aprendiz baseado em seu erro de treinamento, ou seja, o erro que o aprendiz tem perante seus dados de treinamento. Logo, uma forma adequada de se avaliar a performance e atraves de um conjunto de validacao.

Avaliação e comparação

- Isso ocorre porque ambos rotulos no conjunto de treinamento e no conjunto de validação são conhecidos antes do processo de treinamento, e devem ser utilizados juntamente para aferir e ajustar o aprendiz final uma vez que o modelo já tenha sido selecionado.

Avaliação e comparação

- Na maioria dos casos, os conjuntos de treinamento e validação são obtidos através da divisão de um conjunto de dados, já conhecido, em duas partes. Durante a divisão, as propriedades do conjunto original de dados devem ser mantidas ao máximo possível; caso contrário, o conjunto de validação pode fornecer estimativas enganosas, por exemplo, em um cenário onde o conjunto de treinamento contém apenas instâncias positivas enquanto o conjunto de validação possui somente instâncias negativas. Logo, na classificação, quando o conjunto original de dados é dividido aleatoriamente, a porcentagem das classes deve ser mantida para ambos conjuntos de treinamento e validação, processo que é denominado estratificação, evitando o acontecimento do cenário descrito anteriormente.

Avaliação e comparação

- Então, após obter os erros estimados, é possível comparar diferentes algoritmos de aprendizagem. Porém, uma simples comparação baseando-se na média dos erros não é totalmente confiável pois o algoritmo vencedor pode apresentar um desempenho melhor devido a aleatoriedade na divisão dos dados. Sendo assim, há diversos testes que podem ser empregados, entre eles, e talvez o mais comum, é o teste de hipóteses.