Evento	Salão UFRGS 2014: SIC - XXVI SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2014
Local	Porto Alegre
Título	Uma ferramenta para avaliação de algoritmos de classificação de anomalias para tráfego Internet
Autor	ANDERSON SANTOS DA SILVA
Orientador	ALBERTO EGON SCHAEFFER FILHO

Redes de computadores estão se tornando extremamente importantes no suporte a negócios, lazer e atividades em geral. Recursos de rede, tais como largura de banda, precisam ser cuidadosamente dimensionados com respeito às demandas das aplicações e características do tráfego de rede observado. Além do mais, devido à possibilidade de *cyber-attacks* e ameças à segurança, há uma crescente necessidade para que resiliência torne-se uma propriedade chave em redes de computadores. Resiliência em redes de computadores é a habilidade de manter níveis aceitáveis de operação frente a anomalias tais como ataques maliciosos, sobrecarga operacional, problemas de configuração ou falhas de equipamentos. Estratégias de resiliência podem ser definidas em termos de mecanismos de configuração para detecção e remediação. Por um lado, mecanismos de detecção tais como monitores de enlace, sistemas de detecção de anomalias e classificadores de tráfego permitem a identificação e caracterização das condições gerais da rede. Por outro lado, mecanismos de remediação tais como limitadores de tráfego são usados na subsequente mitigação de características indesejáveis na rede. Neste contexto, gerenciamento de resiliência requer que a configuração destes mecanismos seja dinamicamente refinada em resposta a, por exemplo, alta utilização de recursos, degradação de performance ou alarmes específicos de aplicação.

Em particular, classificação de tráfego corresponde a um conjunto de técnicas e algoritmos que visam categorizar o tráfego de rede. Estas técnicas podem ser quebradas em diversos domínios, incluindo classificação de protocolos na Internet (i.e., classificar fluxos de transporte de acordo com seu correspondente protocolo de camada de aplicação), classificação de pacotes (i.e., categorizar pacotes em fluxos de transporte), e classificação de tráfego para detecção de anomalia (i.e., separando fluxos maliciosos de fluxos não maliciosos). De acordo com o resultado obtido na classificação, o tráfego pertencente a uma classe em particular pode ser tratado de forma diferenciada. Devido à variedade de aplicações, protocolos e perfis de tráfego envolvidos, uma abordagem que adapta-se a novas situações e aprende com experiências passadas é necessária. Com isto em mente, técnicas de aprendizagem de máquina representam uma tendência promissora neste campo.

Como parte de um framework integrado para resiliência de redes, o conjunto de ferramentas PReSET (Policy-driven Resilience Strategy Evaluation Toolset) foi utilizado com o objetivo de permitir a avaliação off-line de estratégias de resiliência, através de um ambiente de simulação. Este tipo de abordagem permite aos operadores de rede a análise e identificação de configurações ótimas para combater diferentes tipos de ataques e outras anomalias. PReSET oferece uma série de componentes de rede implementando funções de resiliência e serviços integrados em um framework baseado em políticas. Deste modo, PBM (do inglês, Policy Based Management) pode ser usado para controlar a operação desses mecanismos e para especificar como eles devem ser reconfigurados dinamicamente quando informações sobre o estado da rede são obtidas. O trabalho desenvolvido na bolsa de IC se encaixa dentre desse contexto. O conjunto de ferramentas PReSET foi estendido para permitir a avaliação de técnicas para classificação de anomalias de tráfego baseadas em técnicas de aprendizagem de máquina. Nossa contribuição primária é oferecer aos operadores de rede e administradores um conjunto de ferramentas para a simulação e análise de uma variedade de algoritmos para classificação de tráfego anômalo, assim permitindo a fácil identificação dos melhores parâmetros de configuração e políticas de rede, quando diferentes tipos de ataques e anomalias são simulados. Nós focamos em dois algoritmos que têm sido amplamente usados para classificação de tráfego de rede: K-means e Naive Bayes, os utilizando para caracterizar tráfego de rede malicioso frente a um ataque do tipo *Distributed Denial of Service (DDoS)*.

Os resultados mostram que os dois algoritmos podem ser utilizados em conjunto para oferecer uma estratégia de classificação de tráfego Internet capaz de identificar ataques conhecidos ou não com uma precisão aceitável. Direções futuras apontam para o uso de algoritmos mais sofisticados, como por exemplo, Support Vector Machines (SVM) para classificação de ataques mais elaborados.