Um Sistema de Detecção de Intrusão Baseado em Aprendizagem por Reforço

Roger R. dos Santos[1], **Eduardo K. Viegas**[1], **Altair O. Santin**[1], **Jackson Mallmann**[1,2]

[1]Programa de Pós-Graduação em Informática (PPGIa)
Pontifícia Universidade Católica do Paraná (PUCPR)
80.215-901 - Curitiba - PR

[2]Instituto Federal Catarinense – Jardim Maluche
88.354-300 - Brusque - SC

Abstract. *Over the last years, several techniques were proposed for network-based intrusion detection. However, despite the promising results reported, these techniques do not deal with changes in network traffic over time. In this paper, an approach based on reinforcement learning technique and assessment of the reliability of classifications is proposed to maintain the accuracy of the system over time. With this technique we seek to build models capable of maintaining accuracy for longer periods and assessing reliability while maintains accuracy even with outdated models. Experiments performed in a year of network traffic, showed that the proposed approach is capable of maintaining accuracy for 8 months and reliability for the evaluated period.*

Resumo. *Nos últimos anos foram propostas diversas técnicas para detecção de intrusão em rede. Porém, apesar dos resultados promissores reportados, essas técnicas não lidam com as mudanças de tráfego de rede ao longo do tempo. Neste artigo, uma abordagem baseada em aprendizagem por reforço e avaliação da confiabilidade das classificações é proposta para manter a acurácia do sistema ao longo do tempo. Com essa técnica buscamos construir modelos capazes de manter a acurácia por maiores períodos e avaliar confiabilidade mantendo a acurácia mesmo com modelos desatualizados. Experimentos realizados em 1 ano de tráfego, demonstraram que a abordagem proposta é capaz de manter a acurácia por 8 meses e a confiabilidade pelo período avaliado.*

1. Introdução

Nos últimos anos a quantidade de ataques de rede vem crescendo de forma sistemática. Um relatório apresentado pela CISCO em 2019 [Cisco 2019] demonstra que a previsão do tráfego nuvem pode duplicar em 2020 chegando a 5,3 Hexabytes de tráfego de rede. Neste contexto, a CERT.br [CERT.br 2019], empresa responsável por relatar ataques de rede no Brasil, apresentou um relatório mostrando que em 2019 foram registrados em torno de 845 mil incidentes de rede, sendo que, desses ataques 46% eram *scanners* de rede e 34% ataques de negação de serviço.

A detecção de ataques de rede normalmente é efetuada através de sistemas de detecção de intrusão, que buscam evitar prejuízos que possam resultar em sistemas *off-fline*, vazamento de informações e até mesmo prejuízos monetários para as empresas. Na literatura, diversos autores propõem o uso de técnicas de aprendizagem de máquina para a

detecção de ataques [Viegas et al. 2019]. Em geral, as abordagens propostas utilizam algoritmos de reconhecimento de padrões, que buscam extrair o comportamento do atacante baseado em uma base de treinamento.

Diversos sistemas de detecção de intrusão são encontrados na literatura, sendo que a maior parte se baseia em identificar ataques conhecidos. Porém, devido a ocorrência de novos ataques ao longo do tempo, tais técnicas acabam sendo ineficazes. Abordagens baseadas em aprendizagem de máquina utilizam um conjunto de dados de rede que contém padrões e características de um comportamento de rede comum, de produção [Viegas et al. 2017]. Neste contexto, a base deve possuir os comportamentos da rede de maneira rotulada, considerando ambos os comportamentos normais e do atacante [Abreu et al. 2020]. Sendo assim, um sistema de detecção de intrusão é capaz de aprender a identificar ataques de rede de acordo com os dados dispostos na base de treinamento. Em outras palavras, apenas os ataques conhecidos no conjunto de dados de treinamento podem ser aprendidos pelo modelo de aprendizagem de máquina.

Uma abordagem de aprendizagem de máquina que vem se popularizando nos últimos anos é a técnica de aprendizagem por reforço (*Reinforcement Learning*) [Mallmann et al. 2020]. Esta técnica permite que o algoritmo aprenda a efetuar ações sobre o ambiente sem a necessidade de conhecimento prévio das regras do ambiente. Ela vem sendo amplamente utilizada em cenários como jogos, robótica, dentre outros. Isto ocorre, porque diferente das abordagens tradicionais, esta abordagem é capaz de aprender a classificar uma amostra de dados através de um agente, que é treinado sobre um ambiente simulado aprendendo por meio de interações e recebendo recompensas positivas ou negativas em cada ação, permitindo assim que o modelo seja aperfeiçoado. Abordagens baseadas em aprendizagem por reforço possibilitam que o modelo já existente seja atualizado de maneira incremental, diferente dos modelos tradicionais em que o classificador deve ser treinado novamente sobre a base de treinamento inteira.

Dado este contexto, este artigo propõe a utilização de técnicas de aprendizagem por reforço para o desenvolvimento de um novo modelo que possibilite a detecção de anomalias em fluxos de rede de maneira confiável e sem a assistência humana ao longo do tempo. Para tanto, a proposta é dividida em duas frentes. A primeira, trata o desenvolvimento de um modelo com uma acurácia que se mantenha confiável ao longo do tempo, mesmo sem a realização de atualizações periódicas. A segunda, trata melhorar a confiança do resultado do classificador mediante a avaliação dos resultados ao longo do tempo. Nossa proposta manterá uma confiança na acurácia ao longo do tempo, mesmos sem a realização de retreino do modelo.

Este trabalho apresenta as seguintes contribuições:

- Uma nova abordagem para detecção de intrusão baseada em aprendizagem por reforço capaz de manter a sua confiabilidade e acurácia ao longo do período de um ano sem a necessidade de atualização do modelo;
- Uma técnica para avaliar a confiabilidade da classificação ao longo do tempo. A abordagem proposta possibilita melhorar a confiança do classificador mesmo sem atualização do sistema. Adicionalmente, quando a atualização é efetuada o mecanismo é capaz de aceitar mais classificações e manter a acurácia do mecanismo de detecção.

O restante do artigo esta organizado da seguinte maneira. A seção 2 aborda o

estado da arte em detecção de intrusão em rede e aprendizagem por reforço. A seção 3 descreve os trabalhos relacionados. A seção 4 detalha o modelo proposto, enquanto que a seção 5 o avalia. Finalmente, a seção 6 conclui o trabalho.

2. Estado da Arte

2.1. Sistema de Detecção de Intrusão Baseado em Rede

Um Sistema de Detecção de Intrusão Baseados em Rede (*Network-based Intrusion Detection System*, NIDS) normalmente é composto por 4 módulos. O primeiro, *Aquisição de Dados*: responsável pela coleta dos pacotes com tráfego de rede em um ambiente monitorado por sensores. O segundo, *Extração das Características*: objetiva extrair características do comportamento da rede, normalmente fazendo uso de uma janela de 15 segundos dos pacotes de rede coletados pelo primeiro módulo. O terceiro, *Classificação*: efetua a classificação dos dados extraídos, i.e., trata da classificação de cada instância gerada pelo extrator de características, como sendo uma classe normal (tráfego comum de rede) ou um ataque (evento malicioso). Finalmente, o último modulo, denominado *Alerta*, relata todos os eventos classificados como intrusão. Na literatura diversos NIDS tem sido propostos com o passar dos anos e outros são melhorados a cada novo trabalho conduzido [Chandak et al. 2019].

Aprendizagem de máquina para classificação e construção do modelo em NIDS pode ser alcançada através de técnicas que fazem o reconhecimento de padrões das características extraídas de um conjunto de dados de fluxo de rede. De modo geral, o treinamento do modelo demanda elevados recursos computacionais, uma vez que o modelo gerado pode fazer predições corretas e incorretas de um conjunto de instâncias passadas para classificação [Kugler et al. 2020]. Deste modo, durante o treinamento, o especialista deve diminuir, através de aperfeiçoamento ao modelo, as taxas de erro obtidas. Assim, um modelo poderá trabalhar em um ambiente de produção identificando padrões de rede como normais ou de ataque. Porém, devido a ocorrência de novos ataques e serviços descobertos diariamente, o modelo construído poderá se tornar ineficaz rapidamente. Isto ocorre devido ao classificador aprender o comportamento dos ataques do conjunto de dados utilizado para treinamento, porém não consegue identificar padrões novos do ambiente de produção.

2.2. Aprendizagem por Reforço

Uma abordagem baseada em aprendizagem por reforço (*Reinforcement Learning*) trabalha com diversas políticas durante seu aprendizado. Um agente é colocado em um cenário simulado e realiza ações tentando maximizar suas recompensas. Para tanto, o algoritmo pode fazer uso de diversas políticas que são aplicadas no treinamento, entre elas destaca-se a política de gradiente que otimizam os parâmetros do treinamento tentando alcançar recompensas positivas para o agente. Além da política de gradiente, existe as chamadas cadeias de Markov, que trata-se de um processo que evolui de modo aleatório as decisões do agente, i.e., criando a probabilidade do agente passar de um estado para outro estado, ou a probabilidade dele ficar no mesmo estado maximizando a recompensa atual até ele ficar satisfeito e permitir a passagem para outro estado.

Comparado a uma abordagem tradicional de aprendizagem de máquina, a aprendizagem por reforço possibilita um treinamento aperfeiçoado do modelo, visto que o agente

que o treina pode ficar em *loop* em um mesmo conjunto de instâncias, até alcançar uma recompensa que o permita avançar dentro do cenário simulado, ou até que ele esteja satisfeito com o resultado alcançado. Dessa maneira, o agente atua como se fosse um jogador. Quando ele realiza a ação que permite a evolução do classificador, ele passa para próxima fase. Caso a ação não evolua, como por exemplo o jogador tentar atravessar um obstáculo e não conseguir, então é necessário recalcular sua trajetória para que a próxima ação seja eficaz e o jogador desvie do obstáculo. Geralmente, aprendizagem por reforço é alcançado através da implementação do algoritmo de Q-Learning. Este algoritmo possui como objetivo treinar o agente no cenário simulado e o auxiliar a fim de efetuar melhorias gradativas para a próxima ação do agente, corrigindo-o com pequenas variações na trajetória desempenhada no cenário, i.e., assumindo a maneira ideal para a próxima ação do agente.

Em NIDS, abordagens baseadas em aprendizagem por reforço ainda estão iniciando. O uso da técnica no contexto de classificação de eventos apresenta resultados promissores. Tais abordagens, geralmente, treinam um modelo de aprendizagem de máquina através do algoritmo de Q-Learning, a fim de maximizar a sua recompensa, que é tratada como a acurácia do sistema. Deste modo, o modelo é aperfeiçoado ao longo do tempo de acordo com a sua acurácia obtida na etapa de classificação.

3. Trabalhos Relacionados

Nos últimos anos, diversas técnicas para detecção de intrusão em rede foram propostas. Por exemplo, o trabalho [Van et al. 2017] utiliza técnicas de aprendizagem profunda (*deep learning*) através de *Autoencoder* e RBM empilhados, gerando resultados promissores em 4 grupos de ataques no conjunto de dados KDDCup99. Entretanto, o conjunto de dados utilizado para os testes já se encontra com 20 anos de uso, tornando as abordagens que o utilizam irrealistas, visto que novos ataques surgem diariamente. Um mecanismo de NIDS com alta acurácia e tempo de treinamento reduzido foi apresentado no trabalho [Al-Qatf et al. 2018] onde os autores propuseram uma abordagem utilizando mecanismo de *Autoencoder* esparso, inserindo em seguida o algoritmo de aprendizagem máquina de vetores de suporte aplicado ao conjunto de dados NSL-KDD que resolve alguns problemas do conjunto de dados apresentado anteriormente. Todavia, a sua existência durante 20 anos o torna ineficaz atualmente. No trabalho de [Gupta et al. 2016], os autores propõem uma abordagem de regressão aliada a uma abordagem de agrupamento com K-Means utilizando os dois conjuntos de dados já apresentados anteriormente. Os autores mostraram uma comparação na detecção de intrusão entre esses dois conjuntos de dados obtendo uma acurácia média nos resultados e mostrando diferenças entre as classificação em dois conjuntos de dados.

Também encontramos trabalhos que utilizam o algoritmo de *random forest*, como em [Nanda e Parikh 2019] que utiliza uma abordagem de agrupamento com a classificação utilizando floresta aleatória no conjunto de dados NSL-KDD. As comparações utilizam diversos classificadores que também são utilizados em diversos trabalhos, entre eles, podemos falar sobre o trabalho [Dominique e Ma 2019] que apresenta uma abordagem utilizando 4 classificadores (vetores de suporte, florestas aleatórias, perceptron multicamada e C5.0) no conjunto de dados KDDCup99, mostrando que a sua técnica com florestas aleatórias, se tornou mais eficaz que as demais técnicas comparadas no trabalho. Neste trabalho [Haddad et al. 2016] os autores apresentam a utilização

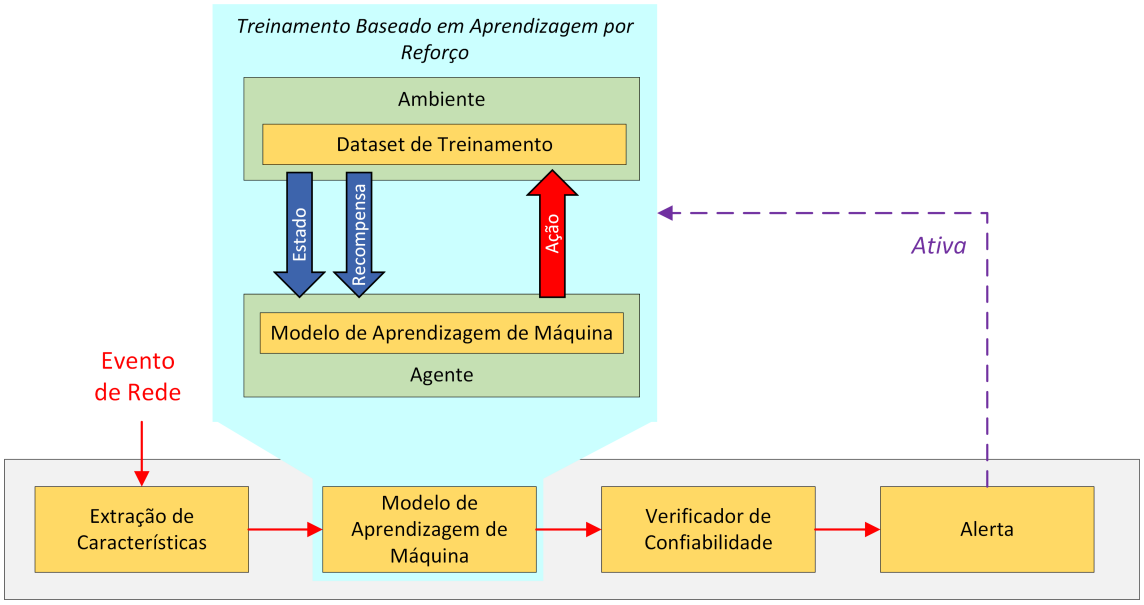

Figura 1. Arquitetura para o modelo de detecção de intrusão baseado em aprendizado por reforço.

de uma técnica colaborativa com a técnica de vetores de suporte para NIDS através da extração do tráfego de rede gerado pelo Snort em uma infraestrutura de nuvem. Os resultados foram promissores e apresentados com uma acurácia para o caso em específico. Já os autores [Cui et al. 2018], conseguiram realizar a incorporação de palavras e extração automática de recursos adequadas por meio de aprendizado profundo, obtendo um bom resultado no campo de detecção de intrusão utilizando o conjunto de dados ISCX2012.

Aprendizagem por reforço é uma técnica que vem se destacando em diversas áreas na literatura, entretanto para detecção de intrusão ainda existem poucos trabalhos que aplicam esta abordagem. O trabalho [Otoum et al. 2019] mostra uma abordagem para identificação de anomalias em redes sem fio, e apresenta uma acurácia considerável se comparado a outras abordagens tradicionais. Neste outro trabalho [Caminero et al. 2019], os autores apresentam uma nova arquitetura baseada na combinação de modelos de aprendizagem por reforço que comparado a outras técnicas apresenta uma acurácia superior, entretanto não buscam o modelo com confiança ao longo do tempo.

Portanto, podemos destacar que este artigo implementa uma nova técnica que não somente proverá uma melhor acurácia, mas também será capaz de avaliar o comportamento e a confiabilidade do classificador ao longo do tempo utilizando técnicas baseadas em aprendizagem por reforço. Além disso, os testes serão realizados através de um conjunto de dados de um ano de tráfego de rede real para que possa avaliar o comportamento e a confiabilidade durante a detecção de anomalias ao longo do tempo.

4. Proposta

A proposta do trabalho se baseia na construção de um novo modelo para detecção de intrusão baseado em aprendizagem por reforço. O esquema proposto possui dois principais objetivos, sendo eles a acurácia e confiança do modelo ao longo do tempo para detecção de comportamento do tráfego de rede real disposto em um ambientes de produção. A

figura1 exibe a visão geral da proposta.

A abordagem proposta recebe como entrada um evento da rede para classificação, como por exemplo um pacote da rede. O evento então possui suas características extraídas pelo módulo de *Extração de Características*. O vetor de características extraído é enviado para classificação pelo módulo *Aprendizagem de Máquina*. O modelo utilizado é construído através de técnicas de aprendizagem por reforço, que, por sua vez, consideram a confiabilidade do modelo durante o treinamento. Posteriormente, o evento classificado é avaliado de acordo com a sua confiabilidade pelo mecanismo *Verificador de Confiabilidade*. O objetivo do mecanismo é avaliar o nível de confiança na classificação do evento pelo modelo de aprendizagem de máquina, de tal modo que apenas eventos com grande confiabilidade sejam aceitos pelo sistema. Eventos com grande confiabilidade são assumidos como eventos conhecidos pelo sistema. Portanto, podem ser aceitos sem impacto na acurácia do mecanismo de detecção. Finalmente, um alerta é gerado caso um ataque seja identificado. Adicionalmente, eventos não confiáveis são utilizados para ativar o procedimento de atualização do modelo, permitindo assim que o sistema se mantenha atualizado ao longo do tempo.

As próximas subseções detalham a abordagem utilizada para treinamento do mecanismo de detecção, assim como o verificador de confiabilidade.

4.1. Treinamento Baseado em Aprendizagem por Reforço

As mudanças de comportamento do tráfego de rede nos ambientes de produção tornam as abordagens baseadas em aprendizagem de máquina não confiáveis, uma vez que a cada mudança de comportamento, o modelo de aprendizagem de máquina deve ser retreinado. Logo, para a devida implantação de modo confiável de abordagens baseadas em aprendizagem de máquina, torna-se necessário prover uma técnica capaz de manter a sua acurácia por longos períodos de tempo, mesmo se a atualização do modelo não for efetuada.

Dado este contexto, o modelo proposto utiliza técnicas de aprendizagem por reforço para manter o modelo de aprendizagem de máquina construído confiável por longos períodos de tempo. Para tanto, a abordagem reproduz o ambiente através da base de dados de treinamento (*Dataset de Treinamento*, Figura 1). O modelo é treinado a fim de aumentar as suas recompensas obtidas ao longo do tempo. As recompensas são calculadas de acordo com a proximidade da classe atribuída ao evento para a sua classe real. Em outras palavras, o modelo é aperfeiçoado de acordo com a corretude da confiança gerada por suas classificações dos eventos da base de treinamento. Ou seja, o *estado* representa o evento, a *ação* a confiança da classificação gerada pelo modelo e a *recompensa* a corretude da confiança gerada, como mostra a Figura 1.

Deste modo, a técnica desenvolvida para treinamento baseado em aprendizagem por reforço busca o treinamento de modelos de aprendizagem de máquina capazes de permanecer confiáveis, de acordo com a corretude das confianças geradas, por maiores períodos de tempo. Consequentemente, a técnica proposta permite a geração de modelos de aprendizagem de máquina que requerem menor frequência de atualização de seus modelos, uma vez que são otimizados baseados em sua confiança gerada, ao invés da mera acurácia.

4.2. Verificador de Confiabilidade

Independentemente da geração de modelos de aprendizagem de máquina capazes de permanecerem confiáveis por maiores períodos de tempo, as mudanças no tráfego de rede demandam a atualização dos modelos utilizados. Porém, a atualização periódica do modelo não é facilmente efetuada, uma vez que requer a disponibilidade dos eventos rotulados, assim como a execução de um processo computacionalmente custoso de retreino. Além disso, devido ao tempo necessário para atualização dos modelos, o modelo utilizado em produção deve permanecer confiável, enquanto o outro modelo, atualizado, está sendo treinado.

Sendo assim, o módulo *Verificador de Confiabilidade* busca prover duas principais propriedades. A primeira, garantir que apenas classificações confiáveis, ou seja, com maior probabilidade de serem corretas são aceitas pelo sistema. A segunda propriedade busca determinar quais eventos devem ser utilizados para o retreino do modelo, diminuindo assim os custos de atualização, uma vez que permite diminuir a quantidade de eventos necessários para a rotulagem e treinamento do modelo.

Para tanto, o módulo avalia o nível de confiança da classificação efetuada pelo modelo de aprendizagem de máquina. Classificações que não atingem um nível predeterminado de confiança, possuem seus alertas suprimidos e são utilizadas para retreino periódico (*Ativa*, Figura 1). Deste modo, a acurácia do sistema permanece confiável por longos períodos de tempo, uma vez que os alertas com baixa confiança são suprimidos, assim como o sistema determina quais eventos podem ser utilizados para manter o modelo atualizado. A limiar de confiabilidade deve ser estabelecida de acordo com o critério do administrador. Um valor maior de limiar, garante a confiabilidade do sistema por maiores períodos de tempo, porém suprime uma maior quantidade de alertas. Por outro lado, o uso de uma limiar menor, gera mais alertas, permitindo assim a geração de mais falso-positivos.

4.3. Discussão

A abordagem proposta busca um método baseado em aprendizagem de máquina para manter a confiabilidade da classificação por longos períodos de tempo. Para tanto, utilizamos aprendizagem por reforço para treinamento de modelos capazes de se manterem confiáveis por longos períodos de tempo, mesmo sem a atualização do mecanismo. Adicionalmente, propomos a utilização de um mecanismo de verificação da confiabilidade da classificação para garantir que apenas classificações confiáveis, e, portanto, com maior probabilidade de serem corretas, são aceitas pelo sistema. Classificações não confiáveis são utilizadas como métrica para ativar a atualização do modelo. Deste modo, a abordagem proposta trata da mudança de comportamento do tráfego de rede em duas frentes, gerando modelos confiáveis por maiores períodos de tempo, assim como avaliando a confiabilidade da classificação ao longo do tempo.

5. Avaliação

As seções a seguir, descrevem o processo de criação do conjunto de dados utilizado e a avaliação do método proposto para tratar mudanças de comportamento de tráfego de rede ao longo do tempo.

5.1. Dataset

Um conjunto de dados adequado que possibilite desenvolver técnicas para classificação de detecção de intrusão com qualidade, não é um trabalho fácil. Para tal, em nosso trabalho, utilizaremos os arquivos de fluxo de rede real, coletados diariamente por um intervalo de 15 minutos, com um link que fica disposto entre EUA e Japão disponibilizado pela MAWI [Mawi]. Para o propósito do nosso trabalho, o conjunto de dados utilizado pertence ao ano de 2016 que compreende em torno de 10 TB de dados, com cerca de 300 bilhões de pacotes de rede. Para a aplicação dos algoritmos de aprendizagem de máquina é necessário aplicação de técnicas de extração de características, sendo assim, utilizaremos uma técnica de aprendizagem de máquina não supervisionada da MAWILAB [Mawilab] que possibilita rotular automaticamente as instâncias encontradas no conjunto de dados como normal ou ataque. A extração das características é feita com intervalo de 15 segundos, onde é feita a extração de 20 características baseadas no fluxo disponibilizado pela MAWI.

5.2. Treinamento dos Modelos

Conforme apresentado na Figura 1, nossa proposta foi construída com um modelo baseado em aprendizagem por reforço. Para tanto utilizamos um modelo de rede neural, o *Multilayer Perceptron* (MLP) com 250 neurônios para gerar as estimativas. Desta maneira permitimos que o algoritmo de aprendizagem por reforço tenha maior chance de correlacionar estimativas próximas ao cenário de treinamento. A construção da aplicação foi realizada utilizando a linguagem de programação Python, e a API TensorFlow, assim como o algoritmo tradicional para comparação também utilizou a mesma configuração abordada pelo aprendizagem por reforço.

Foi utilizado o kit de ferramentas disponibilizado pela OpenAI Gym [Gym] para a construção do ambiente onde o agente irá treinar suas ações. Esta API permitiu o desenvolvimento do cenário de aprendizagem do nosso agente. Por fim, a implementação do algoritmo de *Q-Learning* modificado para nosso cenário foi implementado para executar 10 mil iterações com o processo total construído (*Dataset de Treinamento*, Figura 1). Cada **interação** possui 100 turnos responsáveis a avaliar e predizer a classificação de 1000 instâncias passadas de forma aleatória ao algoritmo, sendo que cada turno é responsável por calcular as estimativas das próximas ações do algoritmo através da política adotada e adaptada de acordo com as recompensas recebidas. Em outras palavras, o modelo se adapta a medida que o agente recebe recompensas positivas ou negativas.

O modelo foi treinado utilizando o conjunto de dados balanceado do mês de janeiro de 2016, em que durante os treinamentos os dados são passados de forma aleatória ao algoritmo e avaliado como é seu comportamento dentro do ambiente. A cada turno passado pelo algoritmo uma nova amostragem de 1000 instâncias é passada de forma aleatória. Sendo assim, as mesmas instâncias podem ser passadas diversas vezes, tornando um processo de reforço para que o algoritmo consiga aprender com maior clareza o comportamento do ambiente durante seu treinamento, e se tornar mais sensível a mudanças de comportamento. Por outro lado, os testes foram realizados com os dados do restante do ano sendo passado todas as instâncias disponibilizadas sem nenhum retreino ao longo deste período de tempo. Deste modo, tornou-se possível replicar um ambiente de produção, em que o modelo é treinado em um período de tempo e avaliado posteriormente, quando utilizado de fato em produção.

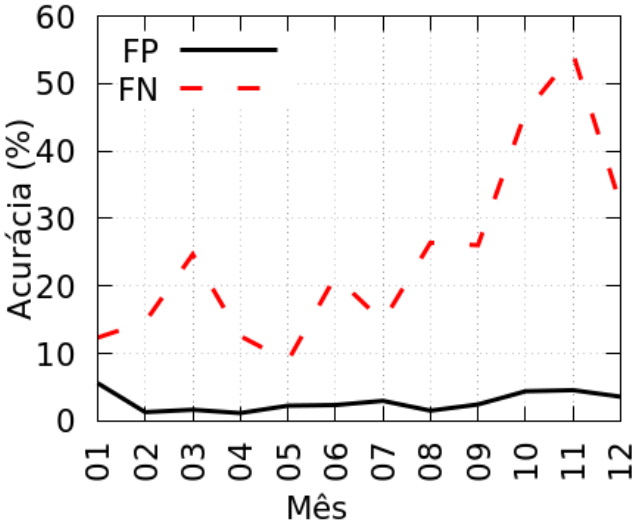

Figura 2. Acurácia ao longo do tempo da abordagem tradicional, através de uma MLP com 250 neurônios. Abordagem é treinada utilizando os dados do mês 01 sem atualizações ao longo do tempo. Taxas de erro aumentam consideravelmente após o treinamento.

5.3. Resultados

A primeira avaliação abordou a performance da abordagem tradicional, sem aprendizagem por reforço e sem a verificação da confiabilidade para fins de comparação. A Figura 2 apresenta a taxa de falsos positivos (FP) e falsos negativos (FN) com o treinamento do algoritmo tradicional perceptron multicamadas com a mesma quantidade de neurônios que foram aplicados em nossa abordagem. O FP é dado pela taxa de eventos normais incorretamente classificados como eventos de ataque. Por outro lado, a taxa de FN é dada pela taxa de ataque incorretamente classificados como normais. É possível observar que o algoritmo tradicional apresenta um aumento significativo nas taxas de FN nos primeiros meses, logo após o treinamento em Janeiro. Adicionalmente, ao atingir a metade do ano, a taxa de erro aumenta significativamente, o tornando ineficaz para a detecção dos ataques de rede.

A Figura 3 exibe a acurácia da abordagem proposta através de aprendizagem por reforço, sem atualizações e sem o verificador de confiabilidade. É possível notar que o modelo é capaz de manter a sua acurácia ao longo do tempo, mesmo sem atualização, quando comparado a abordagem tradicional de geração de modelos. Mais especificamente, a abordagem proposta permanece confiável por 9 meses, apresentando um aumento na sua taxa de FN de até 10% no mês 11. Portanto, a abordagem proposta de geração de modelo através de aprendizagem por reforço é capaz de gerar modelos confiáveis por longos períodos de tempo, quando comparado a abordagem tradicional.

Ao relacionar a acurácia da nossa proposta com a abordagem tradicional, é possível observar que a confiança de fato no algoritmo tradicional dura apenas 1 mês. Isto porque a taxa de erro aumenta significativamente após o período utilizado no treinamento. Além disso, a queda na acurácia da abordagem tradicional é significativamente maior após maiores períodos de tempo, como por exemplo no mês 11, apresentando 20% de aumento médio na taxa de erro. Por outro lado a nossa abordagem se manteve estável

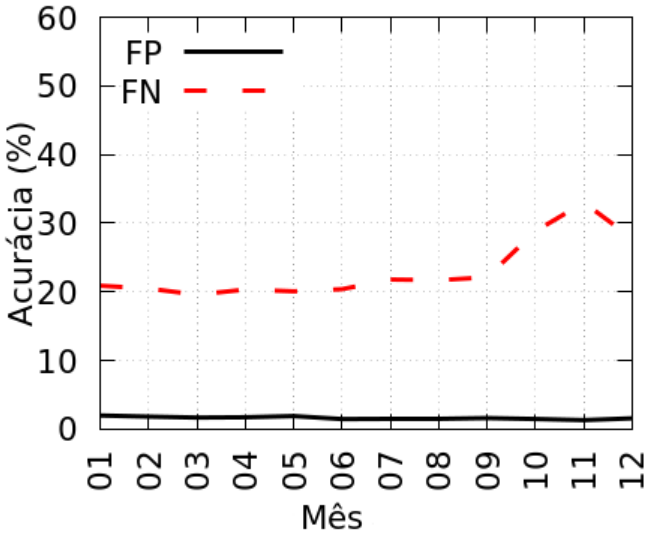

Figura 3. Acurácia ao longo do tempo da abordagem proposta de aprendizagem por reforço, através de uma MLP com 250 neurônios. Abordagem é treinada utilizando os dados do mês 01 sem atualizações ao longo do tempo. Taxas de erro permanecem similares por um período de tempo maior.

por pelo menos 10 meses, enquanto que começou a ter uma queda significava no mês 11, com uma variação muito menor quando comparada ao algoritmo tradicional.

A Figura 4 compara a acurácia da abordagem tradicional com nossa abordagem, sabendo que as duas utilizaram o mesmo conjunto de Janeiro de 2016 para treinamento, e testaram no restante do ano. É possível observar que a nossa abordagem apresentou uma diminuição de apenas 8% na detecção dos ataques de rede no pior caso, enquanto a abordagem tradicional só se manteve confiável até o mês 03 onde teve uma diminuição de 12% na detecção dos ataques de rede. Sendo assim, nossa abordagem melhora significativamente a confiança de um modelo que permanece ao longo do tempo sem a necessidade de atualização, tendo pouca variação se comparado a outras abordagens. Esse é um dos pontos onde nossa proposta começa a atingir um objetivo que mostra a possibilidade de manter a confiança mesmo em grandes espaços de tempo.

Por fim, avaliamos a abordagem proposta aliada ao verificador de confiabilidade. Para tanto, utilizamos o limiar em 90% de confiabilidade. Ou seja, o evento deve ser classificado com mais que 90% de confiança pelo modelo de aprendizagem de máquina para ser aceito pelo sistema. É importante ressaltar que o limiar deve ser definido de acordo com o critério do administrador.

As Figuras 5 e 6 exibem as taxas de aceite e a acurácia ao longo do tempo quando a abordagem do verificador de confiabilidade é utilizada, sem atualizações ao modelo de aprendizagem de máquina. É possível notar que, aliado ao verificador de confiabilidade, a abordagem proposta é capaz de manter, e ainda melhorar a acurácia obtida durante o período de treinamento. Como impacto da falta de atualização no sistema, a taxa de aceite de eventos ao longo do tempo diminui a medida que o modelo de aprendizagem de máquina se torna obsoleto. Ou seja, o mecanismo proposto deve aceitar menos eventos, a fim de manter a confiabilidade nas classificações. Portanto, a técnica proposta de verificação de confiabilidade permitiu garantir que o sistema se mantenha confiável, ou

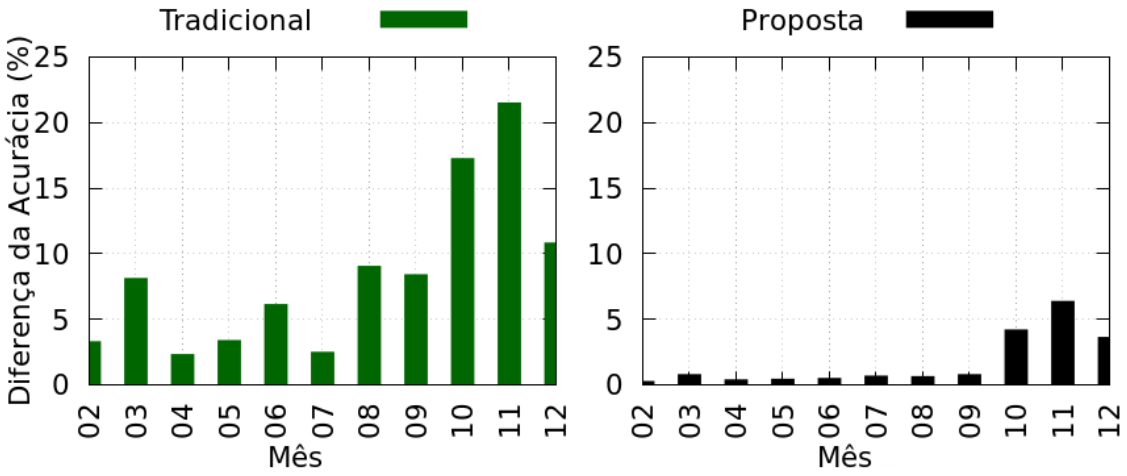

Figura 4. Diferença média da acurácia da abordagem tradicional e da nossa proposta sem o uso do verificador de confiabilidade quando comparado ao mês 01. É possível notar que a abordagem proposta mantém a acurácia obtida durante o período de treinamento por um maior tempo.

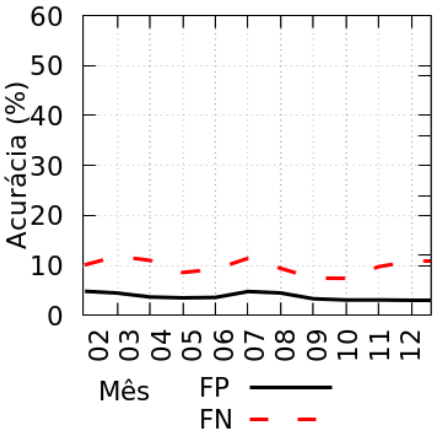

Figura 5. Acurácia ao longo do tempo da abordagem proposta fazendo uso do verificador de confiabilidade.

seja, com a mesma acurácia obtida durante o período de treinamento, mesmo com um modelo desatualizado.

6. Conclusão

As abordagens de aprendizagem de máquina atualmente utilizadas na literatura são incapazes de lidar com novos padrões de conjuntos de ataques, tornando os modelos criados ineficazes após poucos meses de uso. Neste contexto, diversos autores consideram treinamentos em busca da melhor precisão do algoritmo e desconsideram a necessidade que este modelo dure por um longo período de tempo. Tal característica faz com que novos treinamentos sejam necessários de maneira constante e periódica. Porém, isso faz com que um ambiente de produção esteja sujeito a um enorme risco devido à dificuldade de gerar novos modelos.

Neste trabalho, foi mostrado que a nossa proposta pode desenvolver um novo

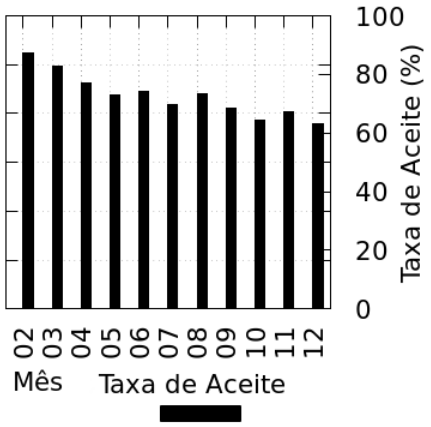

Figura 6. Taxa de aceite ao longo do tempo da abordagem proposta fazendo uso do verificador de confiabilidade. A taxa de eventos aceitos diminui a medida que o modelo se torna desatualizado, porém, a abordagem mantém a taxa de acerto estável ao longo do tempo.

modelo de detecção de intrusão para longos períodos de tempos, sem a necessidade de atualização de modelo, baseado em técnicas de aprendizagem por reforço. Ou seja, nosso trabalho foi capaz de melhorar significativamente a confiança de um algoritmo de detecção de intrusão que pode ser utilizado em uma ambiente de produção por longos períodos de tempo.

Como trabalho futuro, iremos focar nossos esforços em tornar o retreino do algoritmo menos necessário ao longo de um maior período. Isto possibilitará a utilização dos modelos anteriores e descobrindo novos ataques com incrementos ao modelo atual, sem a necessidade de retreinar o modelo novamente, além da melhora da acurácia do algoritmo ao longo do tempo.

Referências

Abreu, V., Santin, A. O., Viegas, E. K., and Cogo, V. V. (2020). Identity and access management for IoT in smart grid. In *Advanced Information Networking and Applications*, pages 1215–1226. Springer International Publishing.

Al-Qatf, M., Lasheng, Y., Al-Habib, M., and Al-Sabahi, K. (2018). Deep learning approach combining sparse autoencoder with svm for network intrusion detection. In *IEEE Access*, volume 6, pages 52843–52856.

Caminero, G., Lopez-Martin, M., and Carro, B. (2019). Adversarial environment reinforcement learning algorithm for intrusion detection. In *Computer Networks*, volume 159, pages 96–109.

CERT.br (2019). Estatísticas dos incidentes reportados ao cert.br.

Chandak, T., Ghorpad, C., and Shukla, S. (2019). Effective analysis of feature selection algorithms for network based intrusion detection system. In *2019 IEEE Bombay Section Signature Conference (IBSSC)*, pages 1–5.

Cisco (2019). Cisco visual networking index: Global mobile data traffic forecast update, 2019 – 2022.

Cui, J., Long, J., Min, E., and Mao, Y. (2018). Wedl-nids: Improving network intrusion detection using word embedding-based deep learning method. In Springer, editor, *Modeling Decisions for Artificial Intelligence*, volume 11144.

Dominique, N. and Ma, Z. (2019). Enhancing network intrusion detection system method (nids) using mutual information (rf-cife). In Springer, editor, *Security with Intelligent Computing and Big-data Services. SICBS 2018. Advances in Intelligent Systems and Computing*, volume 895.

Gupta, D., Singhal, S., Malik, S., and Singh, A. (2016). Network intrusion detection system using various data mining techniques. In *International Conference on Research Advances in Integrated Navigation Systems (RAINS - 2016)*.

Gym. Openai. available online.

Haddad, Z., Hanoune, M., and Manoumi, A. (2016). A collaborative framework for intrusion detection (c-nids) in cloud computing. In *2016 2nd International Conference on Cloud Computing Technologies and Applications (CloudTech)*, pages 261–265.

Kugler, E., Santin, A. O., Cogo, V. V., and Abreu, V. (2020). Facing the unknown: A stream learning intrusion detection system for reliable model updates. In *Advanced Information Networking and Applications*, pages 898–909. Springer International Publishing.

Mallmann, J., Santin, A. O., Viegas, E. K., dos Santos, R. R., and Geremias, J. (2020). PPCensor: Architecture for real-time pornography detection in video streaming. *Future Generation Computer Systems*, 112:945–955.

Mawi. Mawi. available online.

Mawilab. Mawilab. available online.

Nanda, N. and Parikh, A. (2019). Hybrid approach for network intrusion detection system using random forest classifier and rough set theory for rules generation. In Springer, editor, *Advanced Informatics for Computing Research. ICAICR 2019. Communications in Computer and Information Science*, volume 1076.

Otoum, S., Kantarci, B., and Mouftah, H. (2019). Empowering reinforcement learning on big sensed data for intrusion detection. In *2019 IEEE International Conference on Communications (ICC)*, pages 1–7.

Van, N., Thinh, T., and Sach, L. (2017). An anomaly-based network intrusion detection system using deep learning. In *2017 International Conference on System Science and Engineering (ICSSE)*, pages 210–214.

Viegas, E., Santin, A., Bessani, A., and Neves, N. (2019). BigFlow: Real-time and reliable anomaly-based intrusion detection for high-speed networks. *Future Generation Computer Systems*, 93:473–485.

Viegas, E. K., Santin, A. O., and Oliveira, L. S. (2017). Toward a reliable anomaly-based intrusion detection in real-world environments. *Computer Networks*, 127:200–216.