ESCOLA DE COMANDO E ESTADO-MAIOR DO EXÉRCITO

ESCOLA MARECHAL CASTELLO BRANCO

Maj QEM **CRISTIANO ROLIM** PEREIRA

Sistema de Detecção de Intrusão usando Big Data, Inteligência Artificial e Sistemas Colaborativos

Rio de Janeiro
2020

Maj QEM **CRISTIANO ROLIM** PEREIRA

Sistema de Detecção de Intrusão usando Big Data, Inteligência Artificial e sistemas colaborativos

Projeto de pesquisa apresentado à Escola de Comando e Estado-Maior do Exército, como pré-requisito para matrícula no Curso de Especialização em Ciências Militares, com ênfase em Defesa.

Orientador: Ten Cel Com Rodrigo Damasceno Sales

Rio de Janeiro
2020

P436s Pereira, Cristiano Rolim

Sistema de Detecção de Intrusão usando Big Data, Inteligência Artificial e Sistemas Colaborativos. / Cristiano Rolim Pereira. —2020.
39 f. : il. ; 30 cm

Orientação: Rodrigo Damasceno Sales.
Trabalho de Conclusão de Curso (Especialização em Ciências Militares)—Escola de Comando e Estado-Maior do Exército, Rio de Janeiro, 2020.
Bibliografia: f. 36-39

1. SISTEMA DE DETECÇÃO DE INTRUSÃO. 2. BIG DATA. 3. INTELIGÊNCIA ARTIFICIAL. 4. SISTEMAS COLABORATIVOS. I. Título.

CDD 006.3

Maj QEM **CRISTIANO ROLIM** PEREIRA

Sistema de Detecção de Intrusão usando Big Data, Inteligência Artificial e sistemas colaborativos

Projeto de pesquisa apresentado à Escola de Comando e Estado-Maior do Exército, como pré-requisito para matrícula no Curso de Especialização em Ciências Militares, com ênfase em Defesa.

Aprovado em 30 de outubro de 2020.

COMISSÃO AVALIADORA

_____

Rodrigo Damasceno Sales – TC Com - Presidente
Escola de Comando e Estado-Maior do Exército

_____

Luiz Adolfo Sodré de Castro Júnior – TC Cav - Membro
Escola de Comando e Estado-Maior do Exército

_____

Adriano de Paula Fontainhas Bandeira – Maj QEM - Membro
Escola de Comando e Estado-Maior do Exército

RESUMO

O crescimento exponencial da Internet trouxe como consequência negativa o incremento nos incidentes de segurança tais como ataques, invasões de redes e vazamentos de dados que comprometem empresas e instituições governamentais por todo o mundo. Uma das medidas utilizadas para mitigar ou reduzir o risco de ocorrência destes tipos de incidentes é o uso de Sistemas de Detecção de Intrusão (*Intrusion Detection Systems* – IDS) para proteger redes de variados tipos. Contudo, esses sistemas têm perdido a efetividade ao longo dos anos devido a uma série de fatores. O presente trabalho apresenta três das novas tecnologias que trarão a evolução dos sistemas IDS. Particularmente, foram estudadas as pesquisas envolvendo o aprimoramento de IDS usando técnicas de *Big Data*, de Inteligência artificial e de sistemas colaborativos. Foram levantados uma série de trabalhos recentes e relevantes que indicam o estado atual das pesquisas nas três áreas. Verificou-se que a popularização de plataformas de software para o processamento de *Big Data* tem facilitado sua aplicação ao contexto de cibersegurança, com resultados promissores. Processo similar tem ocorrido quando ao uso de bibliotecas de inteligência artificial, mas com mais desafios práticos para obter resultados efetivos. No tocante às pesquisas de sistemas colaborativos, observou-se que há diversas pesquisas propondo soluções para as evoluções que a Internet sofrerá ainda no curto prazo. Além disso, as três tecnologias aparecem combinadas duas a duas ou mesmo todas juntas com certa frequência na literatura, indicando uma interdependência. Por fim, constatou-se a relevância dessas tecnologias para o futuro da cibersegurança.

Palavras-chave: Sistema de Detecção de Intrusão; Big Data; Inteligência Artificial; Sistemas Colaborativos.

ABSTRACT

The exponential growth of the Internet brought as a negative consequence the increase in security incidents such as attacks, network intrusions and data leaks that compromise companies and government institutions around the world. One of the measures used to mitigate or reduce the risk of these types of incidents is the use of Intrusion Detection Systems (IDS) to protect networks of various types. Nevertheless, these systems have lost effectiveness over the years due to a number of factors. This work presents three of the new technologies that will bring the evolution of IDS. In particular, research involving the improvement of IDS using Big Data, Artificial Intelligence and collaborative systems techniques were studied. A series of recent and relevant studies have been raised to indicate the current state of research in the three areas. It was found that the popularization of software platforms for Big Data processing has facilitated its application in the context of cybersecurity, with promising results. A similar process has occurred when using artificial intelligence libraries, but with more practical challenges to obtain effective results. With regard to research on collaborative systems, it was observed that there are several studies proposing solutions for the developments that the Internet will undergo in the short term. In addition, the three technologies appear combined two by two or even all together with a certain frequency in the literature, indicating an interdependence. Finally, the relevance of these technologies for the future of cybersecurity was found.

Keywords: Intrusion Detection System; Big data; Artificial intelligence; Collaborative Systems.

SUMÁRIO

1 INTRODUÇÃO

A utilização de sistemas informatizados ligados por redes de computadores nos sistemas de defesa, sistemas estratégicos e sistemas de infraestruturas críticas é um fato consolidado em todo o mundo. Além disso, se observa uma tendência de crescimento constante nessa utilização, o que acarreta uma dependência cada vez maior das nações sobre suas redes de dados.

Nesse cenário, o Brasil não é uma exceção e apresenta uma situação semelhante a que ocorre em outros países, na qual o setor cibernético tem ganhado mais relevância a cada ano. Essa relevância foi reconhecida pelo Estado Brasileiro que, em sua renovada Estratégia Nacional de Defesa (END) de 2012, definiu o setor cibernético como estratégico para a Defesa do país (BRASIL, 2012), determinando também o seu fortalecimento.

O Ministério da Defesa (MD) do Brasil classifica as ações cibernéticas em três tipos: exploração, ataque e proteção, das quais as duas primeiras têm caráter ofensivo e a última defensivo (BRASIL, 2014). A proteção consiste em neutralizar ações ofensivas dirigidas às nossas redes e é uma atividade permanente.

Em relação à proteção de redes de dados, um sistema de detecção de intrusão (*Intrusion Detection System* – IDS) tem a finalidade de monitorar o tráfego de dados em uma determinada rede para detectar ataques, atividades maliciosas em geral ou violações nas políticas de segurança da organização. Quando uma atividade de rede indesejada é detectada, um alerta é enviado para a equipe de segurança ou para outros sistemas com capacidade de atuar sobre a rede e impedir ou interromper o ataque.

Classicamente, a detecção é realizada por meio da busca de certos padrões previamente conhecidos (assinaturas) ou de anomalias que se destacam em meio ao tráfego considerado normal e que indiquem a ocorrência de ataques (MCHUGH, 2001). O crescimento constante da utilização da Internet tem tornado a tarefa do IDS cada vez mais difícil. Isso se deve ao aumento constante do volume de tráfego a ser analisado. A isso se soma a proliferação de ataques e o desenvolvimento de técnicas que objetivam evadir os mecanismos de detecção clássicos. Disso resulta a necessidade de se buscar novas abordagens no sentido de manter a efetividade dos IDS e em última instância a segurança e a proteção das redes de dados.

Uma abordagem que permite o aprimoramento dos sistemas IDS consiste na utilização de técnicas de processamento denominadas *Big Data*, com as quais

dados de elevado volume, variedade e velocidade (Os “três Vs”) podem ser armazenados, analisados e processados em tempo real, algo que não pode ser obtido com o uso de softwares e bancos de dados clássicos.

Outra proposta de melhoria para os IDS se traduz em incorporar ferramentas de inteligência artificial a esses sistemas, por meio do uso de algoritmos de classificação, aprendizado de máquina automatizado e processamento de linguagem natural.

Por último, se propõe a integração de diferentes IDS, que em um primeiro momento operavam isolados, para formar um sistema distribuído colaborativo com maior desempenho que os seus componentes teriam operando isoladamente, e com isso alcançar uma maior taxa de detecção global. Todas as três abordagens objetivam permitir que os IDS sejam mais eficazes e acompanhem a tendência de crescimento constante das redes de dados.

2 SISTEMAS DE DETECÇÃO DE INTRUSÃO

O Instituto Nacional de Padrões e Tecnologia dos Estados Unidos – *National Institute of Standards and Technology* (NIST) – define como intrusão qualquer conjunto de ações que tente comprometer a integridade, confidencialidade ou a disponibilidade de um recurso computacional (GRANCE e colab., 2003). De acordo com o Glossário de Segurança da Informação do Gabinete de Segurança Institucional (BRASIL, 2019):

> INTEGRIDADE - propriedade pela qual se assegura que a informação não foi modificada ou destruída de maneira não autorizada ou acidental; […] CONFIDENCIALIDADE - propriedade pela qual se assegura que a informação não esteja disponível ou não seja revelada a pessoa, a sistema, a órgão ou a entidade não autorizados nem credenciados; […] DISPONIBILIDADE - propriedade pela qual se assegura que a informação esteja acessível e utilizável sob demanda por uma pessoa física ou determinado sistema, órgão ou entidade devidamente autorizados;

Por sua vez, o IDS consiste em um software ou um equipamento, operando isolado ou combinados em conjuntos, destinado a detectar intrusões ou mesmo tentativas de intrusão em sistemas e redes de dados.

As propostas iniciais de IDS destinados a detectar esse tipo de ação surgiram na década de 80 (DENNING, 1987). Uma das classificações comumente utilizadas divide os IDS em dois tipos principais: baseados em assinaturas e baseados em anomalias (JYOTHSNA e colab., 2011). Os IDS baseados em assinaturas possuem um banco de dados contendo padrões de texto e dados binários de ataques previamente conhecidos. Os dados que trafegam na rede protegida devem ser analisados e comparados com o banco de dados de assinaturas, algo que demanda um esforço computacional relativamente baixo (JYOTHSNA e colab., 2011).

A principal limitação do IDS baseado em assinaturas está na necessidade de conhecimento prévio do ataque a ser detectado, o que demanda uma atualização frequente no banco de dados de assinaturas conhecidas. Além disso, foram desenvolvidas diversas técnicas para evitar a detecção e permitir que ataques superem a proteção de IDS, o que se denomina “evasão de IDS” (CHENG e colab., 2012). Como técnicas de evasão de IDS comuns, podemos citar as seguintes:

- fragmentação de pacotes, o que necessita que o IDS mantenha para remontar os pacotes recebidos, o que consome recursos do equipamento e demanda tempo extra de processamento.
- a ofuscação de *strings*[1] (cadeias de caracteres) por meio de codificações diversas, como hexadecimal ou UTF8, o que exige que o IDS normalize as *strings* recebidas antes das comparações.
- O *spoofing*[2] de endereços, quando o atacante falsifica o endereço dos remetentes, dificultando o bloqueio por parte do IDS, algo comum em ataques de negação de serviço que utilizam o protocolo UDP, a exemplo dos ataques de negação de serviço por meio de amplificação UDP usando o protocolo NTP, DNS, Chargen, SSDP, e memcached (VAUGHAN-NICHOLS, 2018).
- Variações nos ataques já conhecidos, que alteram a assinatura necessária para a correta detecção.

Um ataque que não consta na base de dados de assinaturas ou que consiga evadir o IDS resulta no que se define um “falso negativo”.

IDS baseados em anomalias buscam estimar o comportamento “normal” na rede a ser protegida e geram um alerta de detecção quando a variação entre o tráfego observado e o normal excede um limiar previamente definido (GARCÍA-TEODORO e colab., 2009). Esse tipo de IDS tem a vantagem de possibilitar a detecção de ataques novos e ainda desconhecidos, com a contrapartida de necessitar maior esforço na sua configuração e estar sujeito a gerar alertas de detecção para tráfego normal, no que se configura um “falso positivo”, e que pode gerar problemas na operação das redes protegidas (RAJU, 2005).

A Tabela 1 ilustra os possíveis cenários resultantes de situações de intrusão ou não intrusão em combinação coma ocorrência de detecção por um IDS:

[1] O termo *string* é comumente usado para se referir a uma cadeia de texto em linguagens de programação, mesmo em português.

[2] No contexto de segurança cibernética, *spoofing* se refere a falsificação de endereços ou identificadores de remetente, como endereços de rede, remetentes de correio eletrônico ou mesmo falsificação de números de telefone para o envio de mensagens SMS via rede de telefonia celular. Em geral, estão sujeitos a *spoofing* protocolos e aplicações que não realizam a autenticação de remetentes.

Tabela 1 – Cenários de detecção

		Ocorreu intrusão	
		Sim	Não
IDS identifica intrusão	Sim	Positivo correto	Falso positivo
	Não	Falso negativo	Negativo correto

Fonte: O Autor (2020)

Buscando mitigar as desvantagens dos métodos baseados em assinaturas e dos baseados em anomalias, a maioria das implementações de IDS disponíveis atualmente utiliza uma abordagem híbrida, mesclando ambas as técnicas.

Outra forma de classificar os IDS se baseia na posição em que ele é posicionado na rede. Quando instalado em um computador específico, o IDS é denominado IDS de Hospedeiro (*Host IDS* - HIDS), e tem o objetivo de detectar intrusões para um único equipamento, que pode ser uma estação de trabalho, um servidor ou mesmo um *smartphone* (HALILOVIC e SUBASI, 2012).

Este tipo de IDS tem a capacidade de monitorar alterações no sistema de arquivos e na memória do hospedeiro, e a vantagem de ser mais tolerante a certas técnicas de evasão. Contudo, em caso de comprometimento do sistema onde está instalado, o HIDS poderá ser alvo de ataques mais facilmente. Ademais, para proteger um conjunto de equipamentos, cada um deve receber a instalação do HIDS, gerando um esforço extra de administração e manutenção.

Quando posicionado como um ativo de rede independente, para monitoramento e análise do tráfego de uma rede, o IDS é denominado IDS de Rede (*Network IDS* – NIDS). O NIDS normalmente é posicionado em algum ponto crítico da rede, como um ponto de ligação com a Internet, ou a passagem por um *firewall*[3]. O NIDS tem a vantagem de ter uma maior visibilidade sobre a rede como um todo,

[3] Firewall é um dispositivo de rede destinado ao controle de acesso. Sua configuração define que tráfego de rede pode entrar ou sair. Por sua importância, normalmente opera em conjunto com um roteador, no ponto de transição de uma rede para outra.

em detrimento de necessitar maior capacidade de processamento para cumprir essa tarefa.

No entanto, a visibilidade que caracteriza o NIDS tem sido reduzida devido ao uso cada vez mais frequente de criptografia nas comunicações na Internet que utilizam o protocolo HTTP. Particularmente, a versão 1.3 do protocolo *Transport Layer Security* (TLS)[4], lançada em 2018, possui um mecanismo adicional de proteção que impede a inspeção de dados por IDS, algo que ainda era possível até a versão anterior, TLS 1.2 e em todas as versões do protocolo *Secure Sockets Layer* (SSL)[5], antecessor do TLS.

Outra tendência atual é a criptografia das consultas e respostas envolvendo o uso do protocolo *Domain Name System* (DNS)[6], usado ao longo de toda a internet para a resolução de nomes em endereços IP numéricos, que historicamente utiliza dados em claro, e agora apresenta alternativas como DNSCrypt, DNS-over-TLS e DNS-over-HTTPS. Essa permitem autenticar os servidores de DNS e evitar que as mensagens sejam alteradas ao longo do caminho, algo que ocorre em determinados tipos de ataques. Também evitam ataques de negação de serviço que exploram as vulnerabilidades inerentes ao protocolo UDP utilizado pelo protocolo DNS original. No entanto, essas tecnologias também permitem mascarar as consultas de DNS de modo que os ativos de rede como firewall e IDS não sejam capazes de saber qual páginas estão sendo consultadas pelo usuário final.

A combinação das proteções criptográficas envolvendo HTTP e DNS, que têm o objetivo de dar maior segurança ao usuário final, também representam um desafio para as esquipes de proteção de redes, que terão maior dificuldade de verificar o tráfego em busca de ataques. Isso se reflete no fato que organizações e países que não querem perder a capacidade de inspecionar com profundidade o seu tráfego de rede passaram a bloquear o tráfego que não podem inspecionar (CIMPANU, 2020).

---

[4] TLS é um protocolo de segurança amplamente utilizado na navegação na Internet na atualidade, com a finalidade de prover privacidade e integridade dos dados transmitidos. A descoberta de falhas e a busca pelo aprimoramento do protocolo trazem a demanda por atualizações, que resultam em versões novas.

[5] SSL é o protocolo de segurança que antecedeu o TLS e foi criado na década de 1990. Diversas falhas de segurança foram encontradas e atualmente seu uso é desencorajado.

[6] O protocolo DNS foi criado nos primórdios da Internet, ainda na década de 1980. O seu design não levou em conta diversas questões de segurança que ganhariam relevância nas décadas seguintes, o que provocou o surgimento de iniciativas para aprimorá-lo. Algo que dificulta isso é a necessidade de retrocompatibilidade, pois o protocolo é essencial para o funcionamento da Internet e usado em dispositivos antigos que devem continuar funcionando.

3 **BIG DATA**

O uso da expressão "*Big Data*" com o sentido atual tem origem na década de 1990 (LOHR, 2013) e se popularizou com a grande aceleração na taxa de geração de dados digitais observada pela humanidade desde o início do século XXI (HILBERT e LÓPEZ, 2011). O conceito é relativamente recente e não há uma definição universal aceita, mas a maioria das definições aborda a questão dos três V's: Volume, Variedade e Velocidade.

O **volume** é a principal característica de *big data*. Em 2012, cerca de 2,5 *exabytes* (10^{18} bytes) de dados digitais foram criados a cada dia, e essa taxa dobrou a cada 40 meses desde então (MCAFEE e BRYNJOLFSSON, 2012). Em 2014, o sistema de armazenamento do projeto *Internet Archive*[7] já continha 50 petabytes (10^{15} bytes) de dados. À medida que a capacidade de armazenamento dos dispositivos cresce, o limiar de volume a partir do qual se considera *big data* irá aumentar constantemente. Dessa forma, o que poderia ser considerado um problema de *big data* na década de 1990, hoje pode ser tratado por um computador pessoal comum.

A **variedade** dos dados gerados e transmitidos também é crescente. Considerando somente a quantidade de serviços UDP e TCP registrados junto ao *Internet Assigned Numbers Authority* (IANA)[8], no período da execução desse levantamento, verificamos mais de 10800 entradas (TOUCH e colab., 2020). Além desses, há diversos outros como o protocolo de controle ICMP e protocolos de roteamento como OSPF e BPGP. Dispositivos móveis executam uma ampla gama de aplicativos e a popularização da "Internet das Coisas" tem aumentado a quantidade de dispositivos online, tais como tomadas elétricas inteligentes, câmeras, eletrodomésticos, lâmpadas inteligentes e uma gama de sensores. Comentários postados em redes sociais como Facebook e Twitter se juntam a dados de geolocalização de milhões de telefones celulares e dados de compras em sites de comércio eletrônico.

A massificação do acesso à Internet se juntou à popularização da computação em nuvem ocorrida nos últimos anos, resultando em uma **velocidade**

[7] Projeto criado em 1996, consiste em uma biblioteca digital que guarda livros, filmes, música e, particularmente, páginas da Internet. Isso permite verificar como uma determinada página era em um dado momento do passado. Esse serviço é denominado "*Wayback Machine*".

[8] IANA é organização de padronização responsável por diversas tarefas de gerenciamento da Internet, incluindo as designações de faixas de endereços, serviços, nomes de domínios e banco de dados de fusos horários.

cada vez maior com a qual os dados gerados são transportados pela Internet. Disso resultou o aumento do uso da Internet numa proporção semelhante à da geração de dados, mostrando que esses fatores então interligados e se reforçam mutualmente (RUSSOM, 2011).

Sendo confrontadas com o desafio de armazenar e processar cada vez mais informações, as grandes companhias começaram a desenvolver propostas para a abordagem de *big data*. Uma opção bastante popular para a solução de problemas de *big data* é o *Hadoop* (APACHE SOFTWARE FOUNDATION, 2020), criado em 2006 e mantido pela Fundação Apache. *Hadoop* é uma solução de código aberto que permite o armazenamento e processamento de *big data* em hardware de prateleira. Um dos principais módulos do *Hadoop* é Sistema de Arquivos Distribuído *Hadoop* (*Hadoop Distributed File System* - HDFS). O acesso aos dados é realizado por meio do módulo *Hadoop* MapReduce, destinado a processamento de *big data* de modo paralelo e distribuído em clusters. Uma das críticas a este módulo está no fato de sua programação usa uma linguagem de "baixo nível" (DEWITT e STONEBRAKER, 2008), isto é, de menor abstração, o que resulta em maior dificuldade de maior propensão a erros.

Para contornar essa desvantagem, um módulo adicional comumente utilizado em conjunto com o *Hadoop* é o Apache Hive, que possibilita o acesso aos dados utilizando uma linguagem muito similar à linguagem SQL[9] utilizada nos bancos de dados relacionais, a HiveQL. Essa similaridade facilita sua adoção, uma vez que SQL está entre as linguagens mais utilizadas por desenvolvedores, conforme indicam pesquisas recentes (PUTANO, 2019; STACKOVERFLOW, 2019). Internamente, Hive converte as consultas SQL em comandos MapReduce. Hive foi inicialmente desenvolvido pela Facebook, motivados pelo crescimento exponencial de seus dados, que entre 2007 e 2009 foram de um conjunto de 15 *Terabytes* para mais de 2 *Petabytes* (2000 *Terabytes*) (THUSOO e colab., 2010). Com um crescimento tão grande, os bancos de dados relacionais não eram mais suficientes para o gerenciamento dos dados, e outra solução teve de ser desenvolvida. Facebook optou pelo armazenamento usando o módulo HDFS de *Hadoop*, mas necessitava uma camada de abstração para simplificar a

[9] SQL é a linguagem usada para consulta, acesso, leitura, inserção e remoção de dados nos bancos de dados relacionais, tais como Oracle, MySQL, Microsoft SQL Server, PostgreSQL, IBM DB2, SQLite.

programação pelos desenvolvedores. A principal vantagem do Hive é sua facilidade de uso pelo usuário (PRATHIBHA e DILEESH, 2013).

Outro módulo frequentemente associado a *Hadoop* é o Apache Spark, destinado a análise de alto desempenho de dados e processamento distribuído de larga escala. Enquanto Hive opera como uma camada extra sobre HDFS, Spark é um framework de processamento que pode ser integrado opcionalmente a *Hadoop*, quando acessa dados diretamente do HDFS ou por intermédio do Hive. Spark, contudo, não se restringe a estas fontes de dados, podendo acessar sistemas distintos como sistemas de armazenamento NFS (*Network File System*).

4 INTELIGÊNCIA ARTIFICIAL

A ideia de Inteligência Artificial (IA) acompanha a computação moderna desde seu surgimento no século XX. O “Teste de Turing”[10], que busca identificar um computador se passando por um humano, data da década de 1950. A definição de IA também é muito variada, com enfoques distintos sobre “humanidade” e “racionalidade” (RUSSELL, NORVIG, 2010).

O uso de inteligência artificial envolve disciplinas como o processamento de linguagem natural, a percepção de máquina, redes neurais, raciocínio automatizado e aprendizado de máquina. As aplicações da IA no campo da segurança cibernética são vastas. A empresa Gartner elencou a aplicação de IA no campo da cibersegurança como uma das 10 tecnologias estratégicas para o ano de 2020 (CEARLEY, JONES, et al., 2019).

O processamento de linguagem natural integra os campos de IA com a linguística e tem aplicações em diversos ramos da tecnologia da informação, como os leitores de livros eletrônicos (*e-books*), atendentes virtuais em páginas web e assistentes automáticos para deficientes visuais. Ultimamente, essa área se manifesta na tendência de assistentes virtuais como Google Assistente, Siri (Apple), Cortana (Microsoft) e Alexa (Amazon). Esses assistentes podem estar presentes na forma de aplicativos em telefones celulares e computadores pessoais, televisores ou mesmo em aparelhos dedicados, como caixas de som com microfones integrados, usados para automatizar casas inteligentes por meio de comandos voz. Nesse tipo de casa, o acendimento de lâmpadas, abertura de cortinas e persianas e a temperatura a ser mantida pelos condicionadores de ar pode ser controlada por meio destes equipamentos, por meio de comandos de voz. A análise de comentários em redes sociais para identificar tendências de mercado, com opiniões sobre produtos, ou de posicionamento político, também faz uso do processamento de linguagem natural.

A percepção de máquina se relaciona a capacidade de traduzir os sinais de sensores como câmeras e microfones em aspectos do mundo real. Antes processar a linguagem, isto é, extrair o significado das palavras, os assistentes inteligentes anteriormente mencionado necessitam converter os sons recebidos do ambiente em palavras conhecidas. O reconhecimento de imagens é outra aplicação, e está

[10] Alan Turing, matemático e cientista inglês foi pioneiro nos ramos da computação e inteligência artificial e participou do desenvolvimento dos primeiros computadores modernos, no século XX.

presente desde a recomendações para “etiquetagem” de fotos em redes sociais até os sistemas de vigilância em massa, utilizados em muitos países, com capacidade de reconhecimento facial. O campo da percepção de máquina é uma das tecnologias centrais para a evolução do desenvolvimento de carros autônomos, como os fabricados pela companhia Tesla. A percepção ambiental é crítica para a implementação de um sistema autônomo de direção e por isso está entre os objetivos principais da indústria automotiva de ponta (YURTSEVER e colab., 2020).

O aprendizado de máquina (AM) ou aprendizado automático (AA) é outro campo muito vasto da inteligência artificial e consiste em alimentar um programa de computador com informação previamente coletada para que, a partir daí, o programa possa tomar decisões sobre dados futuros. O objetivo é que, a partir de exemplos, o programa atinja a capacidade de generalização para decidir. Em outras palavras, os programas devem aprimorar sua capacidade de tomar decisões à medida que acumulam experiência.

Alguns exemplos de problemas clássicos envolvendo aprendizado de máquina e que já foram extensivamente estudados são a classificação, regressão, ranqueamento, agrupamento, redução dimensional (MOHRI e colab., 2018).

A **classificação** consiste em associar um item a uma determinada categoria. Animais podem ser classificados quanto à espécie. Cachorros podem ser classificados quanto à raça. Páginas na internet e documentos podem ser classificados como entretenimento, notícias, esportes. No caso específico de tráfego de rede, uma classificação pertinente seria a de tráfego “normal” em oposição a de tráfego “malicioso” ou anômalo. O tráfego malicioso pode ainda ser dividido em subcategorias tais como escaneamento, negação de serviço, ataques contra logins e senhas, ataques a aplicações, etc.

A **regressão** busca determinar o valor futuro que uma determinada variável irá possuir. Isso pode abranger valores de ações na bolsa ou a temperatura ao longo do ano. No contexto das redes de computadores, pode-se estimar o volume de tráfego esperado ou o número de requisições a uma página de Internet em um determinado dia da semana. Uma particularidade da regressão é que ela permite auferir o erro na estimativa em comparação com o valor concreto que ocorre.

O **ranqueamento** é a tarefa de ordenar itens de um conjunto com base em algum critério. Um exemplo bastante comum é o ordenamento de resultados de

uma busca na Internet com base na “relevância” da resposta. Nesse contexto, a definição de relevância será umas das principais características que define o próprio motor de busca utilizado. Na área de cibersegurança, algoritmos de ranqueamento são utilizados para classificar mensagens de e-mail como seguras, spam ou mesmo ataques do tipo *phishing*[11]. Uma outra classificação é determinar a severidade de um ataque detectado por um IDS para facilitar a sua priorização pela equipe de segurança.

O **agrupamento** ou *clustering* se refere a dividir amostras em subconjuntos que apresentam alguma similaridade. Essa abordagem é apropriada para cenários nos quais não se conhecem categorias de antemão. Em redes sociais, é usada para identificar comunidades entre grandes grupos. Em cibersegurança, o agrupamento pode ser utilizado para análise forense, análise de malware e detecção de anomalias (LAKSHMANARAO e SHASHI, 2020).

A **redução dimensional** busca simplificar as amostras de dados a serem analisadas reduzindo a quantidade de dimensões, isto é, de variáveis. Isso é feito por meio de um tratamento estatístico que identifica as variáveis mais relevantes de uma amostra e também e variáveis dependentes, com alto índice de correlação a outras, e que podem ser suprimidas. Isso permite reduzir a quantidade de informação a ser processada e também o tempo e espaço de armazenamento necessários para esse processamento.

Essa abordagem de redução de dimensões se mostra bastante pertinente para o contexto da cibersegurança tanto devido ao grande volume de dados quanto pela complexidade e variedade das informações processadas. Quando consideramos os dados de redes IP, cada pacote IP na versão 4 (IPv4) possui 14 campos no seu cabeçalho que representam variáveis distintas. Um pacote IP na versão 6 (IPv6) possui ao menos 8 campos de cabeçalho. Quanto aos principais protocolos da camada de transporte, UDP e TCP, vemos que o primeiro tem 4 campos de cabeçalho e o segundo possui 10 campos. Quando avançamos para os protocolos de aplicação, como DNS e HTTP, cada um possui seu conjunto particular de atributos. Todos esses protocolos se combinam no tráfego de rede e a análise se mostra complexa, levando em conta somente a quantidade de variáveis envolvida.

[11] Ataques de *phishing* usam mensagens de e-mail forjadas para fazer o usuário acessar sites falsos e roubar credenciais, ou para que este execute programas maliciosos.

As soluções para tratar problemas de aprendizado de máquina se dividem em dois tipos principais baseados nas entradas que se fornece e nas saídas geradas: aprendizado supervisionado e não supervisionado. No aprendizado supervisionado, são fornecidos dados previamente classificados e cada amostra está associada a uma etiqueta. O algoritmo definirá os parâmetros que associam a amostra a uma etiqueta. Esse cenário é típico para problemas de classificação, regressão e ranqueamento (MOHRI e colab., 2018).

No aprendizado não supervisionado, os dados de treinamento não têm uma classificação prévia. Caberá ao algoritmo buscar similaridades e definir grupos baseados nelas. Este cenário é típico de problemas de agrupamento e redução dimensional (MOHRI e colab., 2018).

Um campo de aprendizado de máquina que merece destaque é o das redes neurais artificiais, que usa modelos matemáticos inspirados no comportamento de neurônios no sistema nervoso e são apropriados para o reconhecimento de padrões.

Algoritmos de aprendizado de máquina em geral precisam de dados para serem treinados. Para a pesquisa e para comparar os resultados obtidos por diferentes algoritmos, é interessante que haja conjuntos de dados (*datasets*) públicos e de livre utilização. Para algoritmos de classificação, por exemplo, o conjunto de dados de flores Iris, publicado em um artigo de 1936, se tornou quase que um padrão de fato para avaliações e testes. Há conjuntos de dados disponíveis para diversos campos de pesquisa, como reconhecimento de imagens, facial, de voz, dados físicos, químicos e biológicos.

Com a demanda crescente por aplicações que empreguem recursos de inteligência artificial para obter melhores resultados, e de modo similar ao que ocorre no ramo de *Big Data*, se observa uma oferta crescente de soluções acessíveis a pesquisadores e ao grande público.

O projeto scikit-learn surgiu dentro da iniciativa da empresa Google chamada *Google Summer of Code*, que busca atrair e fomentar o interesse por linguagens de programação e desenvolvimento de software em estudantes. A primeira versão de scikit-learn como um projeto independente foi disponibilizado em 2010 e alcançou grande popularidade entre pesquisadores de AA e foi utilizada por empresas como Spotify, Booking.com, Evernote.

A empresa Google oferece a biblioteca de aprendizado automático TensorFlow desde 2015, e em poucos anos ela ganhou grande popularidade. A biblioteca TensorFlow é voltada para o processamento de AA sobre grandes volumes de dados, usando linguagens de programação Python e C++, e permite a aceleração da execução usando placas de vídeos de alto desempenho. Essa biblioteca ganhou grande atenção na indústria e é utilizada por grandes empresas como Coca Cola, GE, Intel, Twitter, Lenovo, PayPal e a própria Google.

Outra biblioteca que tem alcançado crescente utilização entre pesquisadores e desenvolvedores é Pytorch, lançada pelo Laboratório de pesquisas de Inteligência Artificial da empresa Facebook. Pytorch guarda algumas similaridades com TensorFlow, como o uso de Python e C++ bem como a capacidade de aceleração com placas de vídeo.

Por último, a Fundação Apache mantém o módulo MLlib, para emprego de aprendizado de máquina em conjunto com o motor Spark de processamento de *Big Data* citado anteriormente. MLlib traz uma grande variedade de algoritmos e permite o uso de diversas linguagens de programação distintas, incluindo Java, Python e R.

O desenvolvimento de bibliotecas de aprendizado de máquina patrocinadas por empresas bilionárias do porte de Google e Facebook é um indicador de que a demanda na indústria e na academia por soluções que empregam Inteligência Artificial não se resume aos campos da cibersegurança.

5 SISTEMAS COLABORATIVOS

Um Sistema Colaborativo é um no qual múltiplos usuários ou agentes abordam uma tarefa compartilhada, normalmente a partir de posições remotas (FARLEY, 1998). Nesse contexto, uma das possíveis aplicações permite o estabelecimento de um sistema de computação distribuído com capacidade de processamento várias vezes maior que a de um sistema ou agente isolado possui. Para isso, o uso de redes de dados de alta velocidade é um facilitador e técnicas de programação específicas são necessárias.

Sistemas colaborativos que envolvem interação humana são bastante comuns na atualidade. O aplicativo *Waze*, criado em 2006, integra um sistema de navegação por GPS com uma plataforma de colaboração na qual cada dispositivo pode alimentar o aplicativo com o status do tráfico da via na qual o dispositivo se encontra, bem como a existência de obras e acidentes. A informação passada ao aplicativo é replicada para todos os usuários, gerando uma grande rede na qual todos recebem os dados e todos podem informar ou atualizar informações. Isso ampliou a capacidade de indicar o melhor caminho pelo aplicativo, pois esse passou a levar em conta não somente a distância e a velocidade máxima permitida na via, mas também outros fatores relevantes, como bloqueios e engarrafamentos. O aplicativo se tornou um competidor de peso e foi comprado pela Google em 2013.

Outro exemplo de sistemas colaborativos com interação humana são as plataformas de produtividade de equipes como Trello e Slack.

Quando passamos para os sistemas colaborativos sem interação humana, encontramos exemplos de aplicações que usam sensores de diferentes tipos. Sensores de tráfico podem ser distribuídos pelas vias principais de uma cidade e com isso otimizar o funcionamento dos semáforos, redução da poluição urbana, redução do consumo de combustível (KAFI e colab., 2013) e informar as pessoas sobre tempos de espera em paradas de ônibus de cidades inteligentes (EL MRINI, ANASS e GHACHAM AMRANI, ABDELLATIF, 2018). Na agricultura, sensores de umidade e temperatura no solo podem alimentar um sistema de irrigação automatizado, otimizando o uso de água e energia elétrica, além de irrigar no momento mais apropriado, o que aumenta a eficiência da plantação (KANSARA e colab., 2015).

Além dos exemplos de sensores já mencionados, existem outros tipos a exemplo de sensores de proximidade, pressão, nível de líquidos e grãos, ópticos,

infravermelhos. Atualmente, se pode afirmar que qualquer sensor é passível de receber capacidade de Internet das coisas (*Internet of Things* – IoT) e passar a integrar um sistema colaborativo ligado em rede. Essa tendência será potencializada com chegada das redes de telefonia de 5ª geração (5G), que têm requisitos de tempos de resposta e latência reduzidos em relação às redes atuais.

No contexto da cibersegurança, o IDS pode ser considerado um sensor ou um conjunto de sensores atuando em uma ou mais redes. A partir dessa analogia, os desafios para um sistema IDS colaborativo se mostram similares aos encontrados em outros sistemas colaborativos, nos quais se destaca a comunicação e o desempenho do sistema como um todo.

6 **PROPOSTAS ESTUDADAS**

6.1 IDS USANDO BIG DATA

Big data é atualmente um dos maiores desafios para a Detecção de Intrusão (ZUECH e colab., 2015) e tem sido assim desde a popularização da Internet. Atualmente, no Brasil, há conexões de Internet para instalação em residências com velocidades que atingem 300 ou 400 Mbps. Em outros países, já há opções de planos de Internet residencial com velocidades de 1Gbps ou até mesmo 10Gbps (VAUGHAN-NICHOLS, 2015). Tudo isso ilustra a capacidade de geração de tráfego nas redes atuais. Na era da *big data*, o desafio para o IDS é ser eficiente para processar transmissões de altíssima velocidade (um “V” da conceituação de *big data*), em tempo real, sem perder nenhum fluxo de pacotes relevante (RATHORE e colab., 2016).

Quando tratamos de captura de dados de rede para análise por IDS, um padrão “de fato” usado em todo mundo é o formato pcap (de *packet capture* – captura de pacotes), que permite armazenar os dados transmitidos por redes em arquivos binários. O formato pcap é suportado tanto por soluções de IDS livres e/ou de código aberto como por produtos comerciais. A biblioteca utilizada para isso tem o mesmo nome, libpcap, é fruto do projeto do programa de captura de pacotes tcpdump (THE TCPDUMP GROUP, 2020), e tem implementações para a maioria dos sistemas operacionais usados atualmente, incluindo Unix, Linux, BSD, Mac OS e Windows (NMAP PROJECT, 2020). Como os arquivos pcap são representações fidedignas dos dados enviados e recebidos pelas interfaces de rede, o tamanho dos arquivos gerados cresce na mesma proporção da largura de banda dos links de rede gerando um grande volume de dados (o segundo “V”). Uma alternativa que pode complementar a análise de dados de captura fidedignos consiste na análise de dados do tipo “flow” (SPEROTTO e colab., 2010), a exemplo do padrão NetFlow da empresa Cisco ou Jflow da empresa Juniper, esses padrões consistem em alguns campos dos cabeçalhos dos pacotes e datagramas de rede somados a alguns metadados.

Para pode processar esse enorme volume de dados, o primeiro desafio que se apresenta para um IDS é o de armazenar esses dados. Neste cenário, o armazenamento centralizado se torna cada vez mais difícil e a solução é buscar alternativas que implementem formas de armazenamento distribuído. Além disso,

o uso de bancos de dados relacionais se mostra ineficiente e sofre com dificuldade de escalabilidade (ZUECH e colab., 2015).

Na literatura recente, foram encontradas propostas de IDS utilizando *Hadoop*, e há exemplos de trabalhos no aprimoramentos de IDS já consagrados, como Snort, utilizando *Hadoop* (PRATHIBHA e DILEESH, 2013). O objetivo principal do trabalho de Prathiba foi integrar Snort e *Hadoop* e avaliar o desempenho em cenários de carregamento e de acesso aos dados, utilizando arquivos pcap de diferentes tamanhos, e comprovando a viabilidade da solução apresentada. Nesta solução, o MapReduce é usado para análise dos pacotes e Hive é usado para indexação e consultas *adhoc*.

Outros trabalhos propõem desenhos de IDS inteiramente novos tendo *Hadoop* como o sistema de armazenamento de *big data*, a exemplo de (BANDRE e NANDIMATH, 2015) e (RATHORE e colab., 2016). Bandre apresenta uma solução de NDIS que usa *Hadoop* como solução de armazenamento e executa o processamento e detecção por meio da arquitetura de programação paralela de alto desempenho CUDA, que encaminha os processos para placas gráficas de propósito geral (*General Purpose Graphic Processing Units* - GPGPU), com foco em escalabilidade, desempenho e otimização.

O uso da linguagem CUDA por si só é bastante relevante, pois permite que o processamento seja realizado de modo centralizado ou distribuído. A linguagem CUDA é proprietária da empresa Nvidia e permite a distribuição de processamento em uma ou mais GPGPU e facilita o desdobramento de *clusters* para tarefas de processamento de altíssimo desempenho. Entre os clusters, de maior capacidade de processamento, vários utilizam GPGPU da Nvidia (TOP500.ORG, 2020). Uma alternativa para a linguagem CUDA que não restringe o hardware à uma única empresa é linguagem OpenCL[12], com finalidade similar à CUDA, mas de código aberto.

Rathore, por sua vez, descreve um IDS que usa *Hadoop* na camada de armazenamento, enquanto a detecção é realizada usando o MapReduce e uma série de métodos de aprendizado de máquina executados de modo distribuído por meio de Spark. Essa solução busca um IDS que opera em ambiente de *big data*

[12] OpenCL (Open Computing Language) é uma linguagem de programação paralela que permite que a execução dos programas seja dividida em diferentes dispositivos (processadores e placas de vídeo) ou mesmo arquiteturas distintas.

em redes de alto desempenho, mas mantendo a capacidade de processamento em tempo real, ao mesmo tempo que alcança elevada capacidade de detecção, com baixas taxas de falsos positivos e falso negativos.

Outra questão relevante para a Detecção de Intrusão em ambientes de *big data* é a fusão de dados heterogêneos. De modo simplificado, a fusão de dados consiste em obter sentido a partir de dados com diferentes fontes que comumente tem estruturas distintas (ZUECH e colab., 2015). Isso pode ocorrer quando se deseja integrar alertas de sensores IDS de fabricantes diferentes, fontes diversas como *logs* de servidores, *firewalls*, *switches*, roteadores e outros ativos de redes, alertas de monitoramento SNMP (*Simple Network Management Protocol*), e NetFlow em um resultado que se aproxima do conceito de SIEM (*Security information and event management*) utilizado por alguns fabricantes. Nesse caso estamos abordando o terceiro “V” de *big data* (“Variedade”). Uma vantagem de integrar essas outras fontes ao IDS é prover o IDS de maior visibilidade sobre a rede a ser protegida, que resulta em uma maior consciência situacional e em última instância em maior precisão dos alertas gerados.

Zuech apresenta diversos estudos de IDS que agregam fontes heterogêneas, com abordagens variadas, tanto centralizadas como distribuídas. DShield também é citado como um exemplo de agregador de fontes heterogêneas.

6.2 IDS USANDO INTELIGÊNCIA ARTIFICIAL

Observa-se que maioria dos trabalhos recentes envolvendo pesquisas sobre IA e IDS abordam o uso de técnicas de aprendizado de máquina, envolvendo diferentes tipos de algoritmos, e também técnicas de redução dimensional.

Um exemplo é o trabalho desenvolvido por Akashdeep, que propõe um método de detecção integrando um subsistema de ranqueamento e seleção de variáveis a uma rede neural para detecção de anomalias num IDS (AKASHDEEP e colab., 2017). Outra proposta encontrada consiste no uso de Apache Spark processando algoritmos de AA e redução dimensional, integrando o aprendizado de máquina em um ambiente de *big data* (OTHMAN e colab., 2018).

Um dos primeiros desafios para a pesquisa do uso de IA combinada com sistemas IDS é a dificuldade para encontrar conjuntos de dados (*datasets*) apropriados. A principal causa para essa dificuldade decorre de considerações sobre privacidade, o que faz com que empresas e instituições sejam relutantes em

disponibilizar dados de capturas de tráfego de rede reais. Disso resulta que muitos *datasets* contam com dados de redes simuladas ou virtualizadas ou sofrem um processo de anonimização, pelo qual os dados úteis (*payload*) e campos dos cabeçalhos tais como endereços são removidos dos pacotes de dados.

Outra consequência é que, ainda nos dias atuais, um dos *datasets* mais utilizados em pesquisas envolvendo IA e IDS é o chamado KDD 99 ou KDDCUP'99 (RING e colab., 2019). Esse conjunto de dados foi criado pela agência dos Estados Unidos DARPA para uma competição de mineração de dados em 1999. KDD 99 não está em formato pcap nem em algum formato do tipo flow, e contém dados e ataques de rede que hoje são considerados bastante ultrapassados, uma vez que os dados e ataques usados na Internet atualmente diferem muito do que era comum há mais de duas décadas. Outro questionamento em relação a esse conjunto é o grande número de dados redundantes e a baixa dificuldade que ele oferece para os algoritmos (TAVALLAEE e colab., 2009). Apesar disso, esse *dataset* ainda é encontrado em diversos trabalhos recentes.

Há iniciativas para romper essa situação e, nos últimos anos, foram publicados diversos *datasets* com dados padronizados do tipo pcap, flow ou uma combinação de ambos, e com capturas de ambientes de rede mais atuais e realistas. No entanto, nenhum deles conseguiu obter a mesma abrangência alcançada pelo *dataset* KDD 99. A revisão literária realizada indica a prevalência do uso do dataset KDD 99 em trabalhos relativamente recentes. Em uma levantamento realizado por (FARAH e colab., 2015), contatou-se essa afirmação para os trabalhos publicados entre 2009 e 2014. Em outro levantamento similar, onde foram estudados 39 trabalhos que usam AA sobre dados de cibersegurança, constatou-se que 28 empregaram o *dataset* KDD 99 ou similares do ano 2000 ou mesmo 1998. Isso é atribuído à dificuldade e ao tempo despendido para se encontrar *datasets* representativos (BUCZAK e GUVEN, 2016). Nesse sentido, os trabalhos citados anteriormente, de Akashdeep e Othman, também fazem uso dos *datasets* KDD 99.

Um tentativa de simplificar o trabalho dos pesquisadores interessados em AA aplicado à cibersegurança foi encontrada no trabalho de (RING e colab., 2019), onde são levantados e classificados 35 *datasets*. A classificação inclui o tipo de captura (pcap, flow e outros), o ano e a duração da captura, e o tamanho dos *dataset*. Observa-se que há exemplos variando de alguns kilobytes a 250 gigabytes

de informação capturada. Além disso, há a indicação sobre a disponibilidade e publicidade dos dados para uso geral. Após a análise realizada, Ring recomenda o uso de 4 *datasets* particulares, criados entre 2015 e 2017, que serem abrangentes e relativamente recentes, são apropriados para utilização geral (RING e colab., 2019).

Merece destaque o trabalho recente de Kanimozhi, que usa a biblioteca scikit-learn para propor um IDS baseado em redes neurais. Seu diferencial é a validação usando um *dataset* realista e atual, de 2018, fornecido pelo Instituto Canadense para Cibersegurança (*Canadian Institute for Cybersecurity*). Esse *dataset* constitui uma versão mais atual de um dos 4 recomendados por Ring. Na sugestão para trabalhos futuros, Kanimozhi sugere a migração para um ambiente de AA mais robusto baseado em TensorFlow, bem como a integração com Apache Spark (KANIMOZHI e PREM JACOB, 2019).

6.3 SISTEMAS IDS COLABORATIVOS

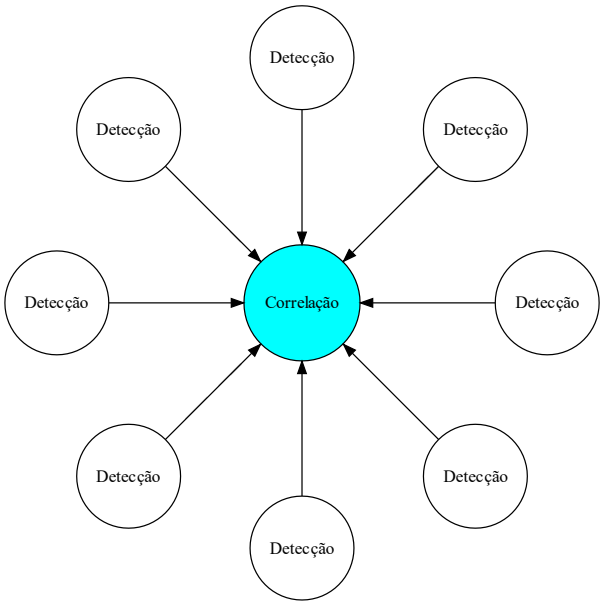
Ao integrar a análise de vários sistemas IDS operando em redes diferentes, é possível obter o IDS colaborativo (*Collaborative Intrusion Detection System* - CIDS), que potencializa a capacidade de detecção. Um CIDS tem a capacidade de detectar ataques que ocorrem ao longo de toda Internet simultaneamente, ao fazer a correlação de assinaturas entre diferente sub-redes da Internet (ZHOU e colab., 2010).

Um CIDS se subdivide em dois subsistemas principais: O Subsistema de Detecção, composto por diferentes sensores IDS (cada um monitorando sua própria rede) e o Subsistema de correlação, que integra os resultados dos diferentes sensores do subsistema de detecção e gera um resultado conjunto. Para o funcionamento do subsistema de correlação há três abordagens ou arquiteturas diferentes: centralizada, hierárquica e totalmente distribuída (ZHOU e colab., 2010).

A abordagem centralizada concentra as informações de todos os sensores em um único servidor de correlação e tem a vantagem de ser mais simples de instalar e operar. No entanto, o servidor de correlação se converte em um ponto único de falha e sua interrupção irá comprometer o funcionamento de todo o CIDS. A **Figura 1** ilustra um CIDS com arquitetura centralizada. Um exemplo de CIDS é o sistema DShield (ISC, 2020), mantido pelo *Internet Storm Center* do Instituto

SANS (*SysAdmin, Audit, Network and Security*). O CIDS DShield funciona desde novembro de 2000 e aceita arquivos de *logs* de *firewall* do mundo inteiro, gerando alertas quando detecta ameaças se alastrando ao longo da Internet, como por exemplo o *Worm*[13] de sequestro de dados WannaCry, que infectou cerca de 300 mil sistemas em 150 países em 2017 (AKBANOV e colab., 2019). As organizações participantes do DShield recebem e-mails com alertas antecipados para potenciais ameaças e os alertas gerados podem facilmente ser integrados aos sistemas de monitoramento de rede das organizações, como Nagios e Zabbix.

Figura 1 – CIDS com arquitetura centralizada



Fonte: O autor

Outros exemplos de CIDS são DIDS (SNAPP e colab., 1991) e NetSTAT (VIGNA e KEMMERER, 1998). A abordagem centralizada se aplica bem a iniciativas de colaboração como DShield, mas as limitações apontadas trazem restrições a sua utilização em sistemas independentes na Internet (ZHOU e colab., 2010).

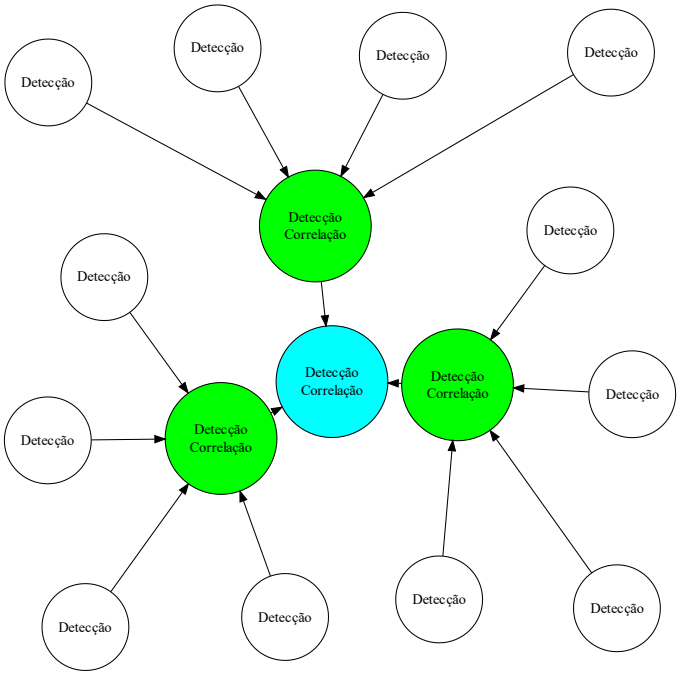
A abordagem hierárquica busca sanar as limitações da abordagem centralizada ao estabelecer uma estrutura de árvore para organizar e distribuir a correlação. Todos os IDS que são nós da árvore mantêm a função de detecção. Os IDS folhas (nós sem filhos) realizam apenas a detecção e todos os demais IDS

[13] Software malicioso como os vírus, mas com foco na replicação automática ao longo de redes de computadores. A internet propiciou que este tipo de software consiga obter alcance mundial em suas infecções.

(nós intermediários) realizam também a correlação. Com isso, é possível organizar o CIDS em subgrupos baseados em fatores com proximidade geográfica, controle administrativo, similaridade de plataformas de software e tipos de ataques esperados (ZHOU e colab., 2010).

Como exemplos de CIDS usando abordagem hierárquica, podemos citar GrIDS (*Graph-Based Intrusion Detection System*) (STANIFORD-CHEN e colab., 1996) e EMERALD (PORRAS e NEUMANN, 1997). A abordagem hierárquica tem maior escalabilidade que a abordagem centralizada. No entanto, os nós dos níveis mais altos ainda são um fator que limita a escalabilidade, e sua falha pode interromper o funcionamento de sua subárvore (ZHOU e colab., 2010). Outra limitação pode ocorrer devido ao fato de os nós mais elevados terem menor capacidade de detecção devido ao resumo das informações sobre as subárvores nos nós intermediários. A **Figura 2** ilustra um CIDS com arquitetura hierárquica.

Figura 2 – CIDS com arquitetura hierárquica

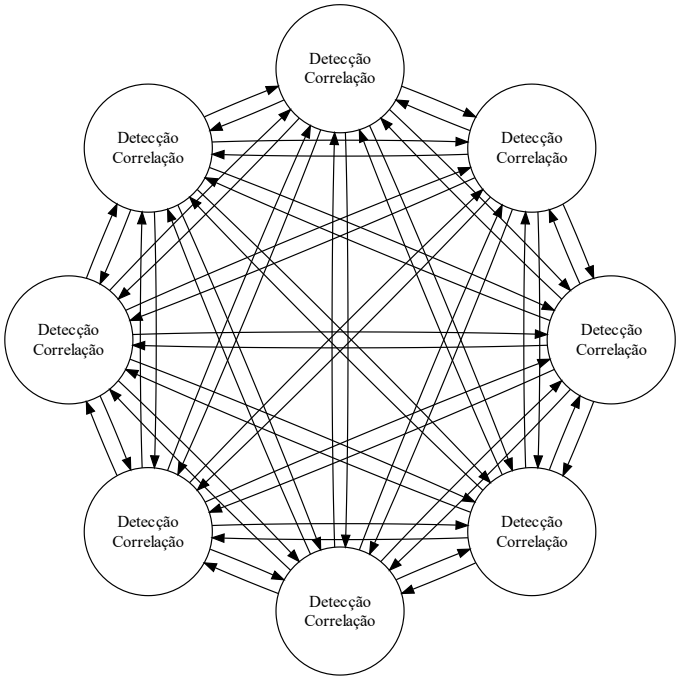


Fonte: O autor

A abordagem totalmente distribuída busca sanar os problemas e limitações das abordagens centralizada e hierárquica utilizando protocolos de transmissão

peer to peer (P2P)[14] para coordenar a detecção e correlação. Dessa forma, busca-se obter um CIDS sem pontos únicos de falha que comprometam todo o sistema, obtendo assim uma maior resiliência, ao mesmo tempo alcançando uma maior escalabilidade. Os desafios para a abordagem totalmente distribuída consistem em minimizar a sobrecarga de comunicação necessária para a operação do CIDS, ao mesmo tempo que evita a perda de precisão de detecção resultante do resumo das informações transmitidas. Além disso, essa abordagem também deve equilibrar a carga de trabalho destinada a cada IDS participante. A **Figura 3** ilustra um CIDS com arquitetura totalmente distribuída.

Figura 3 – CIDS com arquitetura totalmente distribuída



Fonte: O autor

Sharma realizou um estudo comparativo entre um CIDS de arquitetura distribuída e outro CIDS de arquitetura centralizada, em um cenário de redes de computação de borda típico das futuras redes celulares de 5ª Geração. As conclusões do trabalho indicam uma maior capacidade de detecção do CIDS centralizado, ao custo de maior sobrecarga na rede e uma latência que cresce à medida que nós de detecção são adicionados na rede (SHARMA e colab., 2020).

[14] Protocolos P2P são usados em aplicações distribuídas, para evitar um ponto único de falha que comprometa o funcionamento global da aplicação. O termo ganhou visibilidade na década de 1990 com a popularização de softwares de distribuição de arquivo, como Napster e BitTorrent.

Após definir a arquitetura do CIDS, outra questão é escolher o mecanismo de correlação de alertas a ser utilizado. Há quatro grupos de técnicas para essa correlação: baseadas em similaridade, baseadas em cenários de ataque, baseadas em filtros e de múltiplos estágios (ZHOU e colab., 2010). Nos mecanismos de correlação baseados em similaridades, comumente utiliza uma função para determinar o grau de similaridade entre diferentes alertas recebidos para determinar se estes serão correlacionados ou não.

Para os mecanismos baseados em cenários de ataque, diferentes maneiras de caracterizar estes cenários podem ser utilizadas, utilizando desde métodos probabilísticos até máquinas de estados. Este método de correlação se assemelha à detecção baseada em assinaturas no contexto de CIDS, e se mostra mais eficiente para detectar tipos de ataques previamente conhecidos em comparação à ataques inéditos. As técnicas de múltiplos estágios buscam resolver essa deficiência com técnicas para inferir relações de causa e efeito entre alertas quaisquer e assim detectar novos ataques, mas tendem a gerar mais falsos negativos que os métodos anteriores.

Já os mecanismos de correlação baseado em filtros procuram aprimorar sua eficiência buscando um maior conhecimento sobre as redes a serem protegidas, isto é, armazenando informações como os sistemas operacionais e tipos de ativos de rede que operam nestas redes. Desta forma, é possível descartar alertas que não tem potencial para causar impacto. Por exemplo, alertas para ataques a serviços de rede que não estão em execução ou alertas para ataques a sistemas Windows em uma rede onde há somente sistemas Linux ou BSD. Esse tipo de mecanismo de correlação exige um grande esforço de configuração e atualização para ser eficiente.

A questão da privacidade de dados também se mostra relevante para a implementação de CIDS, uma vez que as organizações são relutantes em compartilhar informações que podem revelar dados sensíveis sobre suas redes e seus usuários (HERNANDEZ-ARDIETA e colab., 2013). O que no passado era uma questão cultural agora se converte em questões legais, principalmente após a aprovação de leis como a Lei Geral de Proteção de Dados Pessoais (BRASIL, 2018) - LGPD, brasileira, e o Regulamento Geral sobre a Proteção de Dados (UNIÃO EUROPEIA, 2016), conhecido por GDPR por sua sigla em inglês, da União Europeia. Essa leis naturalmente irão tornar as organizações mais resistentes a

compartilhar informações, uma vez que preveem sanções no caso de ocorrerem perdas ou exposições indevidas de dados de usuários (LI e colab., 2019).

Para os cenários de CIDS, foram propostas diferentes técnicas de filtragem e anonimização das informações consideradas sensíveis. Métodos baseados em filtros de Bloom (BLOOM, 1970) foram encontrados em mais de uma fonte na literatura (ZHOU e colab., 2010).

Uma última questão relevante para a implementação de CIDS diz respeito à padronização dos formatos de dados e protocolos para de troca de informações de alertas, algo que viria a favorecer a interoperabilidade entre equipamentos de fabricantes diferentes. Diversas propostas têm sido realizadas nesse sentido, e merece destaque a iniciativa "*Making Security Measurable* - MSM" (MITRE, 2013), da organização MITRE. MSM compreende uma ampla arquitetura para o gerenciamento de cibersegurança. Segundo Hernandez-Ardieta, os padrões MSM podem ser agrupados em seis áreas do conhecimento, cada uma relacionada a um processo:

- A – Definição de ativos – Processo de inventário
- C – Guias de configuração – Processo de análise
- V – Alertas de vulnerabilidades – Processo de análise
- T – Alertas de ameaças – Processo de análise
- I – Indicadores de risco/ataque – Processo de detecção de intrusão
- R – Relatórios – Processo de gerenciamento

A **Tabela 2** ilustra os 23 padrões do MSM e as áreas do conhecimento que cada um abrange.

Tabela 2 – Padrões MSM e áreas do conhecimento

	CPE	OVAL	SWID	XCCDF	CCE	OCIL	CCSS	CVE	CWE	CVSS	CAPEC	CVRF	MAEC	CybOX	IndEX	STIX	IODEF	CPE	CEE	RID	RID-T	CYBEX	CWSS
A	●	●	●	●														●					
C		●		●	●	●	●																
V		●						●	●	●		●											
T								●	●	●	●		●	●	●	●	●		●	●	●		
I	●							●					●	●	●	●	●	●	●	●	●		
R	●	●			●			●	●	●			●		●	●	●			●	●	●	●

Fonte: (HERNANDEZ-ARDIETA e colab., 2013)

A pesquisa por implementações efetivas de CIDS é um assunto atual e os trabalhos mais recentes agora trazem conceitos que ganharam relevância nos últimos anos, tais como o uso de tecnologias de *blockchain*[15], computação de borda, redes 5G, Internet das coisas e computação em nuvem.

[15] A tecnologia de *blockchain* foi desenvolvida em 2008 e consiste em um "cartório virtual" distribuído, que usa criptografia forte para obter elevada resistência a alterações não autorizadas em seus registros.

7 PROPOSTAS INTEGRADAS

Ao longo da revisão bibliográfica, observou-se que os conceitos de sistemas colaborativos, inteligência artificial e *big data* aplicados a pesquisas recentes sobre IDS aparecem combinados com certa frequência. Essa combinação não se restringe ao domínio da cibersegurança, uma vez que plataformas de uso geral, como Spark, podem ser usadas sobre qualquer tipo de dado e assim mesmo trazem entre seus módulos principais uma robusta biblioteca de aprendizado de máquina. Da mesma forma, os ambientes de processamento de *Big Data* buscam atingir a escalabilidade por meio do uso de computação distribuída, tanto para armazenamento quanto para o processamento dos dados.

No caso particular dos IDS, o problema da detecção de anomalias tem gerado uma demanda crescente por soluções inovadoras, e encontrou na inteligência artificial uma possibilidade promissora. Ao mesmo tempo, o grande volume de dados a ser analisado faz com que as soluções de *Big Data* para uso geral sejam apropriadas para uso nesse contexto.

Entre os trabalhos estudados, o trabalho de Rathore usa algoritmos de AA implementados em Spark e MapReduce para atingir o desempenho pretendido. O levantamento de Zuech, citados no capítulo de *Big Data*, faz extensas considerações a respeito dos desafios do aprendizado de máquina aplicado sobre *Big Data*, incluindo o problema da baixa disponibilidade de *datasets* de boa qualidade e a necessidade de redução dimensional.

Ressalta-se que há pesquisas sobre IDS aplicando AA isoladamente. No entanto, quando os trabalhos consideram cenários de *big data*, há uma forte tendência de trazer conceitos de AA atrelados.

Outra tendência é a integração de IDS colaborativo com *big data*. A bibliografia revela a tendência de se utilizar CIDS integrando diversas fontes, que podem mesmo ser outros IDS de host e de rede, espalhando ao longo de redes de ampla cobertura. Zuech também aborda a questão dos dados heterogêneos vindo de diversas fontes.

Por último, observou-se que essa tendência de integração se acentua nos trabalhos mais recentes, o que reforça a constatação da convergência das três áreas de conhecimento.

8 CONCLUSÃO

A pesquisa realizada e os trabalhos recentes encontrados na revisão mostram que a popularização de ferramentas de *Big Data* favoreceu a pesquisa de IDS com capacidades para tratar grandes volumes de dados e sob altas taxas de transmissão. Para essas ferramentas, os dados de cibersegurança não representam um desafio de per si, conforme mostraram os trabalhos estudados, que atestam a possibilidade de aplicar as soluções já existentes no mercado para realizar o processamento de IDS de alto desempenho, distribuídos ou não, e podendo inclusive utilizar arquiteturas como CUDA, o que viabiliza a criação de *clusters* de processamento escaláveis com capacidade de processar quantidades de dados extremas.

Enquanto ferramentas de *Big Data* resolvem os problemas de escala existentes no cenário de cibersegurança atual, apenas replicar o que é feito no IDS clássico já não é suficiente, pois este mesmo já mostra deficiências que se acentuaram nos últimos anos. A detecção de ameaças nos IDS por meio de assinaturas deve ser complementada por um meio automático e eficiente de detectar anomalias. Essa tarefa está sendo delegada ao campo do aprendizado de máquina.

Dezenas de trabalhos foram publicados nas últimas duas décadas tratando sobre o uso de aprendizado de máquina e outras técnicas de inteligência artificial para aprimorar a detecção de sistemas de detecção de intrusão, e à medida que a quantidade de ataques juntamente com a percepção de insegurança no contexto da cibersegurança aumentam, a demanda por soluções inovadoras para a detecção de anomalias nas redes continua crescente.

Uma dificuldade para atingir esse objetivo reside na histórica impossibilidade de validar as propostas teóricas utilizando *datasets* realistas e representativos. Em trabalhos atuais ainda se observa o uso de *datasets* de cerca de duas décadas para validar propostas. Nesse cenário, mesmo o atingimento de resultados excelentes pode não resultar em soluções eficientes quanto confrontadas com os dados da internet de hoje, que possuem um perfil muito distinto do que havia no século passado. Face a isso, há iniciativas de catalogar *datasets* mais atuais e relevantes, e também trabalhos que rompem com o *status quo* e validam suas propostas usando *datasets* novos. Tanto o levantamento bibliográfico realizado neste trabalho quanto estudos de mercado realizados por entidades como Gartner

indicam que esse campo de pesquisa se encontra em plena atividade, com necessidades de estudos e propostas que atendam à forte demanda proveniente das empresas, governos e da sociedade.

Em outra vertente, a tendência de utilização de IDS colaborativos e distribuídos atualmente é incrementada pelo crescimento das redes móveis de telefonia, incluindo a nova tecnologia 5G, da qual se espera que decorra a massificação da internet das coisas, e a popularização da computação de borda em oposição aos datacenters totalmente centralizados. Todos esses fatores indicam que os sistemas IDS tendem a ser cada vez mais formados por um conjunto de sensores de rede e host distribuídos, e os novos projetos de IDS nesse formato deverão levar em conta as questões relevantes que emergem nesse cenário, tais como a hierarquia e a sobrecarga de comunicação e processamento.

Por fim, concluímos que os conceitos de *Big Data*, Inteligência Artificial e Sistemas Colaborativos aplicados aos sistemas de detecção de intrusão se mostram na atualidade não apenas como *buzz words* utilizadas por empresas e seus departamentos de marketing para aumentar suas vendas, mas também tecnologias e linhas de pesquisa necessárias para o aprimoramento dos sistemas de segurança que suportam uma parte cada vez maior da tecnologia utilizada pela sociedade moderna.

REFERÊNCIAS

AKASHDEEP e MANZOOR, Ishfaq e KUMAR, Neeraj. **A feature reduced intrusion detection system using ANN classifier**. Expert Systems with Applications, 2017.

AKBANOV, Maxat e VASSILAKIS, Vassilios G e LOGOTHETIS, Michael D. **WannaCry ransomware: Analysis of infection, persistence, recovery prevention and propagation mechanisms**. Journal of Telecommunications and Information Technology, 2019.

APACHE SOFTWARE FOUNDATION. **Apache Hadoop**. Disponível em: <https://hadoop.apache.org/>. Acesso em: 4 ago 2020.

BANDRE, Sanraj Rajendra e NANDIMATH, Jyoti N. Design consideration of Network Intrusion detection system using Hadoop and GPGPU. 2015, [S.l: s.n.], 2015. p. 1–6.

BLOOM, Burton H. **Space/time trade-offs in hash coding with allowable errors**. Communications of the ACM, 1970.

BRASIL. **Doutrina Militar de Defesa Cibernética - MD31-M-07**. 1ª ed. [S.l: s.n.], 2014.

BRASIL. **Glossário de Segurança da Informação**. 2019. Disponível em: <http://www.in.gov.br/en/web/dou/-/portaria-n-93-de-26-de-setembro-de-2019-219115663>.

BRASIL. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. 2018. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>.

BRASIL. **Política Nacional de Defesa. Estratégia Nacional de Defesa**. Diário Oficial da União - Seção 1 - 26/9/2013, Página 1, 2012.

BUCZAK, A L e GUVEN, E. **A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection**. IEEE Communications Surveys & Tutorials, v. 18, n. 2, p. 1153–1176, 2016.

CHENG, Tsung Huan e colab. **Evasion techniques: Sneaking through your intrusion detection/prevention systems**. IEEE Communications Surveys and Tutorials, 2012.

CIMPANU, Catalin. **China is now blocking all encrypted HTTPS traffic that uses TLS 1.3 and ESNI**. Disponível em: <https://www.zdnet.com/article/china-is-now-blocking-all-encrypted-https-traffic-using-tls-1-3-and-esni/>. Acesso em: 15 set 2020.

DENNING, Dorothy E. **An Intrusion-Detection Model**. IEEE Transactions on Software Engineering, 1987.

DEWITT, D e STONEBRAKER, Michael. **MapReduce: A major step backwards**. The Database Column, 2008.

EL MRINI, ANASS e GHACHAM AMRANI, ABDELLATIF. **Wireless Sensors Network for Traffic surveillance and management in Smart Cities**. MATEC Web

Conf., v. 200, p. 24, 2018. Disponível em: <https://doi.org/10.1051/matecconf/201820000024>.

FARAH, Nutan e colab. **Application of Machine Learning Approaches in Intrusion Detection System: A Survey**. International Journal of Advanced Research in Artificial Intelligence, 2015.

FARLEY, Jim. **Java Distributed Computing**. [S.l.]: O'Reilly Media, Inc., 1998.

GARCÍA-TEODORO, P. e colab. **Anomaly-based network intrusion detection: Techniques, systems and challenges**. Computers and Security, 2009.

GRANCE, Tim e colab. **SP 800-35: Guide to Information Technology Security Services**. p. 1–84, 2003. Disponível em: <http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-35.pdf>.

HALILOVIC, Muhamed e SUBASI, Abdulhamit. **Intrusion Detection on Smartphones**. 2012.

HERNANDEZ-ARDIETA, Jorge L. e TAPIADOR, Juan E. e SUAREZ-TANGIL, Guillermo. Information sharing models for cooperative cyber defence. 2013, [S.l.]: IEEE Computer Society, 2013. p. 1–28.

HILBERT, Martin e LÓPEZ, Priscila. **The world's technological capacity to store, communicate, and compute information**. Science, 2011.

ISC, SANS. **ISC History and Overview**. Disponível em: <https://www.dshield.org/about.html>. Acesso em: 1 ago 2020.

JYOTHSNA, V. e V. RAMA PRASAD, V. e MUNIVARA PRASAD, K. **A Review of Anomaly based Intrusion Detection Systems**. International Journal of Computer Applications, 2011.

KAFI, Mohamed Amine e colab. A study of wireless sensor networks for urban traffic monitoring: Applications and architectures. 2013, [S.l: s.n.], 2013.

KANIMOZHI, V. e PREM JACOB, T. Artificial intelligence based network intrusion detection with hyper-parameter optimization tuning on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing. 2019, [S.l: s.n.], 2019.

KANSARA, Karan e colab. **Sensor based automated irrigation system with IOT: A technical review**. Int. J. Comp. Sci. Inf. Tech., 2015.

LAKSHMANARAO, A. e SHASHI, M. **A Survey On Machine Learning For Cyber Security**. International Journal of Scientific & Technology Research, v. 9, n. 1, p. 499–502, 2020. Disponível em: <http://www.ijstr.org/research-paper-publishing.php?month=jan2020>.

LI, He e YU, Lu e HE, Wu. **The Impact of GDPR on Global Technology Development**. Journal of Global Information Technology Management, v. 22, n. 1, p. 1–6, 2019. Disponível em: <https://doi.org/10.1080/1097198X.2019.1569186>.

LOHR, Steve. **The Origins of 'Big Data': An Etymological Detective Story**. The

New York Times, 2013.

MCAFEE, Andrew e BRYNJOLFSSON, Erik. **Big data: The management revolution**. Harvard Business Review, 2012.

MCHUGH, John. **Intrusion and intrusion detection**. International Journal of Information Security, 2001.

MITRE. **Making Security Measurable**. Disponível em: <https://measurablesecurity.mitre.org/>. Acesso em: 2 ago 2020.

MOHRI, Mehryar e ROSTAMIZADEH, Afshin e TALWALKAR, Ameet. **Foundations of Machine Learning**. 2ª ed. [S.l.]: MIT Press, 2018.

NMAP PROJECT. **Npcap: Windows Packet Capture Library & Driver**. Disponível em: <https://nmap.org/npcap/>. Acesso em: 5 ago 2020.

OTHMAN, Suad Mohammed e colab. **Intrusion detection model using machine learning algorithm on Big Data environment**. Journal of Big Data, v. 5, n. 1, p. 34, 2018. Disponível em: <https://doi.org/10.1186/s40537-018-0145-4>.

PORRAS, Phillip A e NEUMANN, Peter G. EMERALD: Event monitoring enabling response to anomalous live disturbances. 1997, [S.l: s.n.], 1997. p. 353–365.

PRATHIBHA, P G e DILEESH, E D. **Design of a hybrid intrusion detection system using snort and hadoop**. International journal of computer applications, v. 73, n. 10, 2013.

PUTANO, Ben. **A Look At 5 of the Most Popular Programming Languages of 2019**. Disponível em: <https://stackify.com/popular-programming-languages-2018/>. Acesso em: 25 ago 2020.

RAJU, Peddisetty Naga. **State-of-the-art Intrusion Detection: Technology, Challenges, and Evaluation.** 2005. 1–86 f. 2005. Disponível em: <http://liu.diva-portal.org/smash/record.jsf?pid=diva2:20134>.

RATHORE, M Mazhar e AHMAD, Awais e PAUL, Anand. **Real time intrusion detection system for ultra-high-speed big data environments**. The Journal of Supercomputing, v. 72, n. 9, p. 3489–3510, 2016. Disponível em: <https://doi.org/10.1007/s11227-015-1615-5>.

RING, Markus e colab. **A survey of network-based intrusion detection data sets**. Computers and Security. [S.l: s.n.]. , 2019

RUSSOM, Philip. **Big Data Analytics**. TDWI best practices report, fourth quarter, v. 19, n. 4, p. 1–34, 2011. Disponível em: <https://vivomente.com/wp-content/uploads/2016/04/big-data-analytics-white-paper.pdf>.

SHARMA, R e CHAN, C A e LECKIE, C. Evaluation of Centralised vs Distributed Collaborative Intrusion Detection Systems in Multi-Access Edge Computing. 2020, [S.l: s.n.], 2020. p. 343–351.

SNAPP, Steven R e colab. **DIDS (distributed intrusion detection system)-**

motivation, architecture, and an early prototype. 1991.

SPEROTTO, Anna e colab. **An overview of IP flow-based intrusion detection**. IEEE Communications Surveys and Tutorials, 2010.

STACKOVERFLOW. **StackOverflow Developer Survey**. 2019. Disponível em: <https://insights.stackoverflow.com/survey/2019>.

STANIFORD-CHEN, Stuart e colab. GrIDS-a graph based intrusion detection system for large networks. 1996, [S.l: s.n.], 1996. p. 361–370.

TAVALLAEE, Mahbod e colab. A detailed analysis of the KDD CUP 99 data set. 2009, [S.l: s.n.], 2009.

THE TCPDUMP GROUP. **TCPDUMP/LIBPCAP public repository**. Disponível em: <https://www.tcpdump.org/>. Acesso em: 5 ago 2020.

THUSOO, Ashish e colab. Hive - A petabyte scale data warehouse using hadoop. 2010, [S.l: s.n.], 2010.

TOP500.ORG. **June 2020 | TOP500**. Disponível em: <https://www.top500.org/lists/top500/2020/06/>. Acesso em: 4 ago 2020.

TOUCH, Joe e colab. **Service Name and Transport Protocol Port Number Registry**. Disponível em: <https://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.xhtml>. Acesso em: 9 ago 2020.

UNIÃO EUROPEIA. **Regulamento Geral de Proteção de Dados**. Jornal Oficial da União Europeia, 2016.

VAUGHAN-NICHOLS, Steven J. **Chattanooga: The first 10-gigabit internet city**. 19 Out 2015. Disponível em: <https://www.zdnet.com/article/chattanooga-the-first-10-gigabit-internet-city/>.

VAUGHAN-NICHOLS, Steven J. **Memcached DDoS: The biggest, baddest denial of service attacker yet**. ZDNet, 2018. Disponível em: <https://www.zdnet.com/article/memcached-ddos-the-biggest-baddest-denial-of-service-attacker-yet/>.

VIGNA, Giovanni e KEMMERER, Richard A. NetSTAT: A network-based intrusion detection approach. 1998, [S.l: s.n.], 1998. p. 25–34.

YURTSEVER, Ekim e colab. **A Survey of Autonomous Driving: Common Practices and Emerging Technologies**. IEEE Access, 2020.

ZHOU, Chenfeng Vincent e LECKIE, Christopher e KARUNASEKERA, Shanika. **A survey of coordinated attacks and collaborative intrusion detection**. Computers and Security, 2010.

ZUECH, Richard e KHOSHGOFTAAR, Taghi M e WALD, Randall. **Intrusion detection and Big Heterogeneous Data: a Survey**. Journal of Big Data, v. 2, n. 1, p. 3, 2015. Disponível em: <https://doi.org/10.1186/s40537-015-0013-4>.