Uma Revisão Sobre as Publicações de Sistemas de Detecção de Intrusão

A Review on the Publications of Intrusion Detection Systems

Felipe Cesar Costa Alves[1, 2]
Ed' Wilson Tavares Ferreira[1, 3]
Valtemir Emerencio Nascimento[1, 4]
Ruy de Oliveira[1, 5]

Data de submissão: 12/06/2015, Data de aceite: 12/09/2016

Resumo: O crescente registro de incidentes de segurança em redes de computadores tem motivado o desenvolvimento de estudos em detecção de intrusão, as principais técnicas de identificação de uma intrusão são baseadas em anomalias e assinaturas. Atualmente, a comunidade acadêmica explora preferencialmente pesquisas em redes baseadas em anomalias, entretanto, não existe um modelo comum de desenvolvimento destas propostas de modo que muitos autores descrevem, implementam e validam seus sistemas do modo heterogêneo. Neste artigo foi realizado uma pesquisa que investigou a produção científica de 112 publicações relacionadas a sistemas de detecção de intrusão. Alguns dos critérios utilizados para avaliação destes artigos foram fator de impacto, características de detecção utilizadas e a base de dados implementado. Os resultados obtidos demonstram que ocorreu um aumento da compreensão deste tema, entretanto futuros estudos serão necessários para explorar a validade dos novos métodos de avaliação em detecção de intrusão.

[1] Departamento de Informática, DAI
Instituto Federal de Educação Ciência e Tecnologia do estado de Mato Grosso, IFMT – Cuiabá, Mato Grosso, Brasil.
[2] {felcca@gmail.com}
[3] {edwilson.ferreira@ifmt.edu.br}
[4] {valtemir.nascimento@cba.ifmt.edu.br}
[5] {ruy.oliveira@cba.ifmt.edu.br}

Palavras-Chave: segurança de Redes, Detecção de intrusão, avaliação experimental

Abstract:
The growing record of security incidents on computer networks has motivated the development of intrusion detection studies, the main identification techniques of intrusion are based on anomalies and signatures. Currently, the academic community explores preferably research on anomalies based networks, however, there is no common model to develop these proposals so that many authors describe, implement and validate their heterogeneous mode systems. In this article a survey that investigated the scientific production of X publications related to intrusion detection systems was performed. Some of the criteria used to evaluate these articles were the impact factor, the detection features used and the implemented database. The results demonstrate an increase in understanding about this theme, however future studies are still needed to validate new methods of intrusion detection evaluation.

Keywords: network Security, Intrusion Detection, experimental evaluation

1 Introdução

Nos últimos anos, houve um aumento considerável de pessoas que empregam dispositivos moveis, como celulares e *tablets*, para acessar a Internet, na prática tais dispositivos tornam-se complementar ao uso do computador para o mesmo fim. Estes equipamentos geralmente são equipados com interfaces de rede sem fio, baseada na família de padrões IEEE 802.11 [92]. Com o crescimento do número de usuários de Internet e a facilidade na aquisição de produtos e serviços na rede, também aumentou a quantidade de operações financeiras realizadas através dos sistemas bancários e comércio eletrônico. Este incremento despertou o interesse de atacantes, que perceberam a grande oportunidade para aplicar golpes.

A segurança da informação possui três pilares fundamentais, representados pela tríade: confidencialidade, integridade e disponibilidade [90]. Uma ação maléfica que pode afetar um sistema é caracterizada por uma intrusão. O Sistema de Detecção de Intrusão deve ser capaz de detectar a ação, porém sem comprometer seu funcionamento ou utilizar os recursos computacionais em demasia. O IDS é, portanto uma ferramenta de segurança que, com outras medidas, a exemplos de firewall e antivírus, destina-se a reforçar a segurança da informação.

A detecção de intrusão caracteriza-se por alarmes de invasões que são ativados quando a segurança do ambiente é comprometida [1]. Os sistemas IDS, são empregados por analistas de segurança e administradores de redes como uma ferramenta auxiliar na identificação e

tratamento de ameaças e invasões, ocasionados por ações de atacantes a estruturas de redes ou sistemas [2].

Um evento de intrusão consiste em atividade anormal capaz de originar incidentes de segurança que prejudicam o funcionamento correto de uma rede [4]. As classes de invasões são referenciadas de acordo com tipos de ataques, objetivo do ataque, técnicas utilizadas, arquitetura do IDS e até meio de coletas. As taxonomias de IDS distinguem e descrevem sistemas de acordo com as suas características semelhantes entre os diversos existentes na literatura, como os sistemas de detecção baseados em redes e sistemas baseados em hosts.

O escopo deste artigo foi organizado com base na taxonomia especificada em [1] que organiza os IDS em duas principais técnicas: detecção por anomalias e detecção por assinaturas. Estes são dois dos modelos frequentemente encontrados na literatura, sendo assim, serão descritos detalhadamente na seção 2.1.

A produção científica associada à detecção de intrusão aponta um aumento expressivo em estudos relacionados a este tema, o que pode ser confirmado por uma consulta bibliográfica na plataforma Google Scholar utilizando o termo "*Intrusion Detection"*. Em uma busca efetuada por este termo e filtrada pelo período de 2010 a 2016, os resultados contabilizaram 58.800 estudos até agosto de 2016. Com base nos valores obtidos no Google Scholar e nos registros de incidentes de segurança de instituições como CERT-BR, e CAIS-RNP, conclui-se que os números elevados de ataques registrados a sistemas computacionais têm motivado a produção científica de trabalhos relacionados à segurança da informação.

Parâmetros como confiabilidade e precisão são fatores que interferem na qualidade dos métodos de detecção utilizados nos experimentos e consequentemente nas reais taxas de falso positivo e falso negativo [5]. Os alertas de falso positivo indicam atividades normais do sistema que um IDS identifica incorretamente como intrusivas. Um falso negativo é originado quando uma intrusão ocorre através de atividades maliciosas realizadas por um atacante, sendo avaliadas pelo IDS como um movimento normal na rede [91].

Neste artigo objetivou-se a realização de uma revisão teórica com a análise dos métodos utilizados para representação da eficiência dos IDS propostos pelos autores. Os resultados deste *survey* poderão auxiliar estudos futuros em detecção de intrusão. Serão abordadas hipóteses sobre as preferências no uso conjunto de dados, técnicas de detecção e documentação do experimento.

Este artigo está organizado da seguinte forma: na próxima seção são apresentados os principais conceitos sobre Sistemas de Detecção de Intrusão. Na seção 3 são apresentados trabalhos relacionados, e na seção 4 são detalhados os procedimentos metodológicos empregados, nas seção 5 as discussões sobre os resultados. Na seção 6 são apresentados modelos de detecção de intrusão e finalmente, na sequência as conclusões e sugestões para trabalhos futuros.

2 Sistemas de Detecção de Intrusão

O início dos experimentos envolvendo técnicas de detecção de intrusão surgiu em 1980 com o estudo realizado por James Anderson [6], no qual o autor descreve o comportamento de um usuário baseado em trilhas de auditoria. Estes resultados deram início ao conceito de sistemas de detecção de intrusão [7].

Os sistemas de detecção de intrusão são sistemas capazes de fornecer segurança a sistemas e redes através da detecção de eventos anômalos [8]. Os principais objetivos de um IDS são a realização de análise do comportamento da rede e a integridade da segurança de um sistema computacional para emissão de alertas. Os métodos intrusivos são diversificados e envolvem técnicas que utilizem engenharia social até ataques mais complexos como os *backdoors*.

A eficácia de uma avaliação experimental de um sistema eficiente pode ser influenciada por parâmetros que definem corretamente uma atividade anômala e atividade maliciosa [9]. Em consequência, as taxas de falso positivo e falso negativo são comprometidas e a confiabilidade do sistema torna-se questionável. Estudos antigos em ciência da computação como os realizados por [10], que já argumentavam que as atividades anômalas poderiam representar não apenas intrusões ou códigos maliciosos, mas apenas tráfego incomum.

Tendo em vista os aspectos acima discutidos, o quadro 1 demonstra abordagens distintas sobre o uso do termo anomalia. Foram utilizados estudos frequentemente citados devido ao seu impacto e influência para outras pesquisas em detecção de intrusão. Além dos estudos, o quadro apresenta a definição utilizada pelo glossário de segurança descrita pelo estudo de [12].

Quadro 1 - Relação das definições de anomalia presente na literatura.

Autor	**Definição**
Axelsson, 2000 [1]	A existência de anormalidade em tráfego representa uma suspeita. É necessária formação de um parecer sobre o que é definido como anomalia e qual porcentagem de atividades anômalas são aceitáveis para o sistema.
Allen et al., 2000 [3]	Desvio de comportamento considerado aceitável. Não representa um ataque, mas um tráfego novo ou incomum.
RFC 4949 "Internet Security Glossary Version 2". Shirey, 2007 [12]	Atividade diferente de comportamento, entidades e recursos normais definidos pelo sistema.

Autor	Definição
Wu e Banzhaf, 2010 [2]	Um comportamento anômalo e raro, diferente do comportamento de tráfego comum e aceitável.

2.1 Taxonomia de IDS

A taxonomia de um IDS é definida como uma categorização dos sistemas que descreve sua arquitetura e características funcionais a fim de solucionar problemas que cercam as técnicas de detecção de intrusão. Estudos como os realizados por [1], [13] e [14] descreveram modelos de classificação de um IDS com base em características como arquitetura, tipo de intrusão, técnicas de intrusão, método de coletas, método de análises entre outros. Um estudo mais abrangente sobre organização sistemática dos IDS foi descrito por [15] onde o autor apresenta um diagrama descrevendo principais modelos dos sistemas de detecção de intrusão.

Uma das primeiras classificações de um IDS foi proposto por [1], o autor utilizou como base as características funcionais dos sistemas que são representados pelos modelos de detecção baseado em anomalias ou assinaturas. Além destas abordagens, a seção a seguir também apresenta os modelos com base no local de coleta.

2.2 Detecção baseado em assinaturas

Os eventos identificados por um IDS baseado em assinatura (*Signature-based*) utilizam uma base de conhecimento para comparação e identificação de intrusão. Inicialmente o IDS observa uma característica anormal do fluxo de dados para geração de um identificador, tal como uma assinatura, e em seguida as ações voltam-se a pesquisa e correspondência [24]. A maior dificuldade existente para o uso deste modelo é a necessidade de precisão no reconhecimento das assinaturas, tendo em vista que as características de ataques evoluem continuamente e com isso as assinaturas podem não representar corretamente uma atividade intrusiva [2].

A principal vantagem no uso deste IDS é o baixo processamento no consumo de recursos e a possibilidade de otimização do sistema com ajustes nos algoritmos de correspondência e identificação. Desta forma, as taxas de falsos alarmes são controladas de acordo com o nível de especificação das assinaturas do IDS [27]. Apesar disso, as assinaturas devem ser especificadas frequentemente, pois o sistema detecta apenas atividades intrusivas presentes nesta base de conhecimento e uma tentativa de intrusão com características não descritas nesta base não será reconhecida. As baixas taxas de falso positivo característicos dos IDS baseados em assinatura fazem com que a maior parte do uso comercial seja constituído por neste modelo de detecção [2].

Em um modelo genérico de detecção intrusão baseado em assinatura, o sistema identifica um fluxo anômalo e em seguida estes dados são comparados à base no conjunto de

assinaturas especificadas para identificação de intrusão. Posteriormente o sistema irá emitir alerta para o tratamento da intrusão.

2.3 Detecção baseado em anomalias

Anomalias em redes são descritas como desvios acentuados nos níveis de tráfego normal e de baixo nível de detecção [28]. Um sistema baseado em anomalias (*Anomaly based*) atua com o princípio de formação de perfil de comportamento normal da rede, a detecção ocorre através de busca por alterações nos parâmetros que definem um perfil de comportamento normal [29]. As observações de tráfego de rede resultam em criação de base de perfis de comportamentos estimados. A atividade que divergem deste conjunto de dados é considerada anômala e potencialmente intrusiva [27].

A atualização constante dos modelos comportamentais ideais resulta em um custo elevado deste sistema. A geração de perfil normal envolve não apenas parâmetros técnicos, mas também conhecimentos específicos baseados em preferências de usuários como o horário de uso. Estes IDS foram definidos por [25] em modelo estatístico e de especificações. A detecção por anomalias possui como vantagem a capacidade de identificar ataques recentes e não conhecidos [30]. Neste método, há uma geração de falso-positivo frequente devido a comportamentos anormais, caracterizados como intrusivos.

Os sistemas baseados em anomalias utilizam o parâmetro base de conhecimento que contém um perfil de comportamento ideal no fluxo de atividades. Um evento intrusivo é originado e identificado pelo IDS. Após realizar correspondência de acordo com perfil estabelecido, verifica-se inconsistência nos dados e assim é gerado alerta de intrusão.

2.4 Sistemas baseados em local da colega

As técnicas de detecção de intrusão apresentadas nos tópicos anteriores descrevem o modo como as invasões são identificadas. No entanto, além destas técnicas outros modelos de detecção especificam as características como local de realização da coleta. Sendo assim, esta seção apresenta dois destes modelos de detecção.

A detecção de intrusão baseada em rede ou *Network Intrusion Detection- NID*, tem como objetivo a coleta e análise de registros de rede, do fluxo de pacotes trafegados e dos demais elementos que compõe a pilha de protocolos TCP/IP para extração de informações relacionadas a invasões. A identificação de código malicioso pode ocorrer quando NID o detecta através de interceptações e captura estes dados através da conexão com *switches* e roteadores, que transmitem cópias do tráfego para análise do IDS. Além de não interferir no desempenho dos hosts da rede, NIDs são executados em modo *background* e isso faz com que a detecção seja transparente ao atacante. Estes sistemas são capazes de monitorar e detectar tráfego anômalo de vários hosts simultaneamente [24].

Os sistemas de detecção baseado em hospedeiro ou *Host-based* são capazes de monitorar, detectar e responder aos ataques em sistemas individuais em que um IDS seja executado. Estes IDS foram descritos inicialmente por [7] como sistemas de transmissão em redes, responsáveis por analisar apenas os próprios *hosts* para combate às ameaças. Este método colabora com a coleta de informações através das pistas de auditoria e dos *logs* gerados pelo sistema. Este método é recomendado para análise de arquivos gerados por usuários de *hosts* isoladamente.

3 Trabalhos Relacionados

Apesar das primeiras propostas de IDS como a de [31] terem surgido a cerca de vinte e cinco anos, os questionamentos a respeito da validade destes sistemas têm sido abordados apenas em pesquisas recentes [5]. Desde o início das pesquisas em detecção de intrusão, os esforços são voltados apenas em como construir modelos e técnicas eficientes [2]. As discussões sobre os parâmetros utilizados em experimentos eficazes de detecção de intrusão foram realizadas também por [9]. Diante das análises efetuadas, os autores propõem, as seguintes considerações:

- atividades anômalas e maliciosas devem possuir abordagens distintas, recomenda-se que um IDS possa determinar as características de uma invasão real. Uma abordagem correta atribui taxas de verdadeiro positivo apenas aos índices de detecção de atividade maliciosa e não aos valores de uma anomalia em tráfego;

- é necessário reanalisar a definição de um comportamento anômalo. O estudo realizado por [32] identificou que apenas um terço das atividades em rede eram representadas por atividade intrusiva reforçando a abordagem apresentada por [9]; e

- o conjunto de dados aplicados em análise de IDS deve ser baseado em ambientes reais, sendo disponibilizados publicamente e atualizados anualmente. A base de dados utilizada em um IDS, assim como suas características, exercem influência significativa na avaliação de eficiência destes sistemas.

O estudo de [21] notou a escassez de um conjunto de dados públicos que sejam adequados aos experimentos em detecção de intrusão. Os autores consideram a base de dados como o desafio mais significativo para avaliação de confiabilidade de um sistema. Uma abordagem sistemática foi proposta por [33] para geração de um conjunto de dados baseando-se na criação de perfis com representação do comportamento em rede. Sendo assim, foram considerados importantes atributos como o não anonimato da base de dados, e a necessidade de cobertura integral das interações no tráfego de rede.

Na pesquisa de [5] foram revisados 276 artigos, publicados no período de 2000 até 2008 sobre detecção de intrusão em técnicas baseadas em anomalias. Os experimentos foram

avaliados considerando três atributos para comparação de técnicas eficientes de detecção de intrusão, o conjunto de dados, as características dos experimentos e os métodos utilizados para avaliação de desempenho. Os resultados deste *survey* concluíram que, a maioria dos experimentos não satisfaziam estes quesitos, colocando em questão o rigor científico e a confiabilidade na detecção de eventos anômalos em redes. Além disso, os resultados de [34] e [35], também confirmaram esta tendência experimental em ciência da computação.

O estudo realizado por [8] abordou os desafios e limitações existentes em IDS baseado em anomalias. Algumas de suas considerações referem-se à baixa eficiência em detecção decorrida de ausência de um estudo detalhado sobre a natureza dos eventos intrusivos. Além disso, modelos de detecção híbridos como o de [36] utilizam os conceitos de detecção por anomalia e detecção por assinaturas para melhoria das as taxas de precisão na identificação de invasões.

4 Metodologia

A revisão de literatura realizada nesta pesquisa utilizou diversos critérios para seleção dos estudos dos modelos de detecção. O primeiro deles foi o ano de publicação, sendo considerados ideais apenas artigos recentes, publicados entre nos anos de 2010 a 2016. O segundo critério foi o fator de impacto das revistas nas quais estes artigos estão publicados, o resumo dos parâmetros de seleção destes estudos está disponível no quadro 2.

Os dados de citações foram organizados por indexador, pelo ano de referência e tipo de publicação. O termo *Frequência de Citação – FC* especifica os valores de citações disponíveis no *Google Scholar.* Os dados foram obtidos em 24 de janeiro de 2015. Estes dados foram classificados em três níveis distintos, o FC1 (Quantidade de citações com valores de 0 a 9), FC2 (Quantidade de citações com valores 10 a 30) e FC3 (Quantidade de citações superiores a 30).

Foram selecionados e analisados artigos dos indexadores *ACM Digital Library* que é mantido pela *Association for Computing Machinery – ACM,* do *IEEE Xplore Digital Library* vinculado ao *Institute of Electrical and Electronics Engineers*, e por fim o *ScienceDirect* da editora Elsevier.

Quadro 2 - Resumo dos critérios de seleção dos experimentos.

Critérios de seleção	**Descrição**
Período das publicações	As publicações foram definidas apenas entre os anos de 2010 a 2016.
Indexadores	Selecionados os indexadores *Science Direct, IEEE Xplore,* e *ACM Digital Library.*

Critérios de seleção	**Descrição**
Características das publicações	Artigos completos e resumos expandidos.
Característica de detecção de intrusão	Não foi filtrado pela técnica, todas as características disponíveis foram consideradas.
Fator de impactos das publicações	Não foi considerado o índice de qualidade baseado em frequência de citações, apenas pelos critérios anteriores.

Os artigos que apresentaram disponibilidade em mais de um destes indexadores foram selecionados com base no primeiro local de publicação. A seleção destes estudos fundamentou-se nos resultados da pesquisa pela expressão "*Intrusion detection"* e assim selecionados apenas publicações com as características expostas anteriormente. Foram excluídos os artigos que apesar da referência relacionada ao termo "*Intrusion Detection"* não demonstram características que contemplem uma proposta de IDS. Os *surveys* são exemplos de artigos descartados nesta revisão.

Para seleção destes trabalhos, esta pesquisa analisou a qualidade dos estudos de acordo com o meio em que os periódicos foram publicados e a categoria das publicações. Foram identificadas quatro categorias distintas: os artigos das revistas científicas, de conferências, simpósios e workshops.

Entre os estudos analisados, foram identificadas publicações em nove diferentes revistas científicas, entre eles a *Personal and Ubiquitous Computing*, *IEEE Industrial Electronics Magazine*, *IEEE Transactions on Parallel and Distributed Systems*, Dependable and Secure Computing entre outras. Foi verificado que a maioria destas publicações em revistas possuem alto índice de qualidade e aceitação na comunidade acadêmica de acordo com índice QUALIS.

Nesta pesquisa foi analisado o fator de impacto dos estudos conforme a frequência de citações, além do local das publicações. Para sumarização destes dados, foi utilizado como referência à plataforma *Google Scholar*. A tabela 1 apresenta estes valores de acordo com a frequência das citações. Os registros da tabela 2 demonstram o ano das publicações com base em frequência das citações, foram analisadas as citações de acordo com os anos das publicações selecionadas nesta revisão.

Tabela 1 – Dados dos estudos organizados por frequência de citação e indexador.

Indexador	**Quantidade de estudos avaliados**	**FC1**	**FC2**	**FC3**
ACM Digital Library	36	31	3	2
IEEE Xplore	40	24	12	4
Science Direct	36	21	11	4
Total	112	76	26	10

FC: Frequência de citação.

Tabela 2 - Frequência de citação dos estudos organizados pelo ano.

Ano da publicação	***Quantidade de estudos avaliados***	***FC1***	***FC2***	***FC3***
2010	15	10	3	2
2011	15	7	4	4
2012	21	16	2	3
2013	30	23	5	2
2014	16	10	6	0
2015	9	7	2	0
2016	6	6	0	0
Total	112	79	22	11

Foi identificado a presença de estudos citados com maior frequência na plataforma *Science Direct e IEEE Xplore.* Entretanto, os periódicos dos estudos analisados não foram selecionados de modo uniforme, os trabalhos selecionados do *ACM Digital* Library a maioria foram artigos de conferências, workshops ou simpósios, enquanto em *Science Direct e IEEE Xplorer* a maior parte das publicações foram referentes as revistas científicas.

Os resultados destes registros constataram que os anos de 2012 e 2013 possuem um desenvolvimento científico acentuado, enquanto em 2010 e 2011 e 2014 apontaram índices menores na quantidade de estudos. Os artigos do ano de 2011 possuem maiores valores de trabalhos com frequência de citação FC3, seguido do 2012 com aproximadamente 14% da mesma categoria das citações. Esta revisão iniciou-se no primeiro trimestre do ano de 2014, por isso um volume maior de artigos foi analisado anterior a este período, entretanto foram inseridos artigos dos anos de 2015 e 2016 onde foi observado a ausência de artigos com grande impacto e avaliados como FC3.

Os critérios resumidos neste *survey* estão relacionados no quadro abaixo onde são sintetizados os objetivos desta pesquisa e descritos através dos parâmetros disponíveis nos artigos científicos. O indexador, fator de impacto e características de publicação referem-se aos registros que ocorrem após realização da pesquisa e publicação. Estes fatores não exercem influência direta na qualidade das propostas. Entretanto, as técnicas de detecção, avaliação experimental e tipo de conjunto de dados são fatores de grande relevância capazes de influenciar na eficácia da validade das propostas e no modo como um estudo é aceito pela comunidade acadêmica.

Quadro 3 - Resumo dos critérios de avaliação dos estudos selecionados.

Critérios de Avaliação de IDS	**Descrição**
Indexador	Locais de veiculação dos estudos analisados entre eles IEEE Xplore, Elsevier e ACM Library.
Fator de impacto e data da publicação	Definido fator de impacto de acordo com a frequência de citação dos experimentos. Entre eles FC1, FC2 e FC3.
Modelo de detecção de intrusão	Diversos modelos de detecção de intrusão são descritos na literatura, entre eles sistemas baseados em anomalia, assinaturas [1]. Este segmento objetiva a identificação de características que divergem destes modelos.
Características da publicação	Refere-se ao tipo do periódico vinculado ao estudo, entre eles revistas científicas, conferências, simpósios e workshops.

Critérios de Avaliação de IDS	Descrição
Conjunto de dados	Foram investigadas as especificações e características apontadas pelos autores em relação ao conjunto de dados utilizados. Entre elas se o conjunto de dados possui disponibilidade pública ou são dados sintéticos criados pelos autores apenas para uso em seu estudo.
Documentação e configuração do experimento	A relevância científica de um modelo de detecção ideal é também baseada no potencial de repetição de um experimento, ou seja, a documentação deve ser suficiente para possibilitar a reprodução do sistema proposto. Sendo assim, é de suma importância que os autores descrevam com detalhes o modo como a pesquisa foi conduzida.
Característica de detecção do sistema	Neste segmento foram examinadas as características metodológicas utilizadas nos experimentos. Entre eles, os detalhes do algoritmo proposto, o ambiente de implementação, as justificativas e as técnicas selecionadas como por exemplo redes neurais artificiais.

5 Resultados e discussões

Após análise dos artigos, os parâmetros apresentados no quadro 5 serão descritos com detalhes nesta seção. Além dos registros quantitativos identificados nos modelos propostos, serão apresentadas informações referentes as características da avaliação metodológica e o estado da arte destes padrões. Além disso, para confirmação dos parâmetros examinados, estes artigos passaram por duas rodadas de leitura para redução do índice de erros de avaliação.

5.1 Conjunto de dados

A avaliação das diferentes técnicas de detecção de intrusão necessitam de intervenções humanas para auxilio em um sistema eficiente [37], a criação de assinaturas e algoritmos de aprendizado são algumas das ações realizadas com frequência. Com base nos valores dos dados do comportamento normal e anômalo da rede, estes algoritmos de aprendizado são capazes de definir modelos ideais de dados para avaliar um sistema de detecção de intrusão. Um conjunto de dados (*datasets*) são valores de treinamentos de tráfego de rede que auxiliam em uma análise de eficácia de IDS baseando-se na comparação e reprodução dos resultados dos logs de ataques. A verificação de um conjunto de dados é utilizada como referência para validação do modelo de detecção proposto através de treinamento da técnica de detecção [35].

A realização destes procedimentos é considerada essencial para avanços das pesquisas em segurança da informação. Os estudos em detecção de intrusão utilizam base de dados que são disponibilizados publicamente ou bases proprietárias criadas apenas para

pesquisas específicas. O *survey* realizado por [5] demonstrou que 70% dos estudos em detecção de anomalia utilizaram conjunto de dados públicos, sendo o KDD99 e DARPA (*Defense Advanced Recearch Projects Agency*) utilizados com maior frequência com 24% e 28% respectivamente.

O conjunto de dados KDD99[6] foi financiado pelo DARPA em parceria com o Lincon Laboratory do MIT (*Massachusetts Institute of Technology*) nos anos de 1998 e 1999. Esta base armazenou pacotes originados de conexões TCP com 41 características de identificação que apontavam tráfego normal ou anômalo e inclui 38 tipos de ataques que foram lançados em hosts Unix. O projeto inicial, realizado em 1998, utilizou dados equivalentes a sete semanas de treinamento e duas semanas de testes. No ano de 1999 uma nova base foi gerada, sendo desferidos 58 tipos de ataques gerados em três semanas de treinamentos e duas semanas de testes. Na base KDD99, os perfis de usuários foram projetados e registrados, sendo o tráfego de rede coletado com base no comportamento sintético [38].

Nos anos seguintes, estudos como os de [38] e [39] apresentaram críticas à metodologia utilizada na geração destes dados, como por exemplo, o critério de avaliação e a ausência de atributos para diferenciação entre os dados reais e sintéticos.

O conjunto KDD99 possui outras deficiências e foram descritas por [40]. Alguns destes problemas são a grande quantidade de dados redundantes, em torno de 75%, dados de ataques falsos e antiquados para avaliação. Estudos recentes como o realizado por [33] mantiveram esta percepção em relação a deficiência e ausência de base de dados apropriada para prática de avaliação em sistemas de detecção de intrusão.

Conjuntos de dados propostos por projetos como DARPA e KDD estão disponíveis publicamente e além destas, outras propostas foram produzidas com os mesmos objetivos. O conjunto de dados IES (Internet Exploration Shootout) por exemplo, obteve os dados em capturas de tráfego TCP e UDP de apenas 16 minutos. Este curto período foi criticado por [41] que o considerou insuficiente para a caracterização de um comportamento de normal ou anômalo. Além destas, estão disponíveis publicamente outras opções de conjunto de dados para avaliação experimental de IDS.

O estudo de [2] sumarizou algumas destas opções e entre eles o 10% KDD99. O volume de tráfego deste conjunto de dados utiliza dados originais do KDD99 que é composto por cerca de cinco milhões de registros, entretanto, esse número elevado de dados fez com que a análise fosse dificultada. Por conta disso, inicialmente apenas 500.000 foram analisados e posteriormente 13.094 e 6.900. Com isso 10% foram utilizados para os testes [42].

A avaliação realizada nesta pesquisa identificou uma tendência na escolha de base de dados, a opção pela utilização preferencial de conjunto de dados DARPA e KDD99. Do total de estudos analisados, aproximadamente 47% deles utilizaram conjunto de dados disponíveis publicamente sendo que o KDD com 31 estudos e o DARPA com 13 foram as duas bases de dados mais utilizadas. Foram também identificadas uso de outras opções de conjunto de dados

[6] http:// kdd.ics.uci.edu/ databases/ kddcup99/ kddcup99.html

como por exemplo no trabalho de [43] que optou pelo uso do LANDER (Los Angeles Network Data Exchange and Repository)[7].

Muitos autores utilizam base de dados próprios para seus estudos ou não os especificam em sua pesquisa. Estes critérios representam aproximadamente 52% do total de artigos avaliados. Foi observada também uma escassez de informações a respeito do uso destas bases de dados sintéticas, em diversos experimentos o foco da pesquisa manteve-se em descrições como arquitetura do sistema, modelos de detecção e algoritmos utilizados.

Figura 1 - Registros dos tipos de bases de dados identificados.

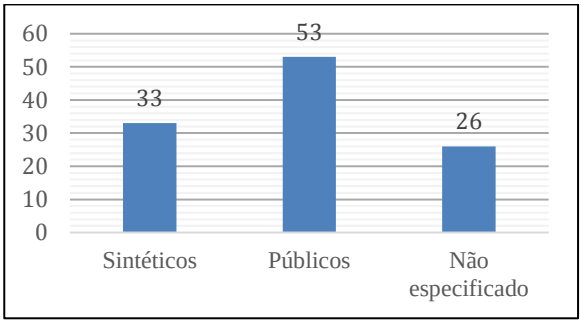
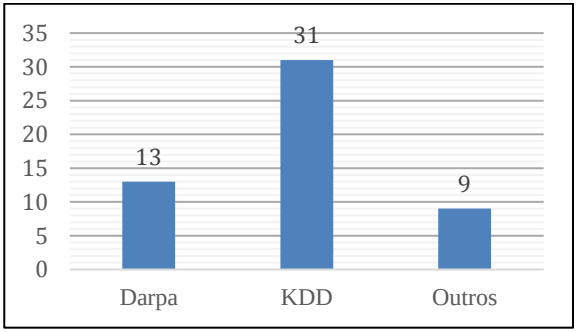


Figura 2 - Registros das bases de dados identificados.



Quando comparado a relação entre a frequência de citação e o tipo de dados utilizados, foi verificado que os estudos onde as bases de dados não são caracterizadas possuem

[7] http://www.isi.edu/ant/lander/index.html

baixa relevância científica pois a maioria estão categorizadas como FC1. Experimentos com bases públicas e sintéticas demonstraram resultados semelhantes, ambos totalizaram 5 artigos na categoria FC3. Os resultados quantitativos da análise sobre o conjunto de dados foram especificados e organizados na tabela 6.

Tabela 3 - Registros das bases de dados com base em frequência de citações.

Tipos de Dados	***FC1***	***FC2***	***FC3***
Sintéticos	11	9	5
Públicos	31	5	5
Não especificados	19	5	0
Total	61	19	10

FC: Frequência de citação.

5.2 Avaliação experimental

A capacidade de reprodução dos experimentos com base nas características metodológicas é essencial para a confiabilidade e aceitação dos modelos de detecção avaliados. A baixa qualidade metodológica e escassez de documentação são fatores que influenciam de modo negativo a viabilidade de um IDS [5]. Nesta revisão, a documentação dos artigos pesquisados foi organizada e definida em três níveis distintos: escassos, regulares ou satisfatórios.

Os artigos de documentações escassas apresentaram poucos detalhes a respeito do modo como a pesquisa foi conduzida, ou seja, informações sobre o método de configuração, ambiente de testes, arquitetura dos sistemas, tipo da rede ou algoritmo utilizado foram reduzidas. Por outro lado, os autores dos artigos definidos como documentação regular ou satisfatória atentaram-se na demonstração dos parâmetros citados anteriormente.

A maior parte destes estudos demonstrou documentação satisfatória quando se tratou das revistas científicas, ainda que estes documentos apresentem número de páginas superiores aos artigos oriundos de eventos como simpósios ou workshop. Entretanto, o estudo de [44] que esteve entre os analisados foi categorizado como frequência de citação FC3, apesar deste trabalho estar categorizado como artigo de conferência. Além deste, outros casos de desconformidade ocorreram quando relacionado o tipo de publicação ao nível de detalhamento da proposta.

Outra importante característica para avaliação de um estudo de IDS é o ambiente de avaliação. Os autores dos artigos selecionados optam pela utilização de um ambiente simulado ou pela implementação em ambientes reais. A maioria destes estudos utilizou simulação para avaliação dos experimentos. No total, apenas 16 autores declararam implementação e avaliação das propostas em ambiente real. Poucos trabalhos expõem minuciosamente as características destes ambientes de implantação como o hardware dos equipamentos, a abrangência desta rede e o modo como estes aspectos podem influenciar na precisão e avaliação experimental das propostas.

Estas análises confirmam que a comunidade acadêmica ainda não possui um modelo ideal uniforme relacionado a documentação de avaliação em pesquisas em mecanismos de segurança. Para comprovação da eficiência de um modelo de detecção, os autores recorrem a organização de técnicas que atendam geralmente apenas às necessidades de seus estudos. Com isso as técnicas são diversificadas e os resultados de precisão tornam-se questionáveis.

6 Modelos de Detecção de intrusão

O estudo realizado por [5] norteou esta pesquisa. Além das características de avaliação já apontadas aqui, para investigação de detecção de anomalias, os autores abordaram diversas

métricas para auxilio a análise de eficiência e precisão descritas em suas pesquisas. Embora esta pesquisa não tenha conduzido uma investigação aprofundada de todas estas características, algumas delas consideradas essenciais também foram observadas, como a especificação das definições de anomalia.

A investigação aos 276 artigos buscou as considerações dos autores sobre o que são caracterizados como anomalias, já que no tráfego da rede esta informação influencia diretamente na caracterização de um evento intrusivo. Neste artigo, apenas 9% dos estudos definiram os parâmetros para caracterização de anomalia em tráfego da rede.

Dos estudos selecionados, todos apresentavam propostas de modelos de IDS. Foram selecionados por ordem de disponibilidade nos indexadores filtrando-os apenas pelos anos de publicação e excluindo capítulos de livros, patentes e *surveys*, os artigos que não se enquadravam em técnicas de intrusões foram descartados. A tabela 4 apresenta análise quantitativa das técnicas de detecção descobertas nesta revisão.

Tabela 4 - Registros das principais técnicas de detecção identificadas.

Técnicas de detecção identificadas	***Quantidade de estudos***
Detecção baseada em Anomalia	26
Detecção baseada em Rede	30
Detecção baseada em Assinatura	15
Detecção baseada em Host	8
Técnicas não especificadas	5
Outras técnicas de detecção	28
Total	112

Inicialmente foi investigado o método de detecção empregado. A tabela abaixo foi sintetizada com base na taxonomia de IDS. Sendo assim, foram mantidos apenas como métodos principais as detecções baseadas em anomalias, rede, host e assinatura. As propostas que divergiam destas, foram categorizadas como "*outras técnicas"*. Dos artigos analisados, 30 deles empregaram método baseado em rede, 18 foram as detecções por anomalia, detecções por assinatura com 15 e por fim detecção baseado em hosts com 8 estudos.

A soma dos artigos que não descreveram mecanismos de detecção ou não se adequam aos citados anteriormente resultaram em 33. Na tabela 5, estes valores foram organizados de acordo com a frequência de citações. Apenas 9% dos trabalhos possuíram citações correspondente a FC3. Nas técnicas baseadas rede, estes valores de citação representam 13% do total de artigos avaliados sendo o maior entre as técnicas encontradas.

Tabela 5 - Registros das técnicas de detecção identificadas com base em frequência de citações.

Técnicas de detecção identificadas	***FC1***	***FC2***	***FC3***
Detecção baseada em Anomalia	19	6	1
Detecção baseada em Rede	22	4	4
Detecção baseada em Assinatura	11	2	2
Detecção baseada em Host	2	4	2
Técnicas não especificadas	4	1	0
Outras técnicas de detecção	18	9	1
Total	76	26	10

FC: Frequência de citação.

Os resultados destas estatísticas evidenciam propostas de detecção apoiadas por técnicas variadas. Verificou-se que os mecanismos baseados em host e assinaturas não possuem grande representação na comunidade acadêmica com base na métrica utilizada nesta revisão. Os tópicos a seguir apresentam alguns dos padrões de taxonomia observados nesta investigação com características funcionais ou nomenclaturas referenciados com pouca frequência na literatura.

- Detecção baseada em Energia [45];
- Detecção baseada em *Storage* [46];
- Detecção baseada em Hypervisior [47];
- Detecção baseada em Jogo [48]; e
- Detecção baseada em *Agent* [49].

Ao realizar as pesquisas em detecção de intrusão, é importante que os autores priorizem e definam parâmetros como o cenário para implementação de suas ferramentas e os recursos que auxiliem na precisão dos sistemas, além das métricas já demonstradas nos tópicos anteriores. Ainda que os recursos utilizados na elaboração de um IDS atendam os parâmetros descritos neste documento, é necessário que os autores demonstrem todas as etapas da produção e avaliação do sistema. Foram examinados o local de implementação do IDS e os recursos de detecção empregados. O quadro 4 sintetiza estas características e apresenta os autores responsáveis pela produção destes trabalhos. O quadro 7 organiza as principais características de detecção encontradas nesta avaliação.

Os registros destes quadros demonstram valores expressivos em propostas implementadas sobre as redes de sensores e redes ad-hoc. Além disso, também são evidentes as propostas sobre tecnologias que apresentam número crescente de aplicações em ciência da computação, como a Computação em Nuvem [50] e as Redes Mesh [51]. Por fim, entre os mecanismos de auxílio a detecção verificou-se o a utilização expressiva de algoritmos baseados em redes neurais e mineração de dados, apenas estes dois padrões são responsáveis por 14% do total de estudos avaliados.

Quadro 4 – Relação dos ambientes de implementação dos experimentos.

Ambiente de implementação	**Autores**
Redes Peer-to-Peer	(BANIK et al., 2013) [52] e (ZHANG, et al., 2014) [53].
Rede ad-hoc	(SAMARAS, et al. 2013) [54], (SHAKSHUKI et al., 2013) [55], (ZHANG e YEO 2011) [56] e (KUMAR e CHILAMKURTI, 2014) [57].
Redes de sensores	(ABUHELALEH e ELLEITHY, 2011) [58], (RIECKER et al., 2013) [45], (COPPOLINO et al., 2013) [59], (WANG et al., 2013) [60], (SUN et al., 2013) [61], (LO e ANSARI, 2013) [62], (TAN et al., 2011) [63], (HAMEDHEIDARI e RAFEH, 2013) [64], (RAZA e WALLGREN, 2013) [65] (WANG et al., 2011) [66], (HASSANZADEH STOLERU, 2013) [67] e (HASSANZADEH et al. 2016) [113].
Computação em nuvem	(MODI, et al., 2012) [68], (NIKOLAI e WANG, 2014) [47], (ALHARKAN e MARTIN, 2012) [69] e (THANG et al. 2016) [105].

Ambiente de implementação	**Autores**
Sistema baseado em Web	(JAMDAGNI et al., 2010) [70], (NAJJAR et al., 2010) [71] e (ARIU e GIACINTO, 2011) [72].
Máquina virtual	(AZMANDIAN et al., 2011) [16], (CHUNG et al., 2013) [73] e (SERESHT e AZMI, 2014) [49].
Smartphones	(ROSHANDEL et al., 2013) [74].
Rede Mesh	(DO CARMO e HOLLICK, 2013) [75]
Cluster	(NGUYEN, et al., 2011) [76], (LINGXI e GUANG, 2013) [77] e (SPATHOULAS e KATSIKAS, 2013) [78].

Quadro 5 – Técnicas utilizadas em algoritmos de detecção.

Caracteristícas do IDS	**Autores**
Redes neurais	(YING et al., 2010) [79], (HUSAGIC et al., 2013) [42], (SALEK e MADANI, 2013) [80], (GOVINDARAJAN e CHANDRASEKARAN, 2011) [23] e (TJHAI et al., 2010) [81]. (CAN et al. 2015) [111]
Mineração de dados	(NAMBIAR, et al., 2010) [82], (NADIAMMAI e HEMALATHA, 2012) [83], (WENGUANG et al., 2011) [84], (MABU, et al., 2011) [85] e (KOC et al., 2012) [86]. (FAISAL et al. 2015) [101]
Aprendizado de máquina	(SYMONS, et al., 2012) [87], (SAHA, et al., 2012) [88] e (SANGKATSANEE et al., 2011) [89].

7 Conclusões e Trabalhos Futuros

Nesta pesquisa, foi realizada uma revisão através da análise de estudos em detecção de intrusão. Afim de examinar a qualidade destas propostas foram observados os parâmetros capazes de afetar a precisão destas técnicas e em consequência alterar eficácia de um IDS. Foram selecionados 112 artigos acadêmicos dos indexadores ACM Digital Library, IEEE Xplore Digital Library e ScienceDirect. As principais métricas de avaliação definidas nesta pesquisa foram os tipos de publicação, descrição do conjunto de dados, características do tipo de detecção e descrição da documentação e configuração do experimento.

A avaliação dos artigos selecionados permitiu verificar que em parte dos artigos foram empregados base de dados públicos como o KDD99 e outros os autores criaram seus próprios conjuntos de dados. Somente esta base de dados representou 58% dos artigos avaliados. Porém, foi significativo o percentual de artigos que foram desenvolvidos empregando bases de dados não disponível publicamente, tal número representou 52%.

Em relação a prática e configuração experimental, foi percebida que a maioria dos autores descrevem poucas informações referentes as definições de anomalias. Em geral, muitos destes estudos apresentam taxas de detecção superiores a 90%, no entanto não especificam a proporção de tráfego normal e malicioso. Este fator representa uma baixa confiabilidade dos sistemas.

Para trabalhos futuros, sugere-se expandir os estudos para avaliação de segurança em redes de sensores, computação em nuvem e redes ad-hoc e comparação e investigação dos atributos e eficiência de outras bases de dados disponíveis publicamente.

Contribuição dos autores:

- Felipe Cesar Costa Alves: Pesquisa nas bases de dados e tabulação.

- Ed' Wilson Tavares Ferreira: Orientação sobre a metodologia empregada.

- Ruy de Oliveira: Revisão dos resultados obtidos e revisão do texto.

- Valtemir Emerencio do Nascimento: Revisão dos resultados obtidos e revisão do texto.

Referências

[1] Axelsson, S. *Intrusion detection systems: A survey and taxonomy* (Vol. 99). Chalmers University of Technology, Goteborg, Sweden: Technical report. 2000

[2] Wu, S. X., Banzhaf, W. *The use of computational intelligence in intrusion detection systems: a review.* Applied Soft Computing, New York, v. 10, n. 1, 35 p., Jan. 2010.

[3] Allen, J. et al. *State of the practice of intrusion detection technologies*. No. CMU/SEI-99-TR-028. Carbegie-Mellon Univ Pittsburgh PA SOFTWARE ENGINEERING INST, 2000.

[4] Alrajeh, Nabil Ali; Khan, Shafiullah; Shams, Bilal. *Intrusion detection systems in wireless sensor networks: a review.* International Journal of Distributed Sensor Networks, v. 2013, 2013.

[5] Tavallaee, Mahbod; Stakhanova, Natalia; Ghorbani, Ali Akbar. *Toward credible evaluation of anomaly-based intrusion-detection methods.* Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, v. 40, n. 5, p. 516-524, 2010.

[6] Anderson, James P. *Computer security threat monitoring and surveillance* Technical report, James P. Anderson Company, Fort Washington, Pennsylvania, 1980.

[7] Innella, Paul. *The evolution of intrusion detection systems.* Security Focus–2001.

[8] Garcia Teodoro, Pedro et al. *Anomaly-based network intrusion detection: Techniques, systems and challenges.* Computers & security, v. 28, n. 1, p. 18-28, 2009.

[9] C. Gates and C. Taylor, *Challenging the anomaly detection paradigm: A provocative discussion* in Proc. 2006 Workshop N. Security Paradigms (NSPW), pp. 21–29.

[10] S. Bellovin. *Packets found on an internet*. Technical report, AT&T Bell Laboratories, 1992.

[11] R. Maxion and F. Feather. *A case study of Ethernet anomalies in a distributed computing environment.* IEEE Transactions on Reliability, 39 (4):433 – 443, 1990.

[12] Shirey, Robert W. *RFC 4949, Internet security glossary, Version 2.* 2007.

[13] Wood, Mark; ERLINGER, Michael A. *Intrusion detection message exchange requirement*s. 2007.

[14] Stakhanova, Natalia; BASU, Samik; WONG, Johnny. *A taxonomy of intrusion response systems.*" International Journal of Information and Computer Security, v. 1, n. 1, p. 169-184, 2007.

[15] Lazarevic, Aleksandar; Kumar, Vipin; Srivastava, Jaideep. *Intrusion detection: A survey*. Managing Cyber Threats. Springer US, 2005. p. 19-78.

[16] Azmandian, Fatemeh et al. *Virtual machine monitor-based lightweight intrusion detection*. ACM SIGOPS Operating Systems Review, v. 45, n. 2, p. 38-53, 2011.

[17] Modi, Chirag et al. *A survey of intrusion detection techniques in cloud.* Journal of Network and Computer Applications, v. 36, n. 1, p. 42-57, 2013.

[18] Houmansadr, Amir; Zonouz, Saman A.; Berthier, Robin. *A cloud-based intrusion detection and response system for mobile phones.* Dependable Systems and Networks Workshops (DSN-W), 2011 IEEE/IFIP 41st International Conference on. IEEE, 2011. p. 31-32.

[19] Dhage, Sudhir N.; Meshram, B. B. *Intrusion detection system in cloud computing environment*. International Journal of Cloud Computing, v. 1, n. 2, p. 261-282, 2012.

[20] Tsai, Chih-Fong, et al. *Intrusion detection by machine learning: A review* Expert Systems with Applications 36.10, 2009: 11994-12000.

[21] Sommer, Robin; Paxson, Vern. *Outside the closed world: On using machine learning for network intrusion detection.* Security and Privacy (SP), 2010 IEEE Symposium on. IEEE, 2010.

[22] Corchado, Emilio; Herrero, Álvaro. *Neural visualization of network traffic data for intrusion detection.* Applied Soft Computing, v. 11, n. 2, p. 2042-2056, 2011.

[23] Govindarajan, M.; Chandrasekaran, R. M. *Intrusion detection using neural based hybrid classification methods.* Computer networks, v. 55, n. 8, p. 1662-1671, 2011.

[24] Kim, Jungwon et al. *Immune system approaches to intrusion detection – A review.* Natural computing, v. 6, n. 4, p. 413-466, 2007.

[25] Gill, R. S. *Intrusion detection techniques in wireless local area networks*. 2009. 264 f. Information Technology (Doctor Of Philosophy) - Faculty de Information Technology, Queensland University Of Technology, Queensland, 2009.

[26] Ilgun, Koral; Kemmerer, Richard A.; Porras, Phillip A. *State transition analysis: A rule-based intrusion detection approach*. Software Engineering, IEEE Transactions on, v. 21, n. 3, p. 181-199, 1995.

[27] Mandujano, Salvador. A *Multiagent Approach to Outbound Intrusion Detection.* Monterrey Campus, 2004.

[28] Zarpelão, Bruno Bogaz, et al. *Detecção de anomalias em redes de computadores*. Simpósio brasileiro de telecomunicações, XXVII. Anais (2009).

[29] Shon, T.; Moon, J. *A hybrid machine learning approach to network anomaly Detection*, Information Sciences, v. 177, no. 18, Sep. 2007, p. 3799-3821.

[30] Boro, Debojit; Nongpoh, Bernard; Bhattacharyya, Dhruba K. *Anomaly based intrusion detection using meta ensemble classifier*. In: Proceedings of the Fifth International Conference on Security of Information and Networks. ACM, 2012. p. 143-147.

[31] Denning, Dorothy E. *An intrusion-detection model*. Software Engineering, IEEE Transactions on, n. 2, p. 222-232, 1987.

[32] K. Xu, Z.-L. Zhang, and S. Bhattacharyya. *Profiling internet backbone traffic: Behavior models and applications.* In Proceedings of the 2005 ACM SIGCOMM Conference, pages 169 – 180, Philadelphia, PA, USA, August 2005.

[33] Shiravi, A., Shiravi, H., Tavallae, M. e Ghorbani, A., A., *Toward Developing a Systematic Approach to Generate Benchmark Datasets for Intrusion Detection*, Computers & Security, vol. 31, p. 357 – 374, 2012.

[34] Wainer, Jacques et al. *Empirical evaluation in Computer Science research published by ACM*. Information and Software Technology, v. 51, n. 6, p. 1081-1085, 2009.

[35] Jyothsna, V.; Prasad, V. V.; Prasad, K. Munivara. *A Review of Anomaly based Intrusion Detection Systems*. International Journal of Computer Applications, v. 28, 2011.

[36] Depren, Ozgur et al. *An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks*. Expert systems with Applications, v. 29, n. 4, p. 713-722, 2005.

[37] Kayacik, H. Günes; Zincir-Heywood, A. Nur; Heywood, Malcolm I. *Selecting features for intrusion detection: A feature relevance analysis on KDD 99 intrusion detection datasets.* Proceedings of the third annual conference on privacy, security and trust. 2005.

[38] Mahoney, Matthew V.; Chan, Philip K. *An analysis of the 1999 DARPA/Lincoln Laboratory evaluation data for network anomaly detection.* Recent Advances in Intrusion Detection. Springer Berlin Heidelberg, 2003. p. 220-237.

[39] Mchugh, John. *Testing intrusion detection systems: a critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory.* ACM transactions on Information and system Security, v. 3, n. 4, p. 262-294, 2000.

[40] Tavallaee, Mahbod et al. *A detailed analysis of the KDD CUP 99 data set*. In: Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defence Applications 2009.

[41] Balthrop, Justin; FORREST, Stephanie; Glickman, Matthew R. "Revisiting lisys: Parameters and normal behavior." Computational Intelligence, Proceedings of the World on Congress on. IEEE, 2002.

[42] Husagic-Selman, Alma; Koker, Rasit; Selman, Suvad. *Intrusion detection using neural network committee machine.* In: Information, Communication and Automation Technologies (ICAT), 2013 XXIV International Symposium on. IEEE, 2013. p. 1-6.

[43] Tartakovsky, Alexander G.; Polunchenko, Aleksey S.; Sokolov, Grigory. *Efficient computer network anomaly detection by changepoint detection methods*. Selected Topics in Signal Processing, IEEE Journal of, v. 7, n. 1, p. 4-11, 2013.

[44] Vasiliadis, Giorgos; Polycronakis, Michalis; Ionnidis, Sotiris. *MIDeA: a multi-parallel intrusion detection architecture.* In: Proceedings of the 18th ACM conference on Computer and communications security. ACM, 2011. p. 297-308.

[45] Riecker, Michael; Biedermann, Sebastian; Hollick, Matthias. *Lightweight energy consumption based intrusion detection system for wireless sensor networks*. In: Proceedings of the 28th Annual ACM Symposium on Applied Computing. ACM, 2013. p. 1784-1791.

[46] Pennington, Adam G. et al. *Storage-based intrusion detection*. ACM Transactions on Information and System Security (TISSEC), v. 13, n. 4, p. 30, 2010.

[47] Nikolai, Jason; Wang, Yong. *Hypervisor-based cloud intrusion detection system. In: Computing*, Networking and Communications (ICNC), 2014 International Conference on. IEEE, 2014. p. 989-993.

[48] Kantzavelou, Ioanna; Katsikas, Sokratis. *A game-based intrusion detection mechanism to confront internal attackers*. Computers & Security, v. 29, n. 8, p. 859-874, 2010.

[49] Seresht, Neda Afzali; Azmi, Reza. *MAIS-IDS: A distributed intrusion detection system using multi-agent AIS approach*. Engineering Applications of Artificial Intelligence, v. 35, p. 286-298, 2014.

[50] Armbrust, Michael et al. *A view of cloud computing*. Communications of the ACM, v. 53, n. 4, p. 50-58, 2010.

[51] Akylidiz, Ian F.; Wang, Xudong; Wang, Weilin. *Wireless mesh networks: a survey*. Computer networks, v. 47, n. 4, p. 445-487, 2005.

[52] Banik, Shankar M.; Bernsen, Derek S.; Javed, Muhammad. *IMAIDS: intelligent mobile agent-based intrusion detection system*. In: Proceedings of the 51st ACM Southeast Conference. ACM, 2013. p. 25.

[53] Zhang, Junjie et al. *Building a scalable system for stealthy p2p-botnet detection*. Information Forensics and Security, IEEE Transactions on, v. 9, n. 1, p. 27-38, 2014.

[54] Samaras, Nicholas S. et al. *On intrusion detection in opportunistic networks*. In: Proceedings of the 17th Panhellenic Conference on Informatics. ACM, 2013. p. 67-74.

[55] Shakshuki, Elhadi M.; Kang, Nan; Sheltami, Tarek R. Eaack - *A Secure Intrusion-Detection System for MANETs*. Industrial Electronics, IEEE Transactions on, v. 60, n. 3, p. 1089-1098, 2013.

[56] Zhang, Da; Yeo, Chai Kiat. *Distributed Court System for intrusion detection in mobile ad hoc networks*. computers & security, v. 30, n. 8, p. 555-570, 2011.

[57] Kumar, Neeraj; Chilamkurti, Naveen. *Collaborative trust aware intelligent intrusion detection in VANETs*. Computers & Electrical Engineering, v. 40, n. 6, p. 1981-1996, 2014.

[58] Abuhelaleh, Mohammed A.; Elleithy, Khaled M. *Security in wireless sensor networks: key intrusion detection module in SOOAWSN*. In: Proceedings of the 14th Communications and Networking Symposium. Society for Computer Simulation International, 2011. p. 56-61.

[59] Coppolino, Luigi et al. *Applying data mining techniques to intrusion detection in wireless sensor networks*. In: P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2013 Eighth International Conference on. IEEE, 2013. p. 247-254.

[60] Wang, Yun; Fu, Weihuang; Agrawal, Dharma P. *Gaussian versus uniform distribution for intrusion detection in wireless sensor networks*. Parallel and Distributed Systems, IEEE Transactions on, v. 24, n. 2, p. 342-355, 2013.

[61] Sun, Bo et al. *Anomaly detection based secure in-network aggregation for wireless sensor networks. Systems Journal*, IEEE, v. 7, n. 1, p. 13-25, 2013.

[62] Lo, Chun-Hao; ANSARI, NIRWAN. *CONSUMER: A novel hybrid intrusion detection system for distribution networks in smart grid*. Emerging Topics in Computing, IEEE Transactions on, v. 1, n. 1, p. 33-44, 2013.

[63] Jamdagni, Aruna et al. *Intrusion detection using Gsad model for HTTP traffic on web services*. In: Proceedings of the 6th International Wireless Communications and Mobile Computing Conference. ACM, 2010. p. 1193-1197.

[64] Hamedheidari, Sina; Rafeh, Reza. *A novel agent-based approach to detect sinkhole attacks in wireless sensor networks*. Computers & Security, v. 37, p. 1-14, 2013.

[65] Raza, Shahid; Wallgren, Linus; Voigt, Thiemo. *Svelte: Real-time intrusion detection in the Internet of Things.* Ad hoc networks, v. 11, n. 8, p. 2661-2674, 2013.

[66] Wang, Shun-Sheng et al. *An integrated intrusion detection system for cluster-based wireless sensor networks*. Expert Systems with Applications, v. 38, n. 12, p. 15234-15243, 2011.

[67] Hassanzadeh, Amin; Storelu, Radu. *On the optimality of cooperative intrusion detection for resource constrained wireless networks*. Computers & Security, v. 34, p. 16-35, 2013.

[68] Modi, Chirag et al. *A novel framework for intrusion detection in cloud*. In: Proceedings of the Fifth International Conference on Security of Information and Networks. ACM, 2012. p. 67-74.

[69] Alharkan, Turki; Martin, Patrick. *IDSaaS: Intrusion detection system as a service in public clouds*. In: Proceedings of the 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (ccgrid 2012). IEEE Computer Society, 2012. p. 686-687.

[70] Jamdagni, Aruna et al. *Repids: A multi tier real-time payload-based intrusion detection system*. Computer Networks, v. 57, n. 3, p. 811-824, 2013.

[71] Najjar, Meisam SA; Abdollahi Azgomi, Mohammad. *A distributed multi-approach intrusion detection system for web services.* In: Proceedings of the 3rd

international conference on Security of information and networks. ACM, 2010. p. 238-244.

[72] ARIU, Davide; TRONCI, Roberto; GIACINTO, Giorgio. *HMMPayl: An intrusion detection system based on Hidden Markov Models*. Computers & security, v. 30, n. 4, p. 221-241, 2011.

[73] Chung, Chun-Jen et al. *NICE: Network intrusion detection and countermeasure selection in virtual network systems*. IEEE transactions on dependable and secure computing, n. 4, p. 198-211, 2013.

[74] Roshandel, Roshanak; Arabshahi, Payman; Poovendran, Radha. *LIDAR: a layered intrusion detection and remediation framework for smartphones*. In: Proceedings of the 4th international ACM Sigsoft symposium on Architecting critical systems. ACM, 2013. p. 27-32.

[75] Do Carmo, Rodrigo; Hollick, Matthias. *DogoIDS: a mobile and active intrusion detection system for IEEE 802.11 s wireless mesh networks*. In: Proceedings of the 2nd ACM workshop on Hot topics on wireless network security and privacy. ACM, 2013. p. 13-18.

[76] Nguyen, Huu Hoa; Harbi, Nouria; Darmount, Jérôme. *An efficient local region and clustering-based ensemble system for intrusion detection*. In: Proceedings of the 15th Symposium on International Database Engineering & Applications. ACM, 2011. p. 185-191.

[77] Lingxi, Meng; Guang, Sun. *An Improved Ant Colony Clustering Method for Network Intrusion Detection. In: Networking,* Architecture and Storage (NAS), 2013 IEEE Eighth International Conference on. IEEE, 2013. p. 312-316.

[78] Spathoulas, Georgios P.; Katsikas, Sokratis K. *Enhancing IDS performance through comprehensive alert post-processing*. Computers & Security, v. 37, p. 176-196, 2013.

[79] Ying, Lin; Yan, Zhang; Yang-JIA, Ou. *The design and implementation of host-based intrusion detection system*. In: Intelligent Information Technology and Security Informatics (IITSI), 2010 Third International Symposium on. IEEE, 2010. p. 595-598.

[80] Salek, Zahra; Madani, Fariborz Mousavi; AZMI, Reza. *Intrusion detection using neuarl networks trained by differential evaluation algorithm*. In: Information Security and Cryptology (ISCISC), 2013 10th International ISC Conference on. IEEE, 2013. p. 1-6.

[81] Tjhai, Gina C. et al. *A preliminary two-stage alarm correlation and filtering system using SOM neural network and K-means algorithm*. Computers & Security, v. 29, n. 6, p. 712-723, 2010.

[82] Nambiar, Athira M.; Vijayan, Asha; Nandakumar, Aishwarya. *Wireless intrusion detection based on different clustering approaches*. In: Proceedings of the 1st Amrita ACM-W Celebration on Women in Computing in India. ACM, 2010. p. 42.

[83] Nadiammai, G. V.; Hemalatha, M. *An evaluation of clustering technique over intrusion detection system*. In: Proceedings of the International Conference on Advances in Computing, Communications and Informatics. ACM, 2012. p. 1054-1060.

[84] Wenguang, Chai; Chunhui, Tan; Yuting, Duan. *Research of Intelligent Intrusion Detection System Based on Web Data Mining Technology*. In: Business Intelligence and Financial Engineering (BIFE), 2011 Fourth International Conference on. IEEE, 2011. p. 14-17.

[85] Mabu, Shingo et al. *An intrusion-detection model based on fuzzy class-association-rule mining using genetic network programming*. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, v. 41, n. 1, p. 130-139, 2011.

[86] Koc, Levent; Mazzuchi, Thomas A.; Sarkani, Shahram. *A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier*. Expert Systems with Applications, v. 39, n. 18, p. 13492-13500, 2012.

[87] Symons, Christopher T.; Beaver, Justin M. *Nonparametric semi-supervised learning for network intrusion detection: combining performance improvements with realistic in-situ training*. In: Proceedings of the 5th ACM workshop on Security and artificial intelligence. ACM, 2012. p. 49-58

[88] Saha, Sriparna et al. *Genetic algorithm combined with support vector machine for building an intrusion detection system*. In: Proceedings of the International Conference on Advances in Computing, Communications and Informatics. ACM, 2012. p. 566-572.

[89] Sangkatsanee, Phurivit; Wattanapongsakorn, Naruemon; Charnsripinyo, Chalermpol. *Practical real-time intrusion detection using machine learning approaches*. Computer Communications, v. 34, n. 18, p. 2227-2235, 2011.

[90] ISO/IEC. ISO/IEC 27002: code of practice for information security management 2005

[91] Joo, Daejoon; HONG, Taeho; HAN, Ingoo. *The neural network models for IDS based on the asymmetric costs of false negative errors and false positive errors*. Expert Systems with Applications, v. 25, n. 1, p. 69-75, 2003.

[92] Crow, Brian P. et al. IEEE 802.11 Wireless local area networks. IEEE Communications magazine, v. 35, n. 9, p. 116-126, 1997.

[93] Kuang, Fangjun; XU, Weihong; ZHANG, Siyang. *A novel hybrid KPCA and SVM with GA model for intrusion detection*. Applied Soft Computing, v. 18, p. 178-184, 2014.

[94] Chaudhary, Alka; Tiwari, V. N.; Kumar, Anil. *Design an anomaly based fuzzy intrusion detection system for packet dropping attack in mobile ad hoc networks*. In: Advance Computing Conference (IACC), 2014 IEEE International. IEEE, 2014. p. 256-261.

[95] Al-Jarrah, Omar; Arafat, Ahmad. *Network Intrusion Detection System using attack behavior classification*. In: Information and Communication Systems (ICICS), 2014 5th International Conference on. IEEE, 2014. p. 1-6.

[96] Mithell, Robert; Chen, Ray. *Adaptive intrusion detection of malicious unmanned air vehicles using behavior rule specifications*. IEEE Transactions on Systems, Man, and Cybernetics: Systems, v. 44, n. 5, p. 593-604, 2014.

[97] Shamshirband, Shahaboddin et al. *D-FICCA: A density-based fuzzy imperialist competitive clustering algorithm for intrusion detection in wireless sensor networks.* Measurement, v. 55, p. 212-226, 2014.

[98] Meng, Yuxin; Kwok, Lam-For. *Adaptive non-critical alarm reduction using hash-based contextual signatures in intrusion detection*. Computer Communications, v. 38, p. 50-59, 2014.

[99] Kim, Sun-il; Edmonds, William; Nwanze, Nnamdi. *On GPU accelerated tuning for a payload anomaly-based network intrusion detection scheme*. In: Proceedings of the 9th Annual Cyber and Information Security Research Conference. ACM, 2014. p. 1-4.

[100] Lin, Wei-Chao; KE, Shih-Wen; Tsai, Chih-Fong. *CANN: An intrusion detection system based on combining cluster centers and nearest neighbors*. Knowledge-based systems, v. 78, p. 13-21, 2015.

[101] Faisal, Mustafa Amir et al. *Data-stream-based intrusion detection system for advanced metering infrastructure in smart grid: A feasibility study*. IEEE Systems Journal, v. 9, n. 1, p. 31-44, 2015.

[102] Summerville, Douglas H.; ZACH, Kenneth M.; CHEN, Yu. *Ultra-lightweight deep packet anomaly detection for Internet of Things devices*. In: 2015 IEEE 34th International Performance Computing and Communications Conference (IPCCC). IEEE, 2015. p. 1-8.

[103] Zhao, Yanjie. *Network intrusion detection system model based on data mining*. In: 2016 17th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD). IEEE, 2016. p. 155-160.

[104] Esmaiily, Jamal; Moradinezhad, Reza; et al. *Intrusion detection system based on Multi-Layer Perceptron Neural Networks and Decision Tree*. In: Information and Knowledge Technology (IKT), 2015 7th Conference on. IEEE, 2015. p. 1-5.

[105] Thang, Dang Duy; Nam, Le Hoai; Khoi, Nguyen Tan. *Poster: Developing an Intrusion Detection System for Cloud Computing*. In: Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services Companion. ACM, 2016. p. 20-20.

[106] Bacs, Andrei et al. *Slick: an intrusion detection system for virtualized storage devices*. In: Proceedings of the 31st Annual ACM Symposium on Applied Computing. ACM, 2016. p. 2033-2040.

[107] Mechtri, Leila et al. *An IDS-based Self-healing Approach for MANET Survival*. In: Proceedings of the International Conference on Intelligent Information Processing, Security and Advanced Communication. ACM, 2015. p. 82.

[108] Tabrizi, Farid Molazem; Pattabiraman, Karthik. *Intrusion Detection System for Embedded Systems*. In: Proceedings of the Doctoral Symposium of the 16th International Middleware Conference. ACM, 2015. p. 9.

[109] Bronte, Robert; Shahriar, Hossain; Haddad, Hisham M. *A Signature-Based Intrusion Detection System for Web Applications based on Genetic Algorithm*. In: Proceedings of the 9th International Conference on Security of Information and Networks. ACM, 2016. p. 32-39.

[110] Kolias, Constantinos et al. *Intrusion detection in 802.11 networks: empirical evaluation of threats and a public dataset*. IEEE Communications Surveys & Tutorials, v. 18, n. 1, p. 184-208, 2015.

[111] Can, Okan; Sahingoz, Ozgur Koray. *An intrusion detection system based on neural network*. In: 2015 23nd Signal Processing and Communications Applications Conference (SIU). IEEE, 2015. p. 2302-2305.

[112] Tamimi, Ali; Kavianpour, Sanaz. *An Intrusion Detection System Based on NSGA*-II Algorithm.

[113] Hassanzadeh, Amin et al. *PRIDE: A practical intrusion detection system for resource constrained wireless mesh networks*. Computers & Security, v. 62, p. 114-132, 2016.

[114] V. Jyothsna, , V.V. Rama Prasad *FCAAIS: Anomaly based network intrusion detection through feature correlation analysis and association impact scale* In he Korean Institute of Communications Information Sciences