IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received 5 April 2024, accepted 23 May 2024, date of publication 29 May 2024, date of current version 12 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3406939

SURVEY

The Missing Link in Network Intrusion Detection: Taking AI/ML Research Efforts to Users

KATHARINA DIETZ [1], (Student Member, IEEE), MICHAEL MÜHLHAUSER [2], JOCHEN KÖGEL [3], STEPHAN SCHWINGER [4], MARLEEN SICHERMANN [1], MICHAEL SEUFERT [5], (Senior Member, IEEE), DOMINIK HERRMANN [2], AND TOBIAS HOßFELD [1], (Senior Member, IEEE)

[1] Chair of Communication Networks, University of Würzburg, 97070 Würzburg, Germany
[2] Privacy and Security in Information Systems Group, University of Bamberg, 96045 Bamberg, Germany
[3] IsarNet Software Solutions GmbH, 85356 Munich, Germany
[4] genua GmbH, 85551 Kirchheim, Germany
[5] Chair of Networked Embedded Systems and Communication Systems, University of Augsburg, 86159 Augsburg, Germany

Corresponding author: Katharina Dietz (katharina.dietz@uni-wuerzburg.de)

This work was supported in part by German Federal Ministry of Education and Research (BMBF) as part of WINTERMUTE under Grant 16KIS1128, Grant 16KIS1129, Grant 16KIS1131, and Grant 16KIS1132; in part by BMBF as part of explanym under Grant 16KISA079K; in part by Deutsche Forschungsgemeinschaft (DFG) as part of UserNet under Grant SE 3163/3-1 (project number: 500105691); and in part by the Bavarian Ministry of Economics, Regional Development and Energy (StMWI) as part of VIPNANO, under Grant DIK-2307-0005 and Grant DIK-2307-0006.

ABSTRACT Intrusion Detection Systems (IDS) tackle the challenging task of detecting network attacks as fast as possible. As this is getting more complex in modern enterprise networks, Artificial Intelligence (AI) and Machine Learning (ML) have gained substantial popularity in research. However, their adoption into real-world IDS solutions remains poor. Academic research often overlooks the interconnection of users and technical aspects. This leads to less explainable AI/ML models that hinder trust among AI/ML non-experts. Additionally, research often neglects secondary concerns such as usability and privacy. If IDS approaches conflict with current regulations or if administrators cannot deal with attacks more effectively, enterprises will not adopt the IDS in practice. To identify those problems systematically, our literature survey takes a user-centric approach; we examine IDS research from the perspective of stakeholders by applying the concept of personas. Further, we investigate multiple factors limiting the adoption of AI/ML in security and suggest technical, non-technical, and user-related considerations to enhance the adoption in practice. Our key contributions are threefold. (*i*) We derive personas from realistic enterprise scenarios, (*ii*) we provide a set of relevant hypotheses in the form of a review template, and (*iii*), based on our reviews, we derive design guidelines for practical implementations. To the best of our knowledge, this is the first paper that analyzes practical adoption barriers of AI/ML-based intrusion detection solutions concerning appropriateness of data, reproducibility, explainability, practicability, usability, and privacy. Our guidelines may help researchers to holistically evaluate their AI/ML-based IDS approaches to increase practical adoption.

INDEX TERMS Anomaly detection, artificial intelligence, intrusion detection, machine learning, network monitoring, privacy, security, usability.

I. INTRODUCTION

Artificial Intelligence (AI) and particularly Machine Learning (ML) have experienced a surge in popularity in network monitoring and management in recent years. Even though we see a lot of AI/ML-driven research on cybersecurity, its adoption in practice remains limited, as it is still facing key challenges [1]. According to Vielberth et al. [2], academic research on Security Operations Centers (SOCs) lacks a holistic view, only looks at fragments, and while it focuses

The associate editor coordinating the review of this manuscript and approving it for publication was Lei Shu.

on both human and technical aspects, it neglects their connection. By including AI/ML approaches, this disparity increases even further, as models introduce more complexity, reducing the trust of AI/ML non-experts.

Intrusion Detection Systems (IDS) are valuable tools in cybersecurity environments, and have been an academic research focus since decades. Intrusion detection is a classical classification problem (both supervised and unsupervised), and IDS in-the-wild are by large based on signatures and thus incapable of coping with new and unseen intrusion attempts. There are different types of IDS, such as host-based IDS (HIDS) and network-based IDS (NIDS). In this survey, we focus on AI/ML-based NIDS, which are placed in the network that perform this classification task mainly on a per-packet- or per-flow-basis.

Academic research in the area of such NIDS faces several issues, which can be grouped into different topics. Challenges exist regarding the data, code, model, practicability, and privacy, as well as user-centric topics like explainability and usability. The former problems mainly relate to technical facets, such as data appropriateness or data and code availability to foster reproducibility. In addition to these more generic challenges, other technical problems persist regarding the required monitoring efforts, overall performance, as well as requirements regarding software, hardware, and confidentiality. Apart from technical challenges, the second group of problems addresses potential users. While explainability is concerned with making decisions of AI/ML models transparent by extracting relevant information, usability goes one step further by providing interfaces to actually display this information appropriately to users, potentially even allowing to interact with the tool.

The above aspects are often disjoint theories in academia, while in reality there is always an interplay between all of them, hindering the practical implementation into the real world. In this work, we look at academic research on the topic of AI/ML-based network intrusion detection and analyze published papers through the eyes of the potential users of such tools to investigate the above challenges. In this context, users are administrators or other stakeholders with different needs, which we extract from two example enterprises. Thus, we follow a user-centric approach for our survey and review recent literature w.r.t. a multitude of technical and usability-related aspects, to analyze the following research questions (RQs):

RQ1: Which technical factors contribute to the poor adoption of AI/ML-based NIDS research?

RQ2: What user-centric aspects should research consider to improve the practical adoption of AI/ML-based IDS?

RQ3: What design guidelines should researchers follow to prepare research results for product integration?

To answer the above questions, the contributions of this paper are as follows:

1) We identify personas that are actually involved with AI/ML-based security tools and discuss their needs in that context.
2) Derived from our personas, we present hypotheses on why AI/ML-based NIDS approaches have not yet been widely adopted.
3) We evaluate existing approaches in the scientific literature based on our hypotheses.
4) We publish[1] our template and reviews, as well as evaluation code.
5) We give design guidelines for future research and discuss these guidelines, i. e., why not all guidelines can always be met simultaneously.

The remainder of this work is structured as follows. Section II reviews existing surveys in the domain of IDS and related topics, with a special focus on AI/ML-based approaches. The differences to our survey and the goal of our survey are highlighted. Section III introduces the applied methodology, i. e., how existing research papers were analyzed in detail. To this end, the concept of personas and their needs for two enterprise scenarios are explained. We also provide our strategy how AI/ML-focused works in literature were selected and analyzed. Then, 17 hypotheses about why those research works are not deployed and adopted in practice are introduced. Section IV presents the results of our literature survey. In particular, the formulated hypotheses are carefully checked for all papers identified in our survey and the key insights are summarized. As a result, Section V gives some concrete guidelines and key take-aways for researchers, which are of utmost importance for AI/ML-based solutions to be adopted in practice. Finally, Section VI concludes this work with a brief summary of the most important insights.

II. RELATED WORK

Several survey papers have already been published on IDS and related topics, shedding light on various aspects of network monitoring and management in general. The surveys by Boutaba et al. [3] and Nguyen and Armitage [4] provide comprehensive insights into traffic classification via AI/ML, in which intrusion detection is one of the use cases. In addition, Guimaraes et al. [5] survey visualization techniques for network management, though, the survey does not specifically zero in on AI/ML-driven approaches. However, intrusion detection is recognized as the by far biggest use case for network management, which shows its significance. Decade-old surveys [6], [7], [8] on IDS demonstrate this long-standing importance even further, and there exists also a plethora of surveys on NIDS specifically [9], [10], [11], focusing on this subcategory of IDS.

Recent surveys on intrusion detection and explainability [12], [13], [14] explore the application of Explainable Artificial Intelligence (XAI) techniques in the context of IDS, partially in different settings compared to enterprise networks, like the Internet of Things (IoT). While these surveys investigate approaches making IDS more transparent, our work goes beyond explainability by considering usability and addressing other secondary concerns for practical

[1] https://github.com/wintermute-project/missing-link-NIDS

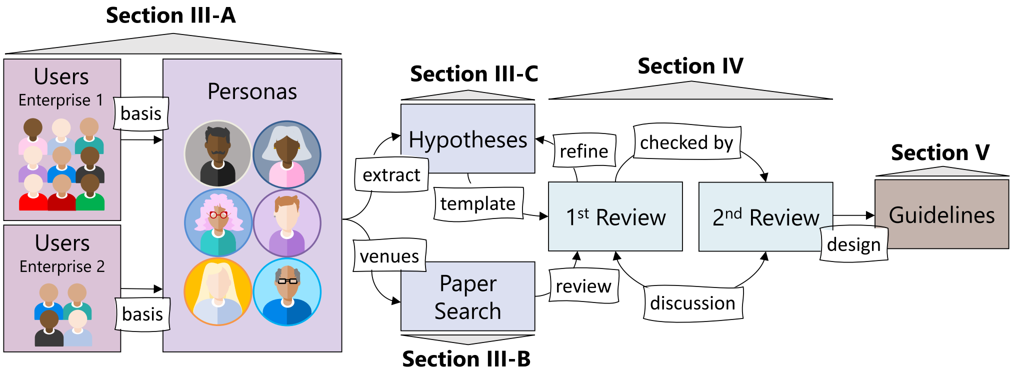

FIGURE 1. The workflow of our literature review. Each step also corresponds to a specific (sub)section, where the components are described in-depth.

deployment in the real world. While explainability is an important aspect, it is just one component among several factors necessary for the successful implementation of IDS in a real-world environment. Furthermore, surveys on Human-Centric Machine Learning (HCML) [15], [16] explore the broader connection between human users and AI/ML, although they do not directly address IDS techniques or are not in the context of network traffic in general.

Table 1 summarizes the above findings in order to make the contribution of this paper clearer. The table showcases a few selected topics we address in this paper and illustrates the disjunction between the various research fields clearly. Most of the related surveys have two or three focal points which they excel in, while neglecting other factors. While the surveys in [3] and [4] present us with thorough insights into anything related to AI/ML-based traffic classification and touch lightly upon privacy issues, they do not necessarily focus on IDS, though it is one of the use cases. Thus they also do not investigate any challenges that come along with it, e. g., practicability, explainability, or usability issues of IDS. Complementary to these two surveys, the work by Guimaraes et al. [5] focuses mainly on visualization with IDS as the most prominent use case, but does not zero in on AI/ML and other factors. Older surveys on IDS also do not focus on AI/ML but on rather more traditional approaches, yet include practicability aspects like performance and possibly recommendations [6], [7], [8]. On the contrary, newer surveys [12], [13], [14] include explainability and – to a limited extent – usability or privacy aspects, while only partially focusing on the practicability and feasibility of IDS. Lastly, there exist surveys on AI/ML methods that focus wholly on aspects like usability [15], [16] and also privacy approaches [17], such as differential privacy or homomorphic encryption, but are not in the context of communication networks and IDS.

TABLE 1. Overview of related surveys and their main focus. (✔: focus, (✔): partial focus, ✘: no focus). The aspects on top are only an excerpt of challenges we discuss in our literature review.

References	AI/ML-based	IDS-centered	Practicability	Explainability	Usability	Privacy
[3], [4]	✔	(✔)	✘	✘	✘	(✔)
[5]	✘	✔	✘	(✔)	✔	✘
[6]–[8]	✘	✔	✔	✘	✘	(✔)
[9]–[11]	✔	✔	✔	✘	✘	✘
[12]–[14]	✔	✔	(✔)	✔	(✔)	(✔)
[15], [16]	✔	✘	✘	✔	✔	✘
[17]	✔	✘	✘	✘	✘	✔

The main difference of the presented literature review is that instead of posing an in-depth discussion of dozens of papers, this survey focuses on a more high-level evaluation of recent literature on NIDS research. This allows to include not only all of the aspects presented in Table 1 into our survey, but many additional factors as well. In other words, we take many aspects necessary for an operational deployment of academic research in a real-world network into consideration, and ultimately derive several guidelines for future research. For this, we design personas, hypotheses, and provide a review template to help researchers identify the most relevant review criteria. The template might be used as a checklist for evaluating future work, such that the potential for deployment in a real-world network security system can be improved.

III. METHODOLOGY

Our goal is to identify potential reasons why academic researchers' AI/ML-based NIDS approaches are not adopted broadly in practice. Fig. 1 illustrates our workflow and how we structured the paper accordingly. To understand the barriers between AI/ML-based academic research and the industry, we first introduce users of two real-world enterprise scenarios, from which we extracted so-called personas, that represent potential users of our IDS tool. As all personas have different objectives, we gain insights from multiple perspectives, i. e., departments or teams in an enterprise. This is the basis for our literature review process, where we searched for recent AI/ML-based NIDS papers published on relevant venues for the personas. Also based on our personas, we formulate hypotheses why the practical application of

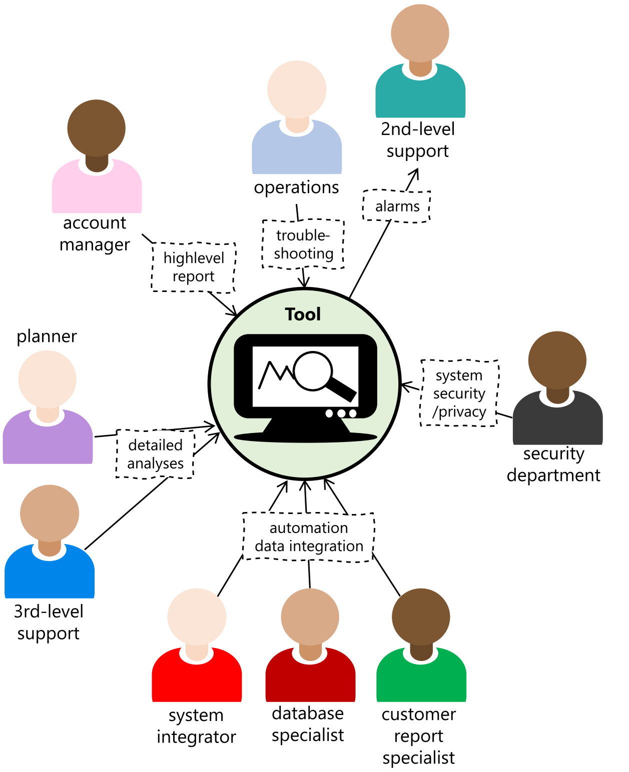

FIGURE 2. Users of example Enterprise 1. Users are characterized by their profession and possible interactions with a monitoring tool.

AI/ML-based NIDS approaches might fail. That is, each persona has potentially different requirements and concerns for introducing such a system. Integrating those different perspectives in our hypotheses, we get a realistic perspective on the adoption of AI/ML-based NIDS approaches.

These hypotheses serve as a template to review each of the papers found in the literature search, where we assess whether the hypotheses are true or false. In fact, there were two review iterations. In the first round, one researcher evaluated the paper individually and created a first review. In the second round, another researcher evaluated the paper and the first review to create a second review. In this way, each paper has been reviewed twice by different researchers and allows the two reviewers to reach a common ground. In practice, we did not perform those three steps (i. e., hypotheses design and first/second review) of our methodology consecutively, but rather in overlapping time periods with multiple feedback loops, as illustrated in Fig. 1. That is, we held weekly meetings and additional discussions between all reviewers to refine the template by adding, removing, and clarifying, until we converged to the final template. Then, the reviews are analyzed to identify the gaps between academia and industry. From these identified barriers we then design and discuss guidelines, that may help future research to increase practical adoption of their proposed IDS solution.

A. CONCERNED PERSONAS AND THEIR NEEDS

To understand the requirements of users of AI/ML-based NIDS approaches, we modeled prototypical users as personas. Personas are fictional characters representing different types of users. While personas initially are results of larger user studies, we used Gothelf's Proto-Persona approach [18], which is based on a company's experiences with customers. These Proto-Personas serve as a starting point for product evaluation, early design hypotheses, as well as raising awareness of the customer's view point [18].

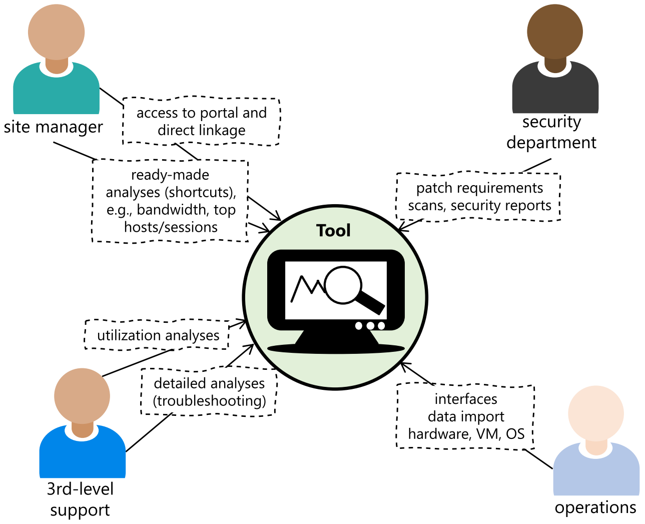
site manager
access to portal and direct linkage
security department
ready-made analyses (shortcuts), e.g., bandwidth, top hosts/sessions
patch requirements scans, security reports
Tool
utilization analyses
detailed analyses (troubleshooting)
interfaces data import hardware, VM, OS
3rd-level support
operations

FIGURE 3. Users of example Enterprise 2. Users depicted in a similar fashion compared to enterprise 1 have similar professions.

However, we found only one work using personas in network monitoring [19]. Therein, the authors introduce two personas (senior and junior network experts) and present user study results mapped and evaluated along these personas. In the visualization domain, McKenna et al. [20] define four personas along a management decision chain as a starting point and template for user-centered design. Additionally, Schufrin et al. [21] specify three personas in a monitoring context in a home network.

1) MONITORING TOOL USERS IN ENTERPRISES

We investigated the typical user base of network monitoring tools in large enterprises, as our experience with enterprise customers shows that the user base is much more diverse than those in related work [19], [20], [21]. There is not only a difference in age, experience, or management level, but experts with different domain knowledge interact with the system and pose distinct requirements onto it.

However, to describe a more diverse user base, we do not rely solely on our own experiences. We conducted interviews with contact persons for German enterprise customers of a monitoring tool vendor. Based on these interviews, we created two exemplary enterprises. Our two example enterprises are used to detail users of monitoring tools in practice. Based on similarities of the users of the two example enterprises, we derived generalized personas to better understand user requirements. To ensure that our personas reflect a diverse user base, each individual author

TABLE 2. Summary of derived personas from the previous exemplary enterprises. Each persona has their specific objectives, workflow, and interfaces that they prefer or require when working with a network monitoring tool.

	Maggie Manager	Charlie Compliance	Sandy Support	Paul Planner	Olga Operations	Nils Networker
Objectives	(internal) customer satisfaction	company security	2nd-level support, solve user problems	future: suitable and cost effective network and system capacity	system stability and up-to-dateness	3rd-level support, understand the network
Workflow	checking dashboards	policy creation, compliance checks	find issues in user/application context	check utilization, trends and compare current state to new requirements	maintain interfaces and keep configuration up-to-date	check and configure technical details
Interfaces	reports (PDF), pre-configured overview pages	health and compliance reports	interactive analysis/browsing logs	long term numbers via UI or exportable formats (CSV)	APIs, admin UI, logs, system-level views	UI with detailed technical views

contributed insights from their own field of expertise, e. g., network management, security, privacy, and usability.

a: EXAMPLE ENTERPRISE 1

The first enterprise runs a core network within Germany for providing connectivity and services to internal and external customers via own or leased lines. Fig. 2 shows the persons involved. Starting on top, people working in network operations and support use a monitoring tool for ticket processing, i. e., troubleshooting and finding the cause of reported network problems. Below, planners and 3rd-level supports perform more detailed data analysis. The security department on the right considers system security, while the remaining persons are concerned with system and data integration.

b: EXAMPLE ENTERPRISE 2

The second enterprise has a network connecting its global production and retail sites. Fig. 3 shows the persons involved in Example Enterprise 2. The site managers get direct access to ready-made simple views for 1st- and 2nd-level support. Below, 3rd-level supports require more sophisticated utilization or detailed analysis for troubleshooting. Top right, the security department is concerned with requirements and operative security, while operations on the bottom right deals with data import and hardware.

2) DERIVED PERSONAS

Our two example enterprises show different users of monitoring tools and serve as input for the abstract personas. Table 2 lists our personas and describes their objectives, usual workflow, and used interfaces. In the following, we explain our different personas in more detail.

a: MAGGIE MANAGER

has a responsible position in the network department and acts as contact point for internal and external customers regarding contracts for service delivery. She maps to the account manager in Example Enterprise 1 and is also interested in cost-effective solutions.

b: CHARLIE COMPLIANCE

is the Chief Information Security Officer (CISO) (or in a similar position) and acts as a proxy for external data security officers. He maps to the security department person of our two examples and manages documents on enterprise policies and ensures the compliance with legal requirements. Charlie specifies requirements on the functionality and security criteria and evaluates the system. He needs detailed threat reports and utilizes the tool to check the fulfillment of general policies.

c: SANDY SUPPORT

is a 2nd-level customer support person and maps to the corresponding support or site manager persons of our examples. Her workflow is ticket-driven and initiated by alarms raised by the tool that automatically opens tickets. She uses the tool to locate the origin of problems. To achieve this, she uses cheat-sheets containing possible solutions and best-practices and also maintains these sheets.

d: PAUL PLANNER

implements new requirements of the business strategy and the compliance into the network configuration. He checks long-term trends and formulates rules, as well as future requirements on the system. In Example Enterprise 1, there is a corresponding planner person, while in Example Enterprise 2, the utilization evaluation lays within the 3rd-level support.

e: OLGA OPERATIONS

takes care of the system itself (e. g., server operation, backups, updates, maintenance) and its integration into the IT landscape. She is concerned with system/data health and resolves tickets related to the system itself (e. g., tickets from Paul Planner). The three persons concerned with automation

TABLE 3. Surveyed conferences and journals. Each venue is associated with a relevant field for our survey and either had a high CORE ranking once during our reviewing period from 2018 to 2023, or was suggested by domain experts.

Focus	Rank	Abbr.	Conference/Journal
Artificial Intelligence	A*	AAAI	National Conference of the American Association for Artificial Intelligence
	A*	ICML	International Conference on Machine Learning
	B	IAAI	Innovative Applications in AI
Network Management	A*	SIGCOMM	ACM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication
	B	CNSM	International Conference on Network and Service Management
	B	ICC	IEEE International Conference on Communications
	B	IM	IFIP/IEEE International Symposium on Integrated Management
	B	NOMS	IEEE Network Operations and Management Symposium
	n/a	TNSM	IEEE Transactions on Network and Service Management
Privacy and Security	A*	CCS	ACM Conference on Computer and Communications Security
	A*	SP	IEEE Symposium on Security and Privacy
	A*	NDSS	USENIX Network and Distributed System Security Symposium
	A*	USS	USENIX Security Symposium
	A	ESORICS	European Symposium on Research in Computer Security
	B	TrustCom	International Conference on Trust, Security and Privacy in Computing and Communications
Usability, HCI, and Visualization	A*	CHI	Conference on Human Factors in Computing Systems
	A	TVCG	IEEE Transactions on Visualization and Computer Graphics
	B	SOUPS	Symposium on Usable Privacy and Security
	B	EuroVis	Eurographics/IEEE Symposium on Visualization
	B	CG&A	IEEE Computer Graphics and Applications
	B	PacificVis	IEEE Pacific Visualization Symposium
	B	AVI	Advanced Visual Interfaces
	C	VizSec	Symposium on Visualization for Cyber Security
	n/a	CGF	Computer Graphics Forum
	n/a	VIS	IEEE Visualization Conference

Data according to latest CORE rankings on http://portal.core.edu.au/conf-ranks/ and http://portal.core.edu.au/jnl-ranks/ if available

of Example Enterprise 1 as well as the operations person of Example Enterprise 2 map to Olga Operations.

f: NILS NETWORKER

solves 3rd-level tickets regarding complex network issues and possesses profound understanding of network-related details. Both of our examples show 3rd-level support persons, who want to understand problems on a fine-grained level and prefer a detailed network/data view, which is important for establishing trust in the system.

B. LITERATURE SELECTION CRITERIA

To find papers on AI/ML-based NIDS approaches, we performed a systematic literature review. In this section, we describe our literature selection criteria by outlining the inclusion and exclusion criteria for the individual papers. Overall, 13 people (8 authors + 5 helpers) participated in searching for or reviewing relevant papers and venues. We decided to consider only papers with AI/ML-based NIDS approaches published from Q1 2018 to Q4 2023[2] to provide a recent overview of this field.

In the first step, we identified established venues in the fields of networking, security, privacy, AI/ML, and UI/UX research to find relevant papers. We decided to focus on these four fields, as AI/ML venues potentially provide the most sophisticated AI/ML models and network management is the exact context of our use case. Furthermore, we derived from our persona-based approach that secondary concerns, such as privacy (Charlie Compliance) and usability/visualization (Sandy Support, Nils Networker, Paul Planner) are vital aspects and hence we included privacy and visualization/usability venues. That means, all these research fields provide different facets important for IDS design.

For our literature review, we selected venues based on their CORE ranking; they should have ranked at least A once in our reviewing period from 2018 to 2023. Additionally, we included some conferences and journals suggested by domain experts in those fields. Table 3 shows the conferences and journals we considered in our literature review. If we found any associated workshops for those venues, we also searched for relevant papers there.

In the second step, we searched for relevant papers in the proceedings for each of these venues. Papers were tagged as relevant in the first place by one person as soon as the paper deals with intrusion or anomaly detection in any way. Additionally, to ensure the comprehensiveness of the selected papers, each of the relevant venues was searched through mainly by domain experts, e. g., the privacy and security conferences were assigned to someone with expertise in that field.

After our initial paper collection, we excluded papers that match at least one of the following exclusion criteria: We did not consider papers that do not contain at least once the term *anomaly detection* or *intrusion detection*. Further, we excluded papers without AI/ML-based IDS approaches. Specifically, we did not include papers where the approach does not use network traffic features since

[2] Some conferences that took place in Q4 22/23 might not be included here, since their proceedings were not published when the literature search was conducted.

TABLE 4. Overview of established hypotheses and relevancy for the six personas. For each hypothesis, the table displays if a persona is concerned with it (✓: directly relevant, (✓): indirectly relevant, ✗: not relevant). Each hypothesis is also assigned a topic for further reference.

ID	Hypothesis	Topic	Manager	Compliance	Support	Planner	Operations	Networker
H1	Data too old[a]	Data	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)
H2	Traffic mix not explained	Data	✓	✗	✗	✗	✗	✓
H3	Data not available	Data	✗	✗	✗	✗	✗	✓
H4	(Pseudo)code[b] not available	Code & Model	✗	✗	✗	✗	✗	✓
H5	Potential model overfitting, not generalizable to own scenario	Code & Model	✗	✓	✗	✗	✗	✓
H6	Unusable in practice, since required data problematic to obtain	Practicability	✗	✓	✗	✗	✗	✗
H7	Unrealistic or high effort for data collection and monitoring	Practicability	✓	✗	✗	✓	✓	✗
H8	Special or excessive processing hardware requirements	Practicability	✓	✗	✗	(✓)	✓	✗
H9	FPR hinders practical adoption	Practicability	✗	✗	✓	✗	✗	✓
H10	Model complexity too high, missing trust and explainability	Understandability	✗	✓	✗	✓	✗	✓
H11	Features not motivated or too complex, missing trust	Understandability	✗	✓	✗	✓	✗	✓
H12	Unclear meaning of decision and result	Understandability	✗	✗	✓	✗	✗	✓
H13	Important usability and interaction features missing	Secondary Concerns	✓	✗	✓	✗	✗	✓
H14	Privacy not in researchers' focus	Secondary Concerns	✗	✓	✗	✗	✗	✗
H15	No comparison with state-of-the-art given	Contextualization	✓	✓	✗	✗	✗	✗
H16	Ignored domain knowledge	Contextualization	✗	✗	✗	✓	✗	✓
H17	No discussion of limitations	Contextualization	✓	✓	✓	✓	✓	✓

[a]with respect to newest utilized dataset, [b]in the strict sense/PASCAL-like

our focus is on network intrusion detection. That means, we excluded, for example, log-based IDS papers and papers analyzing malicious PDF documents or malware executables. Additionally, we excluded SCADA/ICS, Cyber Physical Systems, Internet of Vehicle, Internet of Things, Smart Home, and Smart Grid papers as we focus on enterprise networks. Further, we found some borderline papers, where we had to discuss internally whether we want to include those papers. To account for those borderline papers, two researchers reviewed the papers individually. Afterwards, the two researchers made the final decision about the exclusion of the paper based on their mutual agreement for each paper.

C. HYPOTHESES ON WHY AI/ML APPROACHES FAIL

To evaluate the identified papers of our literature review, we created a template containing hypotheses why AI/ML-based IDS approaches may not broadly be adopted in practice. Table 4 illustrates the various hypotheses, the designed personas, and maps them to each other in terms of relevance. The table already shows that the personas all view the AI/ML tool from different (possibly conflicting) perspectives, i. e., each persona is subject to their own requirements. The table also maps each hypothesis to one of six topics: data, code and model, practicability, understandability, secondary concerns, and contextualization. In the following paragraphs, we briefly describe all of the 17 hypotheses we used in our template and the reasoning behind them by relying on our personas as well as supporting our hypotheses with observations from related studies. We also discuss in-depth which hypothesis is relevant for which personas.

1) H1: DATA TOO OLD

The first three hypotheses (H1 to H3) are about the data used in the paper. It is key to utilize representative, up-to-date data when developing AI/ML tools, as otherwise the derived conclusions may not be meaningful in the context of today's networks. Yet, many decade-old datasets are still frequently used in recent publications as benchmarking sets [22]. While this is a valid approach to compare emerging methods to legacy ones, newer datasets should still be utilized additionally. This hypothesis is indirectly relevant for all personas, as it is concerned with a more general problem in research, rather than being related to a specific persona task or requirement. As reference, we decided on several milestones for this hypothesis, where the Internet traffic landscape changed, with 2016 and the introduction of the QUIC protocol being the most recent one, and 2001 and the creation of Wikipedia the earliest one. Since this hypothesis is a more generic challenge, it is a technical factor related to RQ1.

2) H2: TRAFFIC MIX NOT EXPLAINED

While there exist plenty of datasets from more recent years, the age of the data is not the only issue [22]. Other problems include unrepresentative data due artificial setups in academic networks, testbeds, labs, or simulations and emulations. Thus, at minimum, there should be an objective description of the traffic composition, and, ideally, a justification for using this dataset. This information helps Maggie Manager and Nils Networker to assess if the proposed tool is relevant in context of their use case, from a high-level managing perspective and a low-level networking perspective, respectively. Like H1, this is a technical factor concerned with RQ1.

3) H3: DATA NOT AVAILABLE

Besides the data quality and its justification, public availability and consequently reproducibility is another important factor when considering practical implementation in a real-world system. Without accessible data, it is impossible to replicate results, which hinders the insight and trust into the monitoring tool. This is relevant for Nils Networker, as he is the one eventually implementing the tool into the productive system. Similar to the first two hypotheses, this is also a technical factor linked to RQ1.

4) H4: (PSEUDO)CODE NOT AVAILABLE

The next two hypotheses are about the code and model (H4 and H5). Analogous to data availability, the code availability is another component for reproducible results. Without (pseudo)code, it is a complex, time-consuming task to recreate the methodology, as many AI/ML models exhibit a myriad of parameters to configure. Thus, in the best case there exists a publicly available code repository. Again, this is relevant for Nils Networker, as he is the one incorporating the tool into the running system. Analogous to the data-related hypotheses, H4 is relevant for RQ1, since it is a technical factor.

5) H5: POTENTIAL MODEL OVERFITTING, NOT GENERALIZABLE TO OWN SCENARIO

Generalizability is yet another important factor to consider when assessing the integration of a new approach into a running system. An emerging approach may outperform existing ones on one dataset, but might underperform on other datasets. Consequently, testing an approach on the basis of more than one dataset is desirable. For models that are evaluated on the basis of simulated data, this generalizability is potentially problematic, as these models are often optimized for a single environment [23]. Lastly, it is also important to consider network changes [23] and evaluate the tool's capability to adjust to a new environment [1], e. g., involving transfer learning approaches to reuse old models or evaluating covariate shifts and concept drifts. This is important for Charlie Compliance, as he is responsible to create fitting rules (potentially based on decisions of AI/ML models) and Nils Networker, as – similar to the previous hypotheses – he is responsible for the implementation. This hypothesis is another technical factor impacting RQ1.

6) H6: UNUSABLE IN PRACTICE, SINCE REQUIRED DATA PROBLEMATIC TO OBTAIN

The following four hypotheses are concerned with potential practicability issues (H6 to H9), since efficiency in general is a non-negligible factor for users when choosing AI/ML tools [1]. Due to privacy and confidentiality requirements, the practically available data is often either limited to web proxy and firewall logs [23] (e. g., DNS activity may be problematic from a privacy perspective [24]) or to very comprehensive datasets like full-resolution packet traces. Acquiring packet traces in existing environments is complicated due to multiple compliance problems and organizational overhead. In our opinion, NetFlow/IPFIX is a suitable data foundation for many practical deployments, as it is the predominant data format for many vendors [25]. This hypothesis is important for Charlie Compliance, as he is responsible to fulfill data legal and confidentiality requirements. Further, since this is mainly related to practically available data, it is another influencing factor for RQ1.

7) H7: UNREALISTIC OR HIGH EFFORT FOR DATA COLLECTION AND MONITORING

Besides the data format and type, the amount of the monitored data may pose a problem. In practice, monitored data is often sampled or aggregated (e. g., due to the usage of NetFlow), rendering approaches that require full-resolution network traffic in the form of packet captures (PCAPs) infeasible. Sometimes the initial data needed to train the model spans months, which may make a fast deployment infeasible. This is important for Maggie Manager, as this is a budget-related issue, Olga Operations due to hardware requirements, and Paul Planner due to network requirements (bandwidth, taps, etc.). Like H7, this is related to RQ1, since it is concerned with mainly technical factors.

8) H8: SPECIAL OR EXCESSIVE PROCESSING HARDWARE REQUIREMENTS

In conjunction with the previous two hypotheses, special or excessive hardware requirements may also hinder the applicability in the real world, as extensive resources are not always available [1], e. g., clusters, GPU/CUDA, or specialized hardware such as P4. Similar to before, this is important for Maggie Manager, as this is budget-related, Olga Operations (e. g., due to possible operations on VMs), and – depending on his scope – maybe Paul Planner due to network requirements. Since this hypothesis is concerned with the actual hardware, it depicts another influencing factor for RQ1.

9) H9: FALSE POSITIVE RATE (FPR) HINDERS PRACTICAL ADOPTION

For a practical usage of proposed AI/ML methods, it is essential to keep the false alarm rate low and to provide the severity level of the alarm. Otherwise, admins have to handle too much workload when dealing with security alerts. For example, a FPR of over 1 % might already be too high when looking at big volumes of traffic. As reported by Mink et al. [1], AI/ML models are generally perceived to have a higher FPR than traditional, rule-based solutions by users. Thus, besides measures like accuracy, an evaluation of the false positives (FPs) is important. The FPR should be kept as low as possible by finding a feasible trade-off with the true positive rate (TPR). Alternatively, AI/ML methods should provide measures, such as severity or confidence [26], to the admins to make the workload more bearable. In our use case, Sandy Support and Nils Networker are concerned with these challenges, as Sandy is dealing with and asking others regarding FPs, and Nils Networker handles them. Since this hypothesis is related to the performance of the model, it is mainly a technical factor linked to RQ1. However, since false alarms are potentially handled by the administrators, it indirectly also has an impact on user-centric aspects in RQ2.

10) H10: MODEL COMPLEXITY TOO HIGH FOR PERSONA, MISSING TRUST AND EXPLAINABILITY

The next three hypotheses (H10 to H12) are about explainability issues. After the data is collected, another possible issue may be the chosen AI/ML model, as an understandable and explainable model is important to build trust; otherwise it may result in mistrust and misuse [27]. Models deemed too complex may combine many models in sequence or parallel, may depend on many unintuitive or unexplained hyperparameters, or may be black-box or cutting-edge models. This is important for Nils Networker, Paul Planner, and Charlie Compliance, but not Sandy, Maggie, and Olga, as the latter have no interest in knowing why something was detected, while Nils, Paul, and Charlie do. H10 is mainly concerned with the respective abilities and skills of the users themselves to understand the model and therefore related to RQ2.

11) H11: FEATURES NOT MOTIVATED OR TOO COMPLEX, MISSING TRUST AND EXPLAINABILITY

Not only the model itself should be explainable, but also the input features, as abstract and unfamiliar features will disrupt the trust of the expert, even if the methodology is technically correct [26]. Thus, the feature set should be complete and in the domain context, but it should not include too many features (e. g., Miller's magic number 7±2 [28]) or complex transformations. For the same reasons as before, this is important for Nils Networker, Paul Planner, and Charlie Compliance. Analogous to understanding the AI/ML model, H11 is also a user-related concern and linked to RQ2.

12) H12: UNCLEAR MEANING OF DECISION AND RESULT

Extending on the previous two hypotheses, the output – and thus decision – should be explained properly in the use case context. Further, by adding confidence scores to the decision [1], [26], admins can focus on critical alerts. Other useful information are standard approaches, such as feature importances or the visualization of the decision path for tree-based methods. This is important for Sandy Support and Nils Networker, as both need to understand the output in order to answer tickets. Like the previous two hypotheses, the explainability of the result is another user-centric problem and concerned with RQ2.

13) H13: IMPORTANT USABILITY AND INTERACTION FEATURES MISSING

In the following two hypotheses (H13 and H14), secondary concerns are discussed. An AI/ML tool usable in practice does not only consist of the implemented algorithms, but also includes a GUI, that follows UI/UX/Usability standards and provides functionality, such as sorting or ignoring. Though, according to Oesch et al. [27], GUIs in practice are either non-existent, or violate established usability heuristics. Consequently, AI/ML security tools are generally perceived to have poor usability [1]. An appropriate GUI is important for Sandy Support, Nils Networker, and Maggie Manager, as, for instance, dashboards or management reports visualize all the relevant information to answer management questions or customer tickets. Naturally, since this hypothesis is concerned with usability itself, it is directly related to RQ2.

14) H14: PRIVACY NOT IN RESEARCHERS' FOCUS

Another important aspect for the deployment of a tool in practice are privacy concerns. With constantly evolving data protection laws and regulations, invasive methods such as DPI, SSL interception, or code on user devices cannot be trivially implemented nowadays. Even if the collected data seems unproblematic, adversaries can still infer sensitive data from encrypted or aggregated traffic by analyzing metadata with traffic analysis. Thus, well-considered anonymization or pseudonymization mechanisms are desirable. Naturally, Charlie Compliance is concerned with this hypothesis. This hypothesis is both a technical and a user-related concern. It potentially complicates the technical implementation of an AI/ML model (e. g., due to the implementation of methods such as differential privacy or homomorphic encryption) and may have an impact on the performance of the model, but also protects sensitive data of network participants. The latter generally have unique privacy perspectives and expectations [29].

15) H15: NO COMPARISON WITH STATE-OF-THE-ART (SOTA) GIVEN

The last three hypotheses (H15 to H17) are about contextualization. When considering a new approach to deploy, it is always important to consider alternatives. Thus, a comparison with SOTA approaches as a baseline is obligatory, e. g., comparisons with other AI/ML-based approaches and/or signature-based IDS, such as Suricata[3] or Snort.[4] Up-to-date approaches should also discuss early detection, zero-day attacks, and multi-stage attacks, as these are increasingly problematic to detect [30]. This hypothesis is important for Maggie Manager, as she is the one that wants to buy and include an AI/ML tool. Charlie Compliance is also concerned with this hypothesis as there might be requirements on the detection capabilities. Since this hypothesis is concerned with putting the performance of the proposed approach into context of performances of SOTA approaches, it is another technical factor for RQ1.

16) H16: IGNORED DOMAIN KNOWLEDGE

Often, context information is neglected [31] (e. g., topology information) and open-source or commercial sources are not considered [23] (e. g., CVEs). This is especially crucial when data from different sources is combined (e. g., via federated

[3] https://suricata.io/

[4] https://www.snort.org/

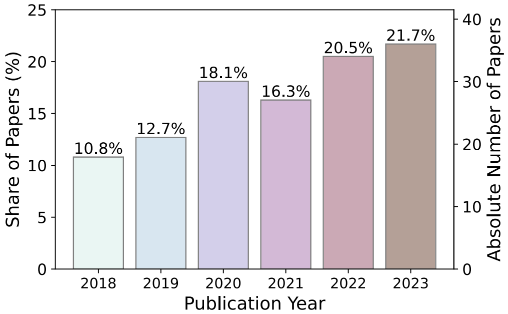

FIGURE 4. Percentage of relevant papers for this literature review published in the surveyed years.

learning [32]). Admins like Nils Networker and Paul Planner need this context information to give feedback [26] (e. g., via active learning approaches), as they are the ones that possess the required technical knowledge about the system. H16 is both a technical and user-related factor, and therefore linked to the first two RQs, since domain knowledge may come from different types of sources. In other words, sources like vulnerability databases and how to include them is more of a technical factor, while implementing a feedback loop to gain knowledge from the admins is directly related to the users.

17) H17: NO DISCUSSION OF LIMITATIONS

Last, it is important to also discuss a proposed model's weaknesses and limitations. Models may often be evaded or cannot handle unknown zero-day attacks. This is an important point to consider in practice, as this raises the questions whether the proposed approach is fit for one's use case or can be adapted properly. Limitations could concern personas from all departments. For instance, Maggie Manager might be interested to buy an AI/ML tool. Charlie Compliance and Nils Networker are also concerned with this hypothesis, as they want to know about potential problems prior to deployment. This allows them to fix these problems in order to comply with privacy-related and other requirements. H17 is mainly a technical factor and thus related to RQ1, as it directly influences the decision to deploy the proposed IDS.

IV. EVALUATION

After defining our set of hypotheses that we want to check for each paper, this section is concerned with the actual evaluation. Note that this section depicts mainly an objective analysis, before diving into a more subjective discussion in the next section, where we give recommendations and describe challenges or trade-offs. For this, we considered 166 papers on AI/ML-based NIDS approaches in our literature review. For all papers, we determined at first some general information. Secondly, we reviewed the hypotheses from the template for each paper individually.

A. GENERAL INFORMATION

In the following, we analyze the general information of the surveyed papers, including the venue, the year of publication, the used datasets, and AI/ML methods.

1) PUBLICATION VENUES

Table 5 lists the relevant papers we found at the publication venues introduced in Table 3. The most prominent venue among the 166 papers on AI/ML-based NIDS approaches is ICC; almost 25 % of all papers have been published there. TNSM is also a popular venue with over 23 % of the papers published there. Further, roughly 10 % of the papers have been published at the TrustCom conference with its associated BigDataSE workshop. We found 9 % of the papers at NOMS, as well as around 8 % at SIGSAC CCS including four workshops (on Moving Target Defense, on Security and Artificial Intelligence, on Cyber-Security Arms Race, and the Cloud Computing Security Workshop). Furthermore, 6 % of papers have been published at CNSM. For other venues, we found only a limited number of papers on AI/ML-based NIDS approaches. For instance, we found 3 % of the papers matching our inclusion criteria at IM and the USENIX Security Symposium. Additionally, we found roughly 2 % of the papers for each venue at AAAI, ICML (including its workshop on Machine Learning for Cybersecurity), SIGCOMM (including its associated Big-DAMA workshop), ESORICS (including its Workshop on Security and Artificial Intelligence), and NDSS. For the rest of the venues, we either found only one or no relevant papers.

The largest portion of relevant papers has been published at network management conferences, and the smallest portion at AI and usability/HCI/visualization conferences. Papers in the latter two fields generally work on other data (often images or text), or focus more on sophisticated methodologies themselves (AI/ML algorithms, visualization techniques, etc.). Only a small fraction of these papers proposes NIDS and utilizes network traffic data. Noteworthy, there are several works we found in the surveyed venues in the visualization domain. Those works are in our context of network security, but focus more on visual analytics, utilizing mainly statistical methods [33], [34], [35], [36], [37], and present the data directly to the user for manual investigation, instead of employing an intermediary AI/ML model.

2) YEAR OF PUBLICATION

All the papers we considered on AI/ML-based NIDS approaches have been published at venues between 2018 and 2023. Fig. 4 shows a notable and steady increase in papers throughout all six years, with exception of the years 2020 and 2021. 18 % of the selected papers have been published in 2020, while in the previous years 2019 and 2018 only 13 % and 11 % of the selected papers have been published, respectively. The number of published papers slightly decreased in 2021 in comparison to 2020 with 16 % of papers published that year. In 2022 and 2023 this number increased again. That is, 20 %, respectively 22 % of the papers have been published in those years.

3) USED DATASETS

Less than 29 % of papers found with our literature review created an own dataset for their use case. The remaining

TABLE 5. Papers per venue (n=166), excluding venues with no papers. The venues might include papers from associated workshops.

Abbr.	#	%	References
Artificial Intelligence (8 papers ∼ 4.8%)			
AAAI	4	2.4 %	[38], [39], [40], [41]
ICML	4	2.4 %	[42], [43], [44], [45]
Network Management (113 papers ∼ 68.1%)			
SIGCOMM	3	1.8 %	[46], [47], [48]
CNSM	10	6.0 %	[49], [50], [51], [52], [53], [54], [55], [56], [57], [58]
ICC	41	24.7 %	[59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86], [87], [88], [89], [90], [91], [92], [93], [94], [95], [96], [97], [98], [99]
IM	5	3.0 %	[100], [101], [102], [103], [104]
NOMS	15	9.0 %	[105], [106], [107], [108], [109], [110], [111], [112], [113], [114], [115], [116], [117], [118], [119]
TNSM	39	23.5 %	[120], [121], [122], [123], [124], [125], [126], [127], [128], [129], [130], [131], [132], [133], [134], [135], [136], [137], [138], [139], [140], [141], [142], [143], [144], [145], [146], [147], [148], [149], [150], [151], [152], [153], [154], [155], [156], [157], [158]
Privacy and Security (42 papers ∼ 25.3%)			
CCS	13	7.8 %	[159], [160], [161], [162], [163], [164], [165], [166], [167], [168], [169], [170], [171]
SP	1	0.6 %	[172]
NDSS	4	2.4 %	[173], [174], [175], [176]
USS	5	3.0 %	[177], [178], [179], [180], [181]
ESORICS	3	1.8 %	[182], [183], [184]
TrustCom	16	9.6 %	[185], [186], [187], [188], [189], [190], [191], [192], [193], [194], [195], [196], [197], [198], [199], [200]
Usability, HCI, and Visualization (3 papers ∼ 1.8%)			
TVCG	1	0.6 %	[201]
PacificVis	1	0.6 %	[202]
VizSec	1	0.6 %	[203]

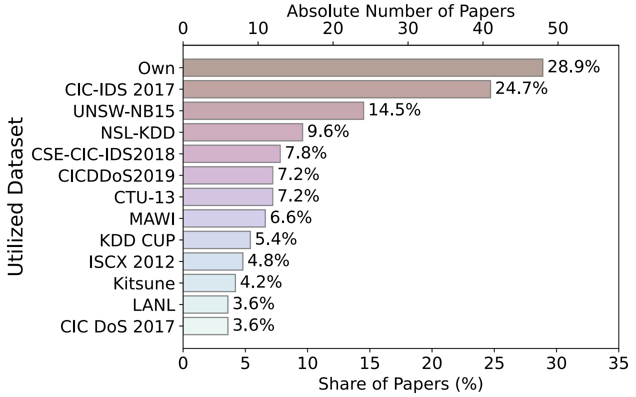

FIGURE 5. Top utilized datasets. Many papers used more than one dataset, so the sum of shares may exceed 100 %.

papers use existing datasets from related work. For our template, we utilized the list proposed by Ring et al. [204].

Fig. 5 depicts the top datasets found in the surveyed papers. Among the included papers, the most used dataset is the CIC-IDS 2017 [205] dataset, which is used in almost 25 % of papers. The UNSW-NB15 [12], NSL-KDD [206], and CSE-CIC-IDS2018 [205] datasets are also used quite often. While over 14 % of all the papers use the UNSW-NB15 dataset, almost 10 % use the NSL-KDD dataset, and 8 % use the CSE-CIC-IDS2018 dataset. Slightly less utilized are the CICDDoS2019 [207] dataset and the CTU-13 dataset [208] (both about 7 % of papers). In over 6 % of the papers, researchers make use of the traffic data repository maintained by the MAWI Working Group[5]/MAWILab [209]. Other datasets such as KDD CUP [206] and ISCX 2012 [210] are each used in roughly 5 % of the papers. In a limited number of papers, researchers also use other datasets, such as the Kitsune dataset [175], datasets from the Los Alamos National Laboratory (LANL),[6] or the CIC DoS 2017 dataset [211] (each roughly 4 % of papers).

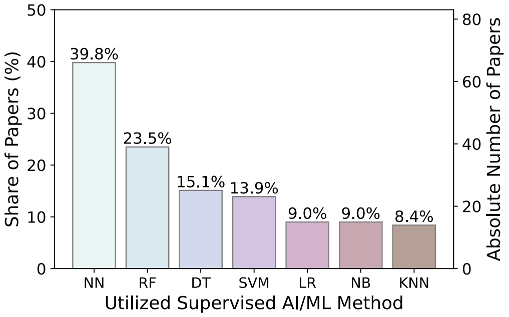

FIGURE 6. Top utilized AI/ML methods for supervised algorithms. Percentages are in relation to total number of papers.

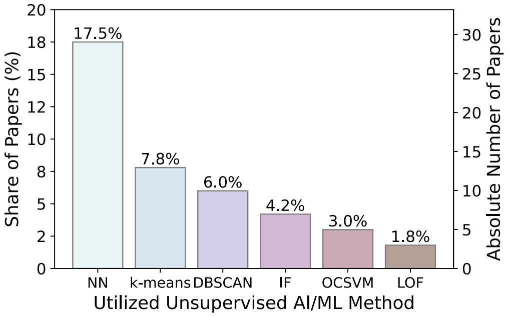

FIGURE 7. Top utilized AI/ML methods for unsupervised algorithms. Percentages are in relation to total number of papers.

4) AI/ML METHODS

As this paper is about AI/ML-based NIDS approaches, we further examine the used methods. 65 % of papers used some kind of supervised learning while 36 % of the papers used unsupervised learning. Note that some of the papers used both, supervised and unsupervised learning, contributing to both of the above numbers. We also found a number of papers using semi-supervised learning (5 %) or reinforcement learning (6 %).

[5] https://mawi.wide.ad.jp/mawi/

[6] https://www.lanl.gov/

Fig. 6 shows an overview about the supervised ML methods. When supervised learning is used in an AI/ML-based NIDS approach, the authors applied mainly Neural Networks (NNs, 40 % of papers), Random Forests (RFs, 23 % of papers), Decision Trees (DTs, 15 % of papers), or Support Vector Machines (SVMs, 14 % of papers). Some other techniques were less frequently used including Logistic Regression (LR), Naïve Bayes (NB, both 9 % of papers), and K-Nearest Neighbors (KNN, 8 % of papers). Besides the AI/ML models in the figure, some papers made use of AI/ML ensembles by combining various of the above models via boosting methods (e. g., AdaBoost, XGBoost, or LogitBoost), which was the case for almost 10 % of papers.

For unsupervised learning, Fig. 7 shows the most used ML methods. In that case, the authors applied also mainly NNs (17 % of papers). Further, the authors used k-means (8 % of papers), DBSCAN (6 % of papers), Isolation Forest (IF), and One-Class SVMs (OCSVMs, both 4 % of papers). Occasionally, we also found other typical anomaly detection algorithms, such as Local Outlier Factor (LOF, 2% of papers), which are less frequently used in the papers.

B. HYPOTHESES

Our hypotheses why AI/ML-based NIDS approaches might not be adopted in the industry are outlined in Section III-C. To answer RQ1 and RQ2, we evaluated whether each hypothesis is true or false for all of the 166 papers we identified in our literature review. Our results are shown in Fig. 8 and detailed in the following.

1) DATA

As shown in Fig. 8, we found for H1 that a non-negligible portion of papers use outdated data. That is, 33 % of the papers use old data, i. e., the latest dataset used is more than 8 years old (introduction of QUIC protocol in 2016). Additionally, there is no information on the data recording date for 15 % of the papers. Thus, only 52 % of the papers make clear to use recent datasets, i. e., not older than 2016 per our definition. However, for H2, we observe that the majority of papers is describing the data or traffic mix to some extent. As shown in Fig. 8, only 18 % of papers do not provide sufficient explanations for the data or traffic mix used in the paper. In H3, we assess whether datasets are mostly publicly available. Datasets are not available for 31 % of papers while 69 % of papers provide their data or use publicly available datasets.

2) CODE AND MODEL

In H4, we evaluate the reproducibility of the authors' work; that comprises the availability of source code or at least pseudo-code. We found that most of the papers do not provide any implementation of their approach. Fig. 8 shows that 79 % of the papers do not give any kind of code. Only 21 % of the papers provide some kind of implementation, e. g., a reference to a Git repository containing the source code. If code and data is directly available, it is often possible to evaluate the quality of the model. However, when a working link is not available in the paper, contacting authors directly may sometimes yield positive results but is unlikely to be successful in most cases as shown by Wieling et al. for a different use case [212].

For H5, we evaluate whether AI/ML models might be prone to overfitting. As shown in Fig. 8, we found that AI/ML models might not be generalizable in 64 % of the papers due to potential overfitting. Indicators for that are, for instance, the use of a single dataset, or simulated data [23].

3) PRACTICABILITY

The availability of data, code and AI/ML models are helpful to estimate the practical utility of any AI/ML-based NIDS approach. If enterprises want to adopt the approach, they still have to assess, whether it is feasible to collect the data in their own setting. In H6, we review how problematic data collection might be for enterprises [23], [24]. We found that data can be easily obtained for the approaches in 40 % of the papers, e. g., using NetFlow. In the other 60 % of papers, the approaches often use more data than provided by NetFlow. For instance, some papers use DNS data or information that require collecting complete PCAP files, e. g., higher statistical moment such as skewness or kurtosis (even variance would be problematic for NetFlow). In H7, we hypothesize that this brings high effort for data collection, so that enterprises might not adopt an approach. Fig. 8 shows that the effort for data collection is too large for 49 % of the papers.

Besides effort for collecting the data, enterprises have to consider running the AI/ML-based NIDS approach in production. That is, enterprises might need special hardware to run the NIDS approaches. Regarding H8, Fig. 8 shows that special hardware, such as a GPU or a cluster for faster calculations, is needed in 23 % of the papers. 43 % of papers provide no specification of the used hardware.

Furthermore, it is not feasible to handle large numbers of false alarms in production, i. e., alarms for intrusions that are actually benign. This is the reason, why we further hypothesize in H9 that the FPR has to be acceptable for production environments. Our results in Fig. 8 show that presented information regarding the FPR hinders practical adoption for 86 % of the papers. Actually, the FPR is above 1 % for 30 % of the papers – what we consider as too high in production. 57 % of papers do not report the FPR at all. Further, many papers do not elaborate on implications of a high FPR and do not interpret the results for the use case (60 % of papers). So, even if the FPR is reported in numbers, it is not explicitly discussed.

4) UNDERSTANDABILITY

If the different personas can not understand an AI/ML-based NIDS approach easily, decision makers likely will not adapt an approach. As shown in Fig. 8, we found for H10 that the model might be too complex for the relevant personas (see

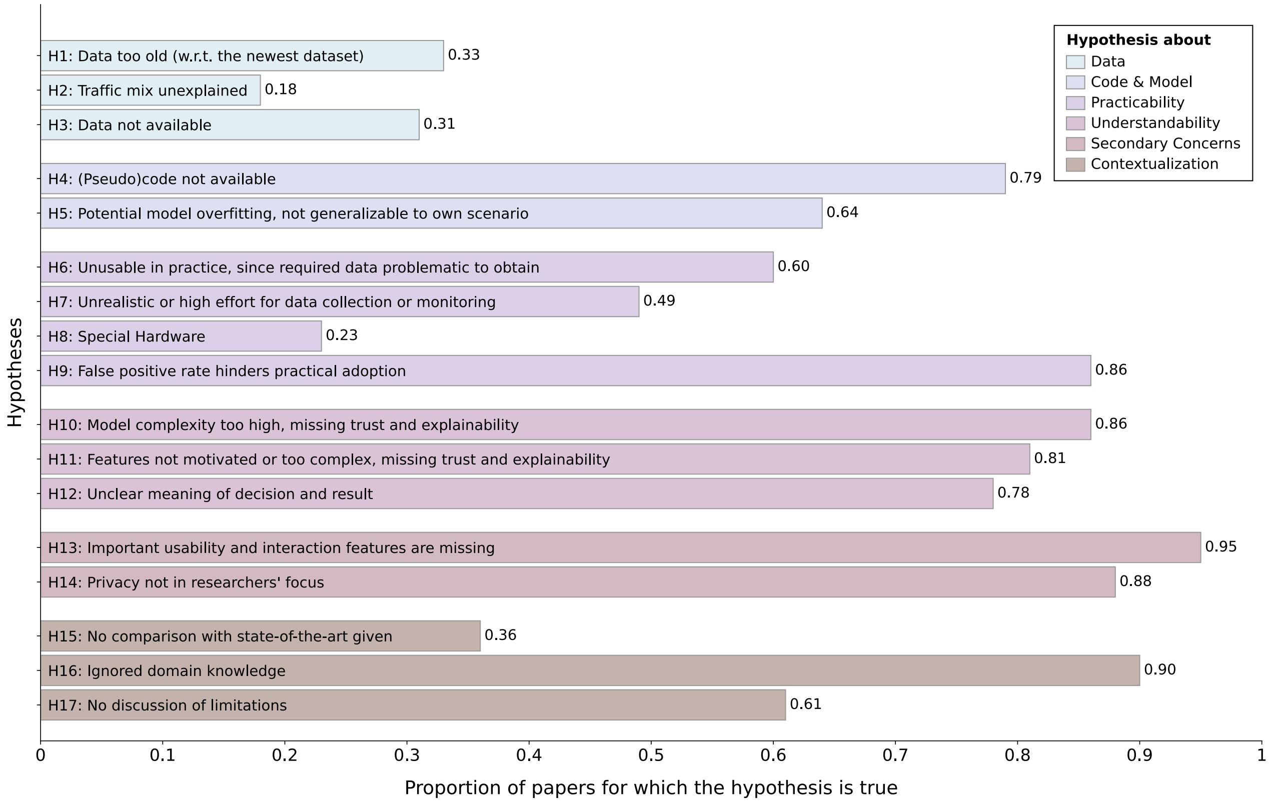

FIGURE 8. Results for all hypotheses and all papers in our literature review. Each hypothesis is also assigned to the topic we previously defined.

Table 4) in 86 % of the papers. That is, personas might not understand the model and thus do not trust the approach. In most papers, this is due to the use of black box models (75 % of papers), unclear model parameters (42 % of papers), or the use of various sequential or parallel models (39 % of papers).

Another factor for understanding the approach is feature selection. In H11, we hypothesize that even if the model is explainable, using complex features leads to missing trust and lack of explainability for the personas. Fig. 8 shows that 81 % of the papers use features that might be too complex for the personas for which this hypothesis is relevant (see Table 4). Those features are often too complex because they are not explicitly explained and no domain interpretation is given (52 % of papers). Further, features might not be comprehensible for the personas (42 % of the papers), also due to heavy preprocessing of the data (27 % of the papers). Additionally, many AI/ML-based NIDS approaches use too many features according to Miller's magical number of seven [28] (63 % of the papers) or do not provide a list of all features used in the approach (45 % of the papers).

In the end, the decision of the AI/ML model has to be explainable. As we found in H12, those decisions are often unclear (78 % of papers). That means, there is no sound explanation of the decision, for instance, researchers do not provide plots of decision trees, any feature importance values, or something comparable (63 % of papers). Further, results are often not explained in the context of the use case (50 % of papers), or the meaning of the results is not communicated to users (57 % of papers).

5) SECONDARY CONCERNS

There are also some secondary concerns introducing AI/ML-based NIDS approaches. Enterprises usually want to provide usable software for their employees. Further, approaches with invasive data collection might conflict with certain privacy regulations. In H13, we evaluated whether the papers address usability concerns. Our results in Fig. 8 show that 95 % of the papers do not integrate important usability features. In fact, 93 % of the papers do not provide any user interface, sophisticated graphical representation (also 93 % of papers), or define possible user interactions (91 % of papers). As we evaluated in H14, privacy is not in the researchers' focus in 88 % of the papers. While privacy is sometimes mentioned in the introduction, the approaches do not consider privacy-enhancing technologies. A small fraction of papers even used invasive methods for data collection (8 % of papers).

6) CONTEXTUALIZATION

There are also several other barriers that cannot be categorized easily. In H15, we evaluate whether the AI/ML-based NIDS approach is compared to other AI/ML-based NIDS approaches or traditional signature-based IDS. We found that there is no comparison with state-of-the-art approaches in

36 % of the papers. Further, regarding H16, we determine if the approaches use domain knowledge to improve the results. Our results in Fig. 8 show that 90 % of the papers do not use domain knowledge. Those papers neglect contextual information [31] (77 % of papers), do not gather information from other sources [23] (87 % of papers), or include no admin feedback [26] (81 % of papers). Furthermore, we believe that enterprises might be interested in limitations of the approaches (H17). However, 61 % of the papers do not describe any limitations of their approach.

V. DISCUSSION AND GUIDELINES

In the previous sections, we have established several hypotheses why AI/ML-based NIDS approaches are not adopted in practice. To answer our first and second research question (RQ1 and RQ2), we performed a literature review, identified relevant papers for our use case, and evaluated whether our hypotheses are true or false with a review template in a more objective manner. Our results have shown that there are multiple factors that prevent the practical adoption of AI/ML-based NIDS approaches. We grouped these limiting factors into six topics: data, code and model, practicability, understandability, secondary concerns, and contextualization.

Based on the above topics, we now establish research guidelines (G1 to G6) for the future in a more subjective and critical manner and discuss why these guidelines might not always be easy to follow in academia. In the following, we detail the six different guidelines, summarized in Table 6, to answer our third research question (RQ3).

A. G1: DATA

Utilizing realistic datasets is essential for obtaining accurate results, as well as making them public to foster reproducibility and, ultimately, practical adoption. While we found that a non-negligible portion of the papers did not utilize up-to-date data, some (few) even did not provide any insightful description of the data, and, most importantly, over a third did not provide publicly accessible data.

However, we acknowledge that obtaining real-world data is often a big challenge due to privacy/compliance concerns, e. g., from ISPs or companies. Even if such data is available, the utilized dataset may then not be shared with other researchers. Thus, utilizing the many readily available public datasets (created in labs or simulations) is an attractive alternative for many, i. e., it is understandable that many works use such, possibly outdated, data. Yet, as Engelen et al. [213] notice, even the newer, and (in our survey) most popular dataset CIC-IDS 2017, exhibits some undesirable properties.

Our recommendation is to at least make use of more recent datasets instead of using decade-old ones, take into consideration more than one dataset for generalizability, and carefully explain the choice of these datasets. If data from real-world scenarios is available but may not be published, researchers can verify their approach additionally on public datasets. This way, the proposed approach can be tested on realistic data, but is also transparent and reproducible to other researchers. The responsibility for establishing this guideline is not on the researchers alone, but also one the publishing venues. Incentives like the ACM Artifact Review and Badging[7] help to reinforce the integrity, reproducibility, and, ultimately, adoptability of academic research. However, Olszewski et al. [214] report no statistically significant difference between the availability of artifacts before and after the introduction of the badge for high-ranking venues. So, more incentives might be needed.

B. G2: CODE AND MODEL

Similarly to data availability, we found that the majority of papers did not provide their code and/or model. In the past years, it has become best practice to make the code available to the community, but incentives for researchers might be needed to establish that as a common practice. Challenges like copyright issues could hinder researchers from making their code available to the public. Additionally, some researchers may also be hesitant to publish their code due to their fear of erroneous code.

Our recommendation is that the model and the code should always be made public, since the community may help to refine and evolve the implementation. So, we advocate to use free-software licenses so that existing code may be altered, copied, and distributed under respective licenses, e. g., GPLv3. It helps to improve the practical applicability, as the approach does not need to be re-implemented based on potentially ambiguous implementation instructions and incomplete (hyper-)parameter settings. We argue that in case of errors it is especially useful to share the code, as the community can help fixing them. Both, commercial and open-source software, receive updates and patches on the regular, since making errors, especially in big projects, is almost unavoidable. Thus, publishing imperfect code should not be something to be afraid or ashamed of. Similar to the previous guideline, the responsibility here is also on the publishing venues, that need to encourage and enforce best practices. As mentioned, the approach can be easily tested on more datasets for the sake of generalizability if the code is public.

C. G3: PRACTICABILITY

Having potential users and/or system environments and their corresponding constraints in mind when designing an IDS is critical for its future deployment. We found that roughly half of the papers use data that is problematic to obtain and exhibits an unrealistic effort for monitoring (e. g., packet granularity, or in general not sampled/aggregated data). In addition, for the majority of the papers, the FPR hinders practical adoption, e. g., the FPR is either too high, or not discussed at all.

[7] https://acm.org/publications/policies/artifact-review-and-badging-current

TABLE 6. Proposed guidelines, including a summary of associated hypotheses, recommendations, challenges, and relevant personas. The guidelines directly correlate to the defined groups and our evaluation from before.

ID	Topic	Hypotheses	Recommendations	Challenges	Personas
G1	Data	H1, H2, H3	Publish & utilize realistic data, use multiple & newer datasets	Privacy/compliance concerns, realism of public datasets	All
G2	Code & Model	H4, H5	Publish code/model, use free-software licenses	Copyright issues, fear of errors	Compliance, Networker
G3	Practicability	H6, H7, H8, H9	Consider real-world constraints	Limited innovation/performance	All
G4	Understandability	H10, H11, H12	Keep it simple and explainable (features/input, model, results)	Limited innovation/performance	Compliance, Support, Planner, Networker
G5	Secondary Concerns	H13, H14	Discuss privacy/usability/etc.	Not in focus/feasible for all researchers	Manager, Compliance, Support, Networker
G6	Contextualization	H15, H16, H17	Expert knowledge & proper baselines	Not always possible/available, high effort	All

The challenge and potential trade-off is that restricting research only to *currently* available and feasible data in practice might hinder innovation. Sometimes, using more intricate and novel data can improve the overall performance of the model. That is, more complex features, that might not be feasible to monitor on enterprise hardware, might increase the AI/ML model's performance in terms of accuracy. Analogously, novel hardware, such as programmable switches, might increase its performance in terms of throughput.

We recommend at least discussing real-world applicability, such as monitoring effort (hardware, software, data volume etc.), or potential user/admin effort due to false alarms. For the FPR, we set an arbitrary cutoff of 1 %. Of course, the feasibility of this percentage is highly dependent on the absolute number of flows in the network; however, this is rarely discussed in any way in papers. So, we also recommend to take different types of performance or accuracy measures into account, and discuss trade-offs accordingly. Nevertheless, according to user studies by Kokulu et al. [215], the relevance of false alerts should not be overstressed, as not all practitioners in the field consider them a major issue.

D. G4: UNDERSTANDABILITY

To foster understanding of AI/ML-based IDS approaches, it is crucial to think about users of the system in advance. If the target audience, e. g., decision makers or administrators in enterprises, do not understand the approach, they will not trust the system and will not adopt it in practice. In this paper, we have shown that AI/ML models and features might not be explainable for our personas in the majority of papers.

Similar to the challenges and trade-offs of the previous guideline, using simpler and more explainable models, such as Decision Trees or Regression models, instead of black box models can lead to poorer performance and might limit innovation. While the previous guideline already mentioned that some features might not be practicable to monitor, this guideline is also concerned with the users' ability to comprehend them, both in terms of the sheer amount of features and their complexity. Again, using less complex but more explainable features might decrease its performance.

We recommend that researchers should always consider explainable AI/ML models in the first place. With XAI methods on the rise, more complex models, such as generalized additive models (GAM) [216] and explainable Boltzmann machines (EBM) [217], aim for both, explainability and performance. Certain visualizations, e. g., plots of the decision tree or insights from feature importances might help users gain trust in the system and understand the decision of a model. For black-box models, there is at least the possibility of post-hoc explainers like SHAP (based on Shapley values [218]) or LIME [219] to gain some insights, even if the model itself might not be explainable at first. In that regard, it is also important to rely on domain knowledge when using explainable features. Those features should be given, limited in number, and clearly communicated to users so that they are not overloaded with information. Finally, results of the AI/ML-based IDS approach should be reported in context of the use case, i. e., authors should describe what "accuracy" means for a certain use case. Our recommendations are further underlined by the fact that user studies already showed the positive reception of potential AI/ML tool users to explainable workflows [1].

Publishing venues have already been taking measures to foster more explainable research, by specifically including explainability into their areas of interest in calls for papers. Thus, we also found some papers during our review that were concerned directly with explainability in IDS approaches, such as the works by Wang et al. [176] and Minh et al. [55]. From 2020 to 2023, the number of articles available when searching for the term *explainable AI* on Google Scholar has almost tripled, from less than 8 000 to over 23 000 hits.

E. G5: SECONDARY CONCERNS

There are also several concerns regarding the practical adoption of AI/ML-based IDS approaches which are not directly related to the AI/ML approach. This includes for instance privacy and usability concerns. We have shown that most of the papers do not focus on usability and privacy aspects. However, if approaches are not compliant with data protection regulations such as the GDPR, enterprises cannot adopt those AI/ML-based IDS approaches in practice and potentially face severe punishments [220]. Additionally, if there is no user interface (or at least a description), the barrier for enterprises might be too high to adopt the system in practice.

The challenge with taking into account multiple secondary concerns is that researchers do not design market-ready software solutions. So, paying major attention to multiple technical and user-related aspects might not be in scope for most academic research and might simply be infeasible.

Our recommendation is to at least think about end users of the system for practical adoption. That is, researchers might use the persona concept to motivate use cases and show the practical applications of their approach. Therefore, we encourage researchers to discuss those concerns when designing AI/ML-based IDS approaches. We also want to emphasize that we found some works during our review where a graphical interface was implemented successfully in the context of AI/ML-based IDS which yielded important insights, e. g., the ILAB graphical user interface by Beaugnon et al. [38].

Regarding privacy-related concerns, there is already a shift in current research to not use payloads in feature engineering. Nevertheless, there are still other privacy concerns, e. g., about IP addresses, DNS data, or traffic patterns. So we also recommend that researchers should follow the principle of data minimization (Art. 5 par. I lit. c GDPR) and collect only absolutely necessary information.

F. G6: CONTEXTUALIZATION

Enterprises want to choose the best IDS solution for their needs in terms of detection performance, capabilities, and limitations. Our literature review shows, however, that this is often not possible. We found that a large number of the papers make no use of domain knowledge, i. e., additional use case-specific internal or external information. In our literature review, more than a third of the papers do not compare the solution with the related work at all and limitations are not always discussed in the papers.

The biggest challenge of this guideline is the high implementation effort that comes along with it, whether it is the synchronization of multiple data sources or the comparison with state-of-the-art approaches by properly reimplementing the related work. Researchers often compare their approaches to the related work in terms of accuracy or false positive rate. However, they often make comparisons based on different datasets or do not use current approaches as a baseline as many IDS solutions are not publicly available.

We recommend that AI/ML-based IDS approaches should use external knowledge, e. g, from open source databases, whenever possible to improve the approach. Candidates for such information that is already gathered by institutions and easily available include threat feeds and vulnerability data, but also (social) media and news [221]. We also highlight that researchers should discuss the capabilities and limitations of their approaches honestly and not cherry-pick the underlying data and baselines. If there are any possible evasions or specific requirements for the IDS solution, enterprises need to know that before running the software in production. Though, we acknowledge that testing the robustness of AI/ML-based IDS has been a topic in quite a few papers we found during our literature review [44], [63], [65], [76], [78], [176], e. g., via countermeasures like data poisoning or evasion attacks. In other words, synthetic data is created by generative methods such as Auto-Encoders (AEs) or Generative Adversarial Networks (GANs), and injected either during training or testing to fool the AI/ML model [44]. Though, as mentioned by Nkashama et al. [44], less attention has been paid to IDS. Even further, expertise for these types of adversarial attacks is considered rare according to user studies [1].

G. LIMITATIONS OF OUR WORK

Our proposed guidelines rely on the persona approach and the findings from our literature survey, which might have some limitations. One limitation might be that we relied on Proto-Personas instead of real Personas. Proto-Personas serve more as an ad-hoc solution for user-related research, instead of in-depth conversations with customers. However, as the personas were developed in close cooperation with domain experts of industry partners, we believe that our personas are sufficiently realistic.

Additionally, we could have derived more personas and hypotheses, as seen in the introduced example enterprises. Not all of the presented users of those enterprises are mapped to our personas. Nevertheless, we tried to break our personas down to the most important ones with feedback from our industry partners. In addition, our personas mainly apply to enterprise networks, as they were specifically designed for that context, and may not be generalizable to other types of networks, since potential tool users require expertise in security as well as AI/ML in general [1]. That is, other network types may have different user requirements, e. g., users in Smart Home environments may have less knowledge about AI/ML and communication networks in general, since they are likely no domain experts or administrators. For other networks, different personas might need to be designed. The same applies to the hypotheses.

During the paper review process, we noticed the gap between academia and practitioners of the industry partners ourselves. Due to the different backgrounds, we sometimes observed diverging opinions among the reviewers. In addition to the different backgrounds of reviewers, their research or working field also varied. Depending on the researcher's focus, this resulted in him/her judging some hypotheses more harshly than others. We tried to overcome these problems and potential biases with regular meetings and multiple iterations of our review template. As far as possible, we also agreed on objective criteria to make decisions on the hypotheses with the help of domain experts.

VI. CONCLUSION

In this literature survey, we asked the question, why academic AI/ML-based NIDS approaches have not been largely adopted in enterprise networks. First, to answer this question, we derived six personas from two different enterprise scenarios. Our personas, that are involved with AI/ML tools, were developed in cooperation with domain experts from our industry partners. Second, based on the personas, we derived 17 hypotheses concerned with six topics, regarding the used data, code and model, practicability, understandability, secondary concerns (usability and privacy), and contextualization.

We found 166 papers on AI/ML-based NIDS approaches from venues of AI, Network Management, Privacy and Security, as well as usability, HCI, and visualization. To answer our initial question, we formulated three fine-grained research questions. In the first one (RQ1), we evaluated technical factors that hinder practical adoption of AI/ML-based NIDS approaches. For that, we analyzed for all 166 papers whether our hypotheses are true. Our results show that the used datasets are not available in over a third of papers and almost 80 % of papers do not provide source code. Additionally, enterprises might have difficulties to obtain the necessary data for the AI/ML-based approaches in roughly half of the papers due to uncommon data formats, unavailable features, or large data volumes. Over 80 % of papers do not examine false alarms or similar metrics in detail. Further, almost 90 % of papers do not focus on privacy, i. e., privacy is rarely mentioned, and when it is, privacy-enhancing technologies are often not used.

In our second research question (RQ2), we analyzed what user-centric aspects should be considered to improve the adoption of AI/ML-based NIDS research. We evaluated RQ2 analogously to RQ1 with our hypotheses. Our analyses show that the AI/ML process is often not understandable. Explainability concerns regarding features, the AI/ML model, or decisions are not addressed in roughly 80 % of the papers. Usability-related issues, e. g., interfaces or sophisticated visualizations, are not a concern in over 90 % of papers.

In our third research question (RQ3), we derived guidelines from the previous two research questions. We conclude that researchers should consider additional concerns like practicability, explainability, privacy, and usability besides the mere performance of the model. Researchers should also be transparent by publishing their source code and data. There is also some responsibility on the publishing venues to foster reproducible research and make room for more niche topics in conjunction with network monitoring.

Our guidelines, that reflect academic and practical views, might help to address the above concerns in future academic research of AI/ML-based NIDS research. With regards to future research, one of the main gaps we discovered was the interplay between AI/ML-based solutions and secondary concerns. While there exist some papers that cover some of the mentioned disciplines in a more isolated manner, i. e., explainability, usability, or privacy for AI/ML-based IDS, there are still a lot of opportunities to expand on these topics from a researcher's point of view. However, not only do these individual topics need more attention in research, but also the combination of multiple disciplines, i. e., combining at least two or more of the aforementioned aspects. This allows to at least cover major concerns in order to make academic research for AI/ML-based IDS more applicable in practice.

ACKNOWLEDGMENT

The author Michael Seufert was with the University of Würzburg, Germany while this work was conducted.

The authors would like to thank Ankit Kariryaa for helping with the personas, Daria Soroko and Gian-Luca Savino for the input on usability-related topics, Alexej Grigorjew for the input on hypotheses and helping with reviewing, and Nicholas Gray for helping with searching through conference proceedings. The authors alone are responsible for the content.

REFERENCES

[1] J. Mink, H. Benkraouda, L. Yang, A. Ciptadi, A. Ahmadzadeh, D. Votipka, and G. Wang, "Everybody's got ML, tell me what else you have: Practitioners' perception of ML-based security tools and explanations," in *Proc. IEEE Symp. Secur. Privacy*, May 2023, pp. 2068–2085.

[2] M. Vielberth, F. Böhm, I. Fichtinger, and G. Pernul, "Security operations center: A systematic study and open challenges," *IEEE Access*, vol. 8, pp. 227756–227779, 2020.

[3] R. Boutaba, M. A. Salahuddin, N. Limam, S. Ayoubi, N. Shahriar, F. Estrada-Solano, and O. M. Caicedo, "A comprehensive survey on machine learning for networking: Evolution, applications and research opportunities," *J. Internet Serv. Appl.*, vol. 9, no. 1, pp. 1–99, 2018.

[4] T. T. Nguyen and G. Armitage, "A survey of techniques for internet traffic classification using machine learning," *IEEE Commun. Surveys Tuts.*, vol. 10, no. 4, pp. 56–76, 4th Quart., 2008.

[5] V. T. Guimarães, C. M. D. S. Freitas, R. Sadre, L. M. R. Tarouco, and L. Z. Granville, "A survey on information visualization for network and service management," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 285–323, 1st Quart., 2016.

[6] T. F. Lunt, "A survey of intrusion detection techniques," *Comput. Secur.*, vol. 12, no. 4, pp. 405–418, 1993.

[7] S. Axelsson, "Intrusion detection systems: A survey and taxonomy," Chalmers Univ. Technol., Gothenburg, Sweden, Tech. Rep. 99-15, 2000.

[8] H.-J. Liao, C.-H. R. Lin, Y.-C. Lin, and K.-Y. Tung, "Intrusion detection system: A comprehensive review," *J. Netw. Comput. Appl.*, vol. 36, no. 1, pp. 16–24, 2013.

[9] Z. Ahmad, A. Shahid Khan, C. Wai Shiang, J. Abdullah, and F. Ahmad, "Network intrusion detection system: A systematic study of machine learning and deep learning approaches," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 1, 2021, Art. no. e4150.

[10] E. Hodo, X. Bellekens, A. Hamilton, C. Tachtatzis, and R. Atkinson, "Shallow and deep networks intrusion detection system: A taxonomy and survey," 2017, *arXiv:1701.02145*.

[11] N. Sultana, N. Chilamkurti, W. Peng, and R. Alhadad, "Survey on SDN based network intrusion detection system using machine learning approaches," *P2P Netw. Appl.*, vol. 12, pp. 493–501, Mar. 2019.

[12] N. Moustafa, N. Koroniotis, M. Keshk, A. Y. Zomaya, and Z. Tari, "Explainable intrusion detection for cyber defences in the Internet of Things: Opportunities and solutions," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 3, pp. 1775–1807, 3rd Quart., 2023.

[13] N. Capuano, G. Fenza, V. Loia, and C. Stanzione, "Explainable artificial intelligence in CyberSecurity: A survey," *IEEE Access*, vol. 10, pp. 93575–93600, 2022.

[14] S. Neupane, J. Ables, W. Anderson, S. Mittal, S. Rahimi, I. Banicescu, and M. Seale, "Explainable intrusion detection systems (X-IDS): A survey of current methods, challenges, and opportunities," *IEEE Access*, vol. 10, pp. 112392–112415, 2022.

[15] F. Sperrle, M. El-Assady, G. Guo, R. Borgo, D. H. Chau, A. Endert, and D. Keim, "A survey of human-centered evaluations in human-centered machine learning," *Comput. Graph. Forum*, vol. 40, no. 3, pp. 543–568, 2021.

[16] T. Kaluarachchi, A. Reis, and S. Nanayakkara, "A review of recent deep learning approaches in human-centered machine learning," *Sensors*, vol. 21, no. 7, p. 2514, 2021.

[17] A. Boulemtafes, A. Derhab, and Y. Challal, "A review of privacy-preserving techniques for deep learning," *Neurocomputing*, vol. 384, pp. 21–45, Apr. 2020.

[18] J. Gothelf, "Using proto-personas for executive alignment," *UX Mag.*, vol. 1, pp. 26–29, Mar. 2012.

[19] F. L. Verdi, H. T. d. Oliveira, L. N. Sampaio, and L. A. M. Zaina, "Usability matters: A human–computer interaction study on network management tools," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 3, pp. 1865–1878, Sep. 2020.

[20] S. Mckenna, D. Staheli, and M. Meyer, "Unlocking user-centered design methods for building cyber security visualizations," in *Proc. IEEE Symp. Visualizat. Cyber Secur. (VizSec)*, Oct. 2015, pp. 1–8.

[21] M. Schufrin, K. Kuban, A. Kuijper, and J. Kohlhammer, "NetVisGame: Mobile gamified information visualization of home network traffic data," in *Proc. Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, 2022, pp. 129–138.

[22] A. Kenyon, L. Deka, and D. Elizondo, "Are public intrusion datasets fit for purpose: Characterising the state of the art in intrusion event datasets," *Comput. Secur.*, vol. 99, Dec. 2020, Art. no. 102022.

[23] O. Akinrolabu, I. Agrafiotis, and A. Erola, "The challenge of detecting sophisticated attacks: Insights from SOC analysts," in *Proc. 13th Int. Conf. Avail., Rel. Secur.*, 2018, pp. 1–9.

[24] M. Fejrskov, J. M. Pedersen, and E. Vasilomanolakis, "Cyber-security research by ISPs: A NetFlow and DNS anonymization policy," in *Proc. Int. Conf. Cyber Secur. Protection Digit. Services (Cyber Secur.)*, Jun. 2020, pp. 1–8.

[25] R. Hofstede, P. Celeda, B. Trammell, I. Drago, R. Sadre, A. Sperotto, and A. Pras, "Flow monitoring explained: From packet capture to data analysis with NetFlow and IPFIX," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2037–2064, 4th Quart., 2014.

[26] A. Sopan, M. Berninger, M. Mulakaluri, and R. Katakam, "Building a machine learning model for the SOC, by the input from the SOC, and analyzing it for the SOC," in *Proc. IEEE Symp. Vis. Cyber Secur.*, Oct. 2018, pp. 1–8.

[27] S. Oesch, R. Bridges, J. Smith, J. Beaver, J. Goodall, K. Huffer, C. Miles, and D. Scofield, "An assessment of the usability of machine learning based tools for the security operations center," in *Proc. Int. Conf. Internet Things (iThings) IEEE Green Comput. Commun. (GreenCom) IEEE Cyber, Phys. Social Comput. (CPSCom) IEEE Smart Data (SmartData) IEEE Congr. Cybermatics (Cybermatics)*, Nov. 2020, pp. 634–641.

[28] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *Psychol. Rev.*, vol. 63, no. 2, p. 81, 1956.

[29] J. Stegman, P. J. Trottier, C. Hillier, H. Khan, and M. Mannan, "'My privacy for their security': Employees' privacy perspectives and expectations when using enterprise security software," in *Proc. 32nd USENIX Secur. Symp.*, 2023, pp. 3583–3600.

[30] H. Neuschmied, M. Winter, B. Stojanović, K. Hofer-Schmitz, J. Božić, and U. Kleb, "APT-attack detection based on multi-stage autoencoders," *Appl. Sci.*, vol. 12, no. 13, p. 6816, 2022.

[31] J. R. Goodall, W. G. Lutters, and A. Komlodi, "The work of intrusion detection: Rethinking the role of security analysts," in *Proc. 10th Amer. Conf. Inf. Syst.*, 2004, p. 179.

[32] Z. Fan, Z. Zhou, J. Pei, M. P. Friedlander, J. Hu, C. Li, and Y. Zhang, "Knowledge-injected federated learning," 2022, *arXiv:2208.07530*.

[33] J. Yan, L. Shi, J. Tao, X. Yu, Z. Zhuang, C. Huang, R. Yu, P. Su, C. Wang, and Y. Chen, "Visual analysis of collective anomalies using faceted high-order correlation graphs," *IEEE Trans. Vis. Comput. Graphics*, vol. 26, no. 7, pp. 2517–2534, Jul. 2020.

[34] F. Böhm, L. Englbrecht, S. Friedl, and G. Pernul, "Visual decision-support for live digital forensics," in *Proc. IEEE Symp. Visualizat. Cyber Secur. (VizSec)*, Oct. 2021, pp. 58–67.

[35] J. Raynor, T. Crnovrsanin, S. Di Bartolomeo, L. South, D. Saffo, and C. Dunne, "The state of the art in BGP visualization tools: A mapping of visualization techniques to cyberattack types," *IEEE Trans. Vis. Comput. Graphics*, vol. 29, no. 1, pp. 1059–1069, Jan. 2023.

[36] Y. Shi, Y. Zhao, F. Zhou, R. Shi, Y. Zhang, and G. Wang, "A novel radial visualization of intrusion detection alerts," *IEEE Comput. Graph. Appl.*, vol. 38, no. 6, pp. 83–95, Nov. 2018.

[37] M. Angelini, G. Blasilli, T. Catarci, S. Lenti, and G. Santucci, "Vulnus: Visual vulnerability analysis for network security," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 1, pp. 183–192, Jan. 2019.

[38] A. Beaugnon, P. Chifflier, and F. R. Bach, "End-to-end active learning for computer security experts," in *Proc. Workshops 32nd AAAI Conf Artif. Intell.*, 2018, pp. 217–224.

[39] X. Han, X. Chen, and L. Liu, "GAN ensemble for anomaly detection," in *Proc. 35th AAAI Conf. Artif. Intell., 33rd Conf. Innov. Appl. Artif. Intell., 11th Symp. Edu. Adv. Artif. Intell.*, 2021, pp. 4090–4097.

[40] M. A. Siddiqui, A. Fern, R. Wright, A. Theriault, D. W. Archer, and W. Maxwell, "Detecting cyberattack entities from audit data via multi-view anomaly detection with feedback," in *Proc. Workshops 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 277–284.

[41] A. R. Tuor, R. Baerwolf, N. Knowles, B. Hutchinson, N. Nichols, and R. Jasper, "Recurrent neural network language models for open vocabulary event-level cyber anomaly detection," in *Proc. Workshops 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 285–293.

[42] A. Sankararaman, B. Narayanaswamy, V. Y. Singh, and Z. Song, "FITNESS: (Fine tune on new and similar samples) to detect anomalies in streams with drift and outliers," in *Proc. Int. Conf. Mach. Learn.*, vol. 162, 2022, pp. 19153–19177.

[43] C. S. de Witt, Y. Huang, P. H. S. Torr, and M. Strohmeier, "Fixed points in cyber space: Rethinking optimal evasion attacks in the age of AI-NIDS," 2021, *arXiv:2111.12197*.

[44] D. K. Nkashama, A. Soltani, J. Verdier, M. Frappier, P. Tardif, and F. Kabanza, "Robustness evaluation of deep unsupervised learning algorithms for intrusion detection systems," 2022, *arXiv:2207.03576*.

[45] H. Xu, Y. Wang, J. Wei, S. Jian, Y. Li, and N. Liu, "Fascinating supervisory signals and where to find them: Deep anomaly detection with scale learning," in *Proc. Int. Conf. Mach. Learn.*, vol. 202, 2023, pp. 38655–38673.

[46] P. Mulinka and P. Casas, "Stream-based machine learning for network security and anomaly detection," in *Proc. Workshop Big Data Analyt. Mach. Learn. Data Commun. Netw.*, 2018, pp. 1–7.

[47] A. Putina, D. Rossi, A. Bifet, S. Barth, D. Pletcher, C. Precup, and P. Nivaggioli, "Telemetry-based stream-learning of BGP anomalies," in *Proc. Workshop Big Data Analyt. Mach. Learn. Data Commun. Netw.*, 2018, pp. 15–20.

[48] M. Wichtlhuber, E. Strehle, D. Kopp, L. Prepens, S. Stegmueller, A. Rubina, C. Dietzel, and O. Hohlfeld, "IXP scrubber: Learning from blackholing traffic for ML-driven DDoS detection at scale," in *Proc. Annu. Conf. ACM Special Int. Group Data Commun. (SIGCOMM)*, 2022, pp. 707–722.

[49] H. Bian, T. Bai, M. A. Salahuddin, N. Limam, A. A. Daya, and R. Boutaba, "Host in danger? Detecting network intrusions from authentication logs," in *Proc. 15th Int. Conf. Netw. Service Manage. (CNSM)*, Oct. 2019, pp. 1–9.

[50] R. Copstein and N. Zincir-Heywood, "Temporal representations for detecting BGP blackjack attacks," in *Proc. 16th Int. Conf. Netw. Service Manage. (CNSM)*, Nov. 2020, pp. 1–7.

[51] R. Copstein, B. Niblett, A. Johnston, J. Schwartzentruber, M. I. Heywood, and N. Zincir-Heywood, "MIMC: Anomaly detection in network data via multiple instances of micro-cluster detection," in *Proc. 19th Int. Conf. Netw. Serv. Manage.*, 2023, pp. 1–7.

[52] P. Golchin, C. Zhou, P. Agnihotri, M. Hajizadeh, R. Kundel, and R. Steinmetz, "CML-IDS: Enhancing intrusion detection in SDN through collaborative machine learning," in *Proc. 19th Int. Conf. Netw. Serv. Manage.*, 2023, pp. 1–9.

[53] C. Hardegen, "Scope-based flow monitoring to improve traffic analysis in programmable networks," in *Proc. 18th Int. Conf. Netw. Service Manage. (CNSM)*, Oct. 2022, pp. 254–260.

[54] J. Koumar, K. Hynek, and T. Cejka, "Network traffic classification based on single flow time series analysis," in *Proc. 19th Int. Conf. Netw. Serv. Manage.*, 2023, pp. 1–7.

[55] C. Minh, K. Vermeulen, C. Lefebvre, P. Owezarski, and W. Ritchie, "An explainable-by-design ensemble learning system to detect unknown network attacks," in *Proc. 19th Int. Conf. Netw. Serv. Manage.*, 2023, pp. 1–9.

[56] V. Minkevics and J. Kampars, "Artificial intelligence and big data driven IS security management solution with applications in higher education organizations," in *Proc. 17th Int. Conf. Netw. Service Manage. (CNSM)*, Oct. 2021, pp. 340–344.

[57] M. A. Salahuddin, Md. Faizul Bari, H. A. Alameddine, V. Pourahmadi, and R. Boutaba, "Time-based anomaly detection using autoencoder," in *Proc. 16th Int. Conf. Netw. Service Manage. (CNSM)*, Nov. 2020, pp. 1–9.

[58] S. Wassermann, T. Cuvelier, P. Mulinka, and P. Casas, "ADAM & RAL: Adaptive memory learning and reinforcement active learning for network monitoring," in *Proc. 15th Int. Conf. Netw. Service Manage. (CNSM)*, Oct. 2019, pp. 1–9.

[59] F. Musumeci, V. Ionata, F. Paolucci, F. Cugini, and M. Tornatore, "Machine-learning-assisted DDoS attack detection with P4 language," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[60] O. Aouedi, K. Piamrat, G. Müller, and K. D. Singh, "Intrusion detection for softwarized networks with semi-supervised federated learning," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 5244–5249.

[61] H. Alanazi, S. Bi, T. Wang, and T. Hou, "Exquisite feature selection for machine learning powered probing attack detection," in *Proc. IEEE Int. Conf. Commun.*, May 2023, pp. 783–789.

[62] S. Araki, K. Takahashi, B. Hu, K. Kamiya, and M. Tanikawa, "Subspace clustering for interpretable botnet traffic analysis," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[63] G. Bovenzi, A. Foggia, S. Santella, A. Testa, V. Persico, and A. Pescapé, "Data poisoning attacks against autoencoder-based anomaly detection models: A robustness analysis," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 5427–5432.

[64] Y. Chen, J. Pei, and D. Li, "DETPro: A high-efficiency and low-latency system against DDoS attacks in SDN based on decision tree," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[65] F. Cui, Q. Ye, and P. L. Kibenge, "A Wasserstein GAN-based framework for adversarial attacks against intrusion detection systems," in *Proc. IEEE Int. Conf. Commun.*, May 2023, pp. 3187–3192.

[66] J. Geng, S. Li, Y. Zhang, Z. Liu, and Z. Cheng, "LIFH: Learning interactive features from HTTP payload using image reconstruction," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2021, pp. 1–6.

[67] N. H. Hamilton, S. McKinney, E. Allan, and E. W. Fulp, "An efficient multi-stage approach for identifying domain shadowing," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–7.

[68] X. Han, P. Dong, S. Liu, B. Jiang, Z. Lu, and Z. Cui, "IV-IDM: Reliable intrusion detection method based on involution and voting," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 4162–4167.

[69] Z. A. El Houda and L. Khoukhi, "A hierarchical fog computing framework for network attack detection in SDN," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 4366–4371.

[70] M. M. Isa and L. Mhamdi, "Hybrid deep autoencoder with random forest in native SDN intrusion detection environment," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 1698–1703.

[71] A. E. Kamali, K. Chougdali, and A. Kobbane, "A new intrusion detection system based on convolutional neural network," in *Proc. IEEE Int. Conf. Commun.*, May 2023, pp. 2994–2999.

[72] R. A. Khamis, M. O. Shafiq, and A. Matrawy, "Investigating resistance of deep learning-based IDS against adversaries using min-max optimization," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–7.

[73] H. Kye, M. Kim, and M. Kwon, "Hierarchical autoencoder for network intrusion detection," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 2700–2705.

[74] Z. Li, J. Wu, S. Mumtaz, A. M. Taha, S. Al-Rubaye, and A. Tsourdos, "Machine learning and multi-dimension features based adaptive intrusion detection in ICN," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–5.

[75] H. Li, Z. Chen, R. Spolaor, Q. Yan, C. Zhao, and B. Yang, "DART: Detecting unseen malware variants using adaptation regularization transfer learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[76] P. Li, Q. Liu, W. Zhao, D. Wang, and S. Wang, "Chronic poisoning against machine learning based IDSs using edge pattern detection," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–7.

[77] Z. Li, Y. Sang, Z. Cheng, T. Zang, S. Zhao, and H. Wang, "TCCN: A network traffic classification and detection model based on capsule network," in *Proc. IEEE Int. Conf. Commun.*, May 2023, pp. 2319–2324.

[78] S. Li, W. Wu, Y. Meng, J. Li, H. Zhu, and X. S. Shen, "Data poisoning attack against anomaly detectors in digital twin-based networks," in *Proc. IEEE Int. Conf. Commun.*, May 2023, pp. 13–18.

[79] A. G. P. Lobato, M. A. Lopez, I. J. Sanz, A. A. Cardenas, O. C. M. B. Duarte, and G. Pujolle, "An adaptive real-time architecture for zero-day threat detection," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.

[80] A. Mudgerikar, E. Bertino, J. Lobo, and D. Verma, "A security-constrained reinforcement learning framework for software defined networks," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2021, pp. 1–6.

[81] L. G. Nguyen and K. Watabe, "A method for network intrusion detection using flow sequence and BERT framework," in *Proc. IEEE Int. Conf. Commun.*, May 2023, pp. 3006–3011.

[82] O. R. Sanchez, M. Repetto, A. Carrega, R. Bolla, and J. F. Pajo, "Feature selection evaluation towards a lightweight deep learning DDoS detector," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2021, pp. 1–6.

[83] Y. Shen, C. Wu, D. Kong, and M. Yang, "TPDD: A two-phase DDoS detection system in software-defined networking," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[84] Y. Sun, N. S. T. Chong, and H. Ochiai, "Network flows-based malware detection using a combined approach of crawling and deep learning," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2021, pp. 1–6.

[85] R. L. Tomio, E. K. Viegas, A. O. Santin, and R. R. dos Santos, "A multi-view intrusion detection model for reliable and autonomous model updates," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2021, pp. 1–6.

[86] E. K. Viegas, A. O. Santin, V. V. Cogo, and V. Abreu, "A reliable semi-supervised intrusion detection model: One year of network traffic anomalies," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[87] L. Vu, V. L. Cao, Q. U. Nguyen, D. N. Nguyen, D. T. Hoang, and E. Dutkiewicz, "Learning latent distribution for distinguishing network traffic in intrusion detection system," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[88] X. Wang, K. J. Kim, Y. Wang, T. Koike-Akino, and K. Parsons, "DeepEAD: Explainable anomaly detection from system logs," in *Proc. IEEE Int. Conf. Commun.*, May 2023, pp. 771–776.

[89] D. Wu, B. Fang, J. Wang, Q. Liu, and X. Cui, "Evading machine learning botnet detection models via deep reinforcement learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[90] S. Xu, X. Han, T. Tian, B. Jiang, Z. Lu, and C. Zhang, "Few-shot network traffic anomaly detection based on Siamese neural network," in *Proc. IEEE Int. Conf. Commun.*, May 2023, pp. 3012–3017.

[91] M. Xu, X. Li, J.-f. Ma, C. Zhong, and W. Yang, "Detection of multi-stage attacks based on multi-layer long and short-term memory network," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[92] S. Xu, Y. Qian, and R. Q. Hu, "A semi-supervised learning approach for network anomaly detection in fog computing," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[93] R. Yaegashi, E. Takeshita, and Y. Nakayama, "Two-stage DDoS mitigation with variational auto-encoder and cyclic queuing," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 5421–5426.

[94] Y. Yan, L. Qi, J. Wang, Y. Lin, and L. Chen, "A network intrusion detection method based on stacked autoencoder and LSTM," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[95] S. Yang, X. Zheng, Z. Xu, and X. Wang, "A lightweight approach for network intrusion detection based on self-knowledge distillation," in *Proc. IEEE Int. Conf. Commun.*, May 2023, pp. 3000–3005.

[96] J. Zhang, F. Li, and F. Ye, "An ensemble-based network intrusion detection scheme with Bayesian deep learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[97] R. Zhao, Y. Chen, Y. Wang, Y. Shi, and Z. Xue, "An efficient and lightweight approach for intrusion detection based on knowledge distillation," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2021, pp. 1–6.

[98] R. Zhao, T. Tang, G. Gui, and Z. Xue, "A lightweight semi-supervised learning method based on consistency regularization for intrusion detection," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 3124–3129.

[99] B. Zhou, J. Li, J. Wu, S. Guo, Y. Gu, and Z. Li, "Machine-learning-based online distributed denial-of-service attack detection using spark streaming," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.

[100] J. Bakker, B. Ng, W. K. G. Seah, and A. Pekar, "Traffic classification with machine learning in a live network," in *Proc. IFIP/IEEE Symp. Integr. Netw. Service Manage. (IM)*, Apr. 2019, pp. 488–493.

[101] A. A. Daya, M. A. Salahuddin, N. Limam, and R. Boutaba, "A graph-based machine learning approach for bot detection," in *Proc. IFIP/IEEE Symp. Integr. Netw. Service Manage. (IM)*, Apr. 2019, pp. 144–152.

[102] E. Jalalpour, M. Ghaznavi, R. Boutaba, and T. Ahmed, "TMAS: A traffic monitoring analytics system leveraging machine learning," in *Proc. IFIP/IEEE Symp. Integr. Netw. Service Manage. (IM)*, Apr. 2019, pp. 408–414.

[103] S. G. Macías, L. P. Gaspary, and J. F. Botero, "ORACLE: An architecture for collaboration of data and control planes to detect DDoS attacks," in *Proc. IFIP/IEEE Int. Symp. Integr. Netw. Manage. (IM)*, May 2021, pp. 962–967.

[104] J. S. M. Osorio, J. A. V. Tejada, and J. F. B. Vega, "Detection of DoS/DDoS attacks: The UBM and GMM approach," in *Proc. IFIP/IEEE Int. Symp. Integr. Netw. Manage. (IM)*, May 2021, pp. 866–871.

[105] I. Akbari, E. Tahoun, M. A. Salahuddin, N. Limam, and R. Boutaba, "ATMoS: Autonomous threat mitigation in SDN using reinforcement learning," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2020, pp. 1–9.

[106] P. Arregoces, J. Vergara, S. A. Gutiérrez, and J. F. Botero, "Network-based intrusion detection: A one-class classification approach," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2022, pp. 1–6.

[107] G. Baye, P. Silva, A. Broggi, L. Fiondella, N. D. Bastian, and G. Kul, "Performance analysis of deep-learning based open set recognition algorithms for network intrusion detection systems," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, May 2023, pp. 1–6.

[108] A. Blaise, M. Bouet, V. Conan, and S. Secci, "BotFP: FingerPrints clustering for bot detection," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2020, pp. 1–7.

[109] H.-F. Chang, M. I.-C. Wang, C.-H. Hung, and C. H. Wen, "Enabling malware detection with machine learning on programmable switch," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2022, pp. 1–5.

[110] S. Dvorak, P. Prochazka, and L. Bajer, "GNN-based malicious network entities identification in large-scale network data," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2022, pp. 1–4.

[111] P. Golchin, R. Kundel, T. Steuer, R. Hark, and R. Steinmetz, "Improving DDoS attack detection leveraging a multi-aspect ensemble feature selection," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2022, pp. 1–5.

[112] M. Hajizadeh, S. Barua, and P. Golchin, "FSA-IDS: A flow-based self-active intrusion detection system," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, May 2023, pp. 1–9.

[113] K. Hara and K. Shiomoto, "Intrusion detection system using semi-supervised learning with adversarial auto-encoder," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2020, pp. 1–8.

[114] C. Lübben and M.-O. Pahl, "Distributed device-specific anomaly detection using deep feed-forward neural networks," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, May 2023, pp. 1–9.

[115] R. Paudel and H. H. Huang, "Pikachu: Temporal walk based dynamic graph embedding for network anomaly detection," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2022, pp. 1–7.

[116] K. Yang, J. Zhang, Y. Xu, and J. Chao, "DDoS attacks detection with AutoEncoder," in *Proc. NOMS IEEE/IFIP Netw. Operations Manage. Symp.*, Apr. 2020, pp. 1–9.

[117] M. Žádník, "Towards inference of DDoS mitigation rules," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2022, pp. 1–5.

[118] M. Zaman and C.-H. Lung, "Evaluation of machine learning techniques for network intrusion detection," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, Apr. 2018, pp. 1–5.

[119] K. Zhang, N. Samaan, and A. Karmouch, "An intelligent data-plane with a quantized ML model for traffic management," in *Proc. NOMS IEEE/IFIP Netw. Oper. Manage. Symp.*, May 2023, pp. 1–9.

[120] A. Alsirhani, S. Sampalli, and P. Bodorik, "DDoS detection system: Using a set of classification algorithms controlled by fuzzy logic system in apache spark," *IEEE Trans. Netw. Service Manage.*, vol. 16, no. 3, pp. 936–949, Sep. 2019.

[121] P. F. de Araujo-Filho, M. Naili, G. Kaddoum, E. T. Fapi, and Z. Zhu, "Unsupervised GAN-based intrusion detection system using temporal convolutional networks and self-attention," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 4, pp. 4951–4963, Aug. 2023.

[122] G. Apruzzese, M. Andreolini, M. Marchetti, A. Venturi, and M. Colajanni, "Deep reinforcement adversarial learning against botnet evasion attacks," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 4, pp. 1975–1987, Dec. 2020.

[123] O. Barut, Y. Luo, P. Li, and T. Zhang, "R1DIT: Privacy-preserving malware traffic classification with attention-based neural networks," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 2, pp. 2071–2085, Oct. 2023.

[124] R. Doriguzzi-Corin, S. Millar, S. Scott-Hayward, J. Martínez-del-Rincón, and D. Siracusa, "Lucid: A practical, lightweight deep learning solution for DDoS attack detection," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 2, pp. 876–889, Jun. 2020.

[125] S. Das, S. Saha, A. T. Priyoti, E. K. Roy, F. T. Sheldon, A. Haque, and S. Shiva, "Network intrusion detection and comparative analysis using ensemble machine learning and feature selection," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 4, pp. 4821–4833, Dec. 2022.

[126] A. A. Daya, M. A. Salahuddin, N. Limam, and R. Boutaba, "BotChase: Graph-based bot detection using machine learning," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 1, pp. 15–29, Mar. 2020.

[127] P. Kr. Deka, Y. Verma, A. B. Bhutto, E. Elmroth, and M. Bhuyan, "Semi-supervised range-based anomaly detection for cloud systems," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 2, pp. 1290–1304, Sep. 2023.

[128] D. IR and S. K, "DAD: Domain adversarial defense system against DDoS attacks in cloud," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 1, pp. 554–568, Mar. 2022.

[129] S. Dong, Y. Xia, and T. Peng, "Network abnormal traffic detection model based on semi-supervised deep reinforcement learning," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 4, pp. 4197–4212, Dec. 2021.

[130] S. Garg, K. Kaur, N. Kumar, G. Kaddoum, A. Y. Zomaya, and R. Ranjan, "A hybrid deep learning-based model for anomaly detection in cloud datacenter networks," *IEEE Trans. Netw. Service Manage.*, vol. 16, no. 3, pp. 924–935, Sep. 2019.

[131] M. J. Hashemi, E. Keller, and S. Tizpaz-Niari, "Detecting unseen anomalies in network systems by leveraging neural networks," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 3, pp. 2515–2528, Jun. 2023.

[132] Z. A. El Houda, L. Khoukhi, and A. S. Hafid, "Bringing intelligence to software defined networks: Mitigating DDoS attacks," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 4, pp. 2523–2535, Dec. 2020.

[133] M. Injadat, A. Moubayed, A. B. Nassif, and A. Shami, "Multi-stage optimized machine learning framework for network intrusion detection," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 2, pp. 1803–1816, Jun. 2021.

[134] P. Krishnamurthy, F. Khorrami, S. Schmidt, and K. Wright, "Machine learning for NetFlow anomaly detection with human-readable annotations," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 2, pp. 1885–1898, Jun. 2021.

[135] D. C. Le, N. Zincir-Heywood, and M. I. Heywood, "Analyzing data granularity levels for insider threat detection using machine learning," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 1, pp. 30–44, Mar. 2020.

[136] W. T. Lunardi, M. A. Lopez, and J.-P. Giacalone, "ARCADE: Adversarially regularized convolutional autoencoder for network anomaly detection," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 2, pp. 1305–1318, Aug. 2023.

[137] M. Lyu, H. H. Gharakheili, C. Russell, and V. Sivaraman, "Hierarchical anomaly-based detection of distributed DNS attacks on enterprise networks," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 1, pp. 1031–1048, Mar. 2021.

[138] D. Madariaga, J. Madariaga, M. Panza, J. Bustos-Jiménez, and B. Bustos, "Detecting anomalies at a TLD name server based on DNS traffic predictions," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 1, pp. 1016–1030, Mar. 2021.

[139] M. D. Mauro, G. Galatro, and A. Liotta, "Experimental review of neural-based approaches for network intrusion management," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 4, pp. 2480–2495, Dec. 2020.

[140] R. Mills, A. K. Marnerides, M. Broadbent, and N. Race, "Practical intrusion detection of emerging threats," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 1, pp. 582–600, Mar. 2022.

[141] G. Mohi-ud-din, L. Zhiqiang, Z. Jiangbin, W. Sifei, L. Zhijun, M. Asim, Y. Zhong, and Y. Chen, "Intrusion detection using hybrid enhanced CSA-PSO and multivariate WLS random-forest technique," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 4, pp. 4937–4950, May 2023.

[142] T. V. Phan, T. G. Nguyen, N.-N. Dao, T. T. Huong, N. H. Thanh, and T. Bauschert, "DeepGuard: Efficient anomaly detection in SDN with fine-grained traffic flow monitoring," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 3, pp. 1349–1362, Sep. 2020.

[143] T. V. Phan and T. Bauschert, "DeepAir: Deep reinforcement learning for adaptive intrusion response in software-defined networks," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 3, pp. 2207–2218, Sep. 2022.

[144] A. Putina and D. Rossi, "Online anomaly detection leveraging stream-based clustering and real-time telemetry," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 1, pp. 839–854, Mar. 2021.

[145] M. A. Salahuddin, V. Pourahmadi, H. A. Alameddine, M. F. Bari, and R. Boutaba, "Chronos: DDoS attack detection using time-based autoencoder," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 1, pp. 627–641, Mar. 2022.

[146] B. H. Schwengber, A. Vergütz, N. G. Prates, and M. Nogueira, "Learning from network data changes for unsupervised botnet detection," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 1, pp. 601–613, Mar. 2022.

[147] Z. Shao, T. Chen, G. Cheng, X. Hu, W. Li, and H. Wu, "AF-FDS: An accurate, fast, and fine-grained detection scheme for DDoS attacks in high-speed networks with asymmetric routing," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 4, pp. 4964–4981, Apr. 2023.

[148] K. A. Simpson, S. Rogers, and D. P. Pezaros, "Per-host DDoS mitigation by direct-control reinforcement learning," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 1, pp. 103–117, Mar. 2020.

[149] D. Tang, C. Gao, W. Liang, J. Zhang, and K. Li, "FTMaster: A detection and mitigation system of low-rate flow table overflow attacks via SDN," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 4, pp. 5073–5084, Apr. 2023.

[150] S. Tan, X. Zhong, Z. Tian, and Q. Dong, "Sneaking through security: Mutating live network traffic to evade learning-based NIDS," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 3, pp. 2295–2308, Sep. 2022.

[151] M. Verkerken, L. D'hooge, D. Sudyana, Y.-D. Lin, T. Wauters, B. Volckaert, and F. D. Turck, "A novel multi-stage approach for hierarchical intrusion detection," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 3, pp. 3915–3929, Aug. 2023.

[152] X.-S. Vu, M. Ma, and M. Bhuyan, "MetaVSID: A robust meta-reinforced learning approach for VSI-DDoS detection on the edge," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 2, pp. 1625–1643, May 2023.

[153] Z. Wu, P. Gao, L. Cui, and J. Chen, "An incremental learning method based on dynamic ensemble RVM for intrusion detection," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 1, pp. 671–685, Mar. 2022.

[154] B. M. Xavier, R. S. Guimarães, G. Comarela, and M. Martinello, "MAP4: A pragmatic framework for in-network machine learning traffic classification," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 4, pp. 4176–4188, Dec. 2022.

[155] L. Yang, A. Moubayed, A. Shami, P. Heidari, A. Boukhtouta, A. Larabi, R. Brunner, S. Preda, and D. Migault, "Multi-perspective content delivery networks security framework using optimized unsupervised anomaly detection," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 1, pp. 686–705, Mar. 2022.

[156] L. Yang, Y. Song, S. Gao, A. Hu, and B. Xiao, "Griffin: Real-time network intrusion detection system via ensemble of autoencoder in SDN," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 3, pp. 2269–2281, Sep. 2022.

[157] Y. Yue, X. Chen, Z. Han, X. Zeng, and Y. Zhu, "Contrastive learning enhanced intrusion detection," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 4, pp. 4232–4247, Dec. 2022.

[158] Z. Zeng, W. Peng, and D. Zeng, "Improving the stability of intrusion detection with causal deep learning," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 4, pp. 4750–4763, Dec. 2022.

[159] G. Andresini, F. Pendlebury, F. Pierazzi, C. Loglisci, A. Appice, and L. Cavallaro, "INSOMNIA: Towards concept-drift robustness in network intrusion detection," in *Proc. 14th ACM Workshop Artif. Intell. Secur.*, 2021, pp. 111–122.

[160] P. Casas, P. Mulinka, and J. M. Vanerio, "Should I (re)Learn or should I Go(on): Stream machine learning for adaptive defense against network attacks," in *Proc. 6th ACM Workshop Moving Target Def.*, 2019, pp. 79–88.

[161] A. Drichel, B. Holmes, J. von Brandt, and U. Meyer, "The more, the better: A study on collaborative machine learning for DGA detection," in *Proc. 3rd Workshop Cyber-Secur. Arms Race*, 2021, pp. 1–12.

[162] C. Fu, Q. Li, M. Shen, and K. Xu, "Realtime robust malicious traffic detection via frequency domain analysis," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2021, pp. 3431–3446.

[163] C. Fu, Q. Li, K. Xu, and J. Wu, "Point cloud analysis for ML-based malicious traffic detection: Reducing majorities of false positive alarms," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2023, pp. 1005–1019.

[164] D. Han, Z. Wang, W. Chen, Y. Zhong, S. Wang, H. Zhang, J. Yang, X. Shi, and X. Yin, "DeepAID: Interpreting and improving deep learning-based anomaly detection in security applications," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2021, pp. 3197–3217.

[165] J. Holland, P. Schmitt, N. Feamster, and P. Mittal, "New directions in automated traffic analysis," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2021, pp. 3366–3383.

[166] A. S. Jacobs, R. Beltiukov, W. Willinger, R. A. Ferreira, A. Gupta, and L. Z. Granville, "AI/ML for network security: The emperor has no clothes," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2022, pp. 1537–1551.

[167] C. Novo and R. Morla, "Flow-based detection and proxy-based evasion of encrypted malware C2 traffic," in *Proc. 13th ACM Workshop Artif. Intell. Secur.*, 2020, pp. 83–91.

[168] M. Piskozub, R. Spolaor, M. Conti, and I. Martinovic, "On the resilience of network-based moving target defense techniques against host profiling attacks," in *Proc. 6th ACM Workshop Moving Target Def.*, 2019, pp. 1–12.

[169] Y. Shen, E. Mariconti, P. Vervier, and G. Stringhini, "Tiresias: Predicting security events through deep learning," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2018, pp. 592–605.

[170] A. Yehezkel, E. Elyashiv, and O. Soffer, "Network anomaly detection using transfer learning based on auto-encoders loss normalization," in *Proc. 14th ACM Workshop Artif. Intell. Secur.*, 2021, pp. 61–71.

[171] C. Zhang, X. Costa-Pérez, and P. Patras, "Tiki-Taka: Attacking and defending deep learning-based intrusion detection systems," in *Proc. ACM SIGSAC Conf. Cloud Comput. Secur. Workshop*, 2020, pp. 27–39.

[172] S. T. K. Jan, Q. Hao, T. Hu, J. Pu, S. Oswal, G. Wang, and B. Viswanath, "Throwing darts in the dark? Detecting bots with limited data using neural data augmentation," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2020, pp. 1190–1206.

[173] C. Fu, Q. Li, and K. Xu, "Detecting unknown encrypted malicious traffic in real time via flow interaction graph analysis," in *Proc. 30th Annu. Netw. Distrib. Syst. Secur. Symp.*, 2023. Accessed: Jun. 3, 2024. [Online]. Available: https://www.ndss-symposium.org/wp-content/uploads/2023/02/ndss2023_s80_paper.pdf

[174] D. Han, Z. Wang, W. Chen, K. Wang, R. Yu, S. Wang, H. Zhang, Z. Wang, M. Jin, J. Yang, X. Shi, and X. Yin, "Anomaly detection in the open world: Normality shift detection, explanation, and adaptation," in *Proc. 30th Annu. Netw. Distrib. Syst. Secur. Symp.*, 2023. Accessed: Jun. 3, 2024. [Online]. Available: https://www.ndss-symposium.org/wp-content/uploads/2023/02/ndss2023_f830_paper.pdf

[175] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: An ensemble of autoencoders for online network intrusion detection," in *Proc. 25th Annu. Netw. Distrib. Syst. Secur. Symp.*, 2018. Accessed: Jun. 3, 2024. [Online]. Available: https://www.ndss-symposium.org/wp-content/uploads/2018/02/ndss2018_03A-3_Mirsky_paper.pdf

[176] K. Wang, Z. Wang, D. Han, W. Chen, J. Yang, X. Shi, and X. Yin, "BARS: Local robustness certification for deep learning based traffic analysis systems," in *Proc. 30th Annu. Netw. Distrib. Syst. Secur. Symp.*, 2023. Accessed: Jun. 3, 2024. [Online]. Available: https://www.ndss-symposium.org/wp-content/uploads/2023/02/ndss2023_f508_paper.pdf

[177] J. Piet, A. Sharma, V. Paxson, and D. A. Wagner, "Network detection of interactive SSH impostors using deep learning," in *Proc. 32nd USENIX Secur. Symp.*, 2023, pp. 4283–4300.

[178] S. Schüppen, D. Teubert, P. Herrmann, and U. Meyer, "FANCI: Feature-based automated NXDomain classification and intelligence," in *Proc. 27th USENIX Secur. Symp.*, 2018, pp. 1165–1181.

[179] F. Wei, H. Li, Z. Zhao, and H. Hu, "xNIDS: Explaining deep learning-based network intrusion detection systems for active intrusion responses," in *Proc. 32nd USENIX Secur. Symp.*, 2023, pp. 4337–4354.

[180] L. Yang, W. Guo, Q. Hao, A. Ciptadi, A. Ahmadzadeh, X. Xing, and G. Wang, "CADE: Detecting and explaining concept drift samples for security applications," in *Proc. 30th USENIX Secur. Symp.*, 2021, pp. 2327–2344.

[181] G. Zhou, Z. Liu, C. Fu, Q. Li, and K. Xu, "An efficient design of intelligent network data plane," in *Proc. 32nd USENIX Secur. Symp.*, 2023, pp. 6203–6220.

[182] Y. Daihes, H. Tzaban, A. Nadler, and A. Shabtai, "MORTON: Detection of malicious routines in large-scale DNS traffic," in *Proc. 26th Eur. Symp. Res. Comput. Secur.*, vol. 12972, 2021, pp. 736–756.

[183] X. Hong, Z. C. Papazachos, J. M. del Rincón, and P. Miller, "Network intrusion detection by variational component-based feature saliency Gaussian mixture clustering," in *Proc. ESORICS Int. Workshops CPS4CIP, ADIoT, SecAssure, WASP, TAURIN, PriST-AI, SECAI*, vol. 14399, 2023, pp. 761–772.

[184] G. Stergiopoulos, A. Talavari, E. Bitsikas, and D. Gritzalis, "Automatic detection of various malicious traffic using side channel features on TCP packets," in *Proc. 23rd Eur. Symp. Res. Comput. Secur.*, vol. 11098, 2018, pp. 346–362.

[185] S.-C. Chen, Y.-R. Chen, and W.-G. Tzeng, "Effective botnet detection through neural networks on convolutional features," in *Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./12th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2018, pp. 372–378.

[186] H. Dhillon and A. Haque, "Towards network traffic monitoring using deep transfer learning," in *Proc. IEEE 19th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Dec. 2020, pp. 1089–1096.

[187] T. Fernandes, L. Dias, and M. Correia, "C2BID: Cluster change-based intrusion detection," in *Proc. IEEE 19th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Dec. 2020, pp. 310–319.

[188] P. Fu, Q. Yang, Y. Guan, B. Wang, G. Gou, Z. Li, G. Xiong, and Z. Li, "Towards multi-source extension: A multi-classification method based on sampled NetFlow records," in *Proc. IEEE 20th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Oct. 2021, pp. 1069–1076.

[189] X. Han, R. Yin, Z. Lu, B. Jiang, Y. Liu, S. Liu, C. Wang, and N. Li, "STIDM: A spatial and temporal aware intrusion detection model," in *Proc. IEEE 19th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Dec. 2020, pp. 370–377.

[190] G. Jia, P. Miller, X. Hong, H. Kalutarage, and T. Ban, "Anomaly detection in network traffic using dynamic graph mining with a sparse autoencoder," in *Proc. 18th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./13th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2019, pp. 458–465.

[191] Z. Li, Y. Wang, P. Wang, and H. Su, "PGAN: A generative adversarial network based anomaly detection method for network intrusion detection system," in *Proc. IEEE 20th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Oct. 2021, pp. 734–741.

[192] J. Li, Z. Zhang, Y. Li, X. Guo, and H. Li, "FIDS: Detecting DDoS through federated learning based method," in *Proc. IEEE 20th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Oct. 2021, pp. 856–862.

[193] L. Lu, X. Zhu, X. Zhang, J. Liu, M. Z. A. Bhuiyan, and G. Cui, "One intrusion detection method based on uniformed conditional dynamic mutual information," in *Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./12th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2018, pp. 1236–1241.

[194] B. Qi, J. Jiang, Z. Shi, R. Mao, and Q. Wang, "BotCensor: Detecting DGA-based botnet using two-stage anomaly detection," in *Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./12th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2018, pp. 754–762.

[195] L. Sacramento, I. Medeiros, J. Bota, and M. Correia, "FlowHacker: Detecting unknown network attacks in big traffic data using network flows," in *Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./ 12th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2018, pp. 567–572.

[196] S. Saganowski, "A three-stage machine learning network security solution for public entities," in *Proc. IEEE 19th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Dec. 2020, pp. 1097–1104.

[197] L. Su, Y. Yao, Z. Lu, and B. Liu, "Understanding the influence of graph kernels on deep learning architecture: A case study of flow-based network attack detection," in *Proc. 18th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./13th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2019, pp. 312–318.

[198] L. Su, Y. Yao, N. Li, J. Liu, Z. Lu, and B. Liu, "Hierarchical clustering based network traffic data reduction for improving suspicious flow detection," in *Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./12th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2018, pp. 744–753.

[199] Y. Yao, L. Su, Z. Lu, and B. Liu, "STDeepGraph: Spatial–temporal deep learning on communication graphs for long-term network attack detection," in *Proc. 18th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./13th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2019, pp. 120–127.

[200] R. Zhang, M. Tong, L. Chen, J. Xue, W. Liu, and F. Xie, "CMIRGen: Automatic signature generation algorithm for malicious network traffic," in *Proc. IEEE 19th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Dec. 2020, pp. 736–743.

[201] J. R. Goodall, E. D. Ragan, C. A. Steed, J. W. Reed, G. D. Richardson, K. M. T. Huffer, R. A. Bridges, and J. A. Laska, "Situ: Identifying and explaining suspicious behavior in networks," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 1, pp. 204–214, Jan. 2019.

[202] N. Danneman and R. Gove, "Tuning automatic summarization for incident report visualization," in *Proc. IEEE 15th Pacific Visualizat. Symp. (PacificVis)*, Apr. 2022, pp. 191–195.

[203] J. L. Guerra, E. Veas, and C. A. Catania, "A study on labeling network hostile behavior with intelligent interactive tools," in *Proc. IEEE Symp. Visualizat. Cyber Secur. (VizSec)*, Oct. 2019, pp. 1–10.

[204] M. Ring, S. Wunderlich, D. Scheuring, D. Landes, and A. Hotho, "A survey of network-based intrusion detection data sets," *Comput. Secur.*, vol. 86, pp. 147–167, Sep. 2019.

[205] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *Proc. Int. Conf. Inf. Syst. Secur. Privacy*, vol. 1, 2018, pp. 108–116.

[206] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *Proc. IEEE Symp. Comput. Intell. Secur. Defense Appl.*, Jul. 2009, pp. 1–6.

[207] I. Sharafaldin, A. H. Lashkari, S. Hakak, and A. A. Ghorbani, "Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy," in *Proc. Int. Carnahan Conf. Secur. Technol. (ICCST)*, Oct. 2019, pp. 1–8.

[208] S. Garcia, M. Grill, J. Stiborek, and A. Zunino, "An empirical comparison of botnet detection methods," *Comput. Secur.*, vol. 45, pp. 100–123, Sep. 2014.

[209] R. Fontugne, P. Borgnat, P. Abry, and K. Fukuda, "MAWILab: Combining diverse anomaly detectors for automated anomaly labeling and performance benchmarking," in *Proc. ACM Int. Conf. Emerg. Netw. Exp. Technol.*, Dec. 2010, pp. 1–12.

[210] A. Shiravi, H. Shiravi, M. Tavallaee, and A. A. Ghorbani, "Toward developing a systematic approach to generate benchmark datasets for intrusion detection," *Comp. Secur.*, vol. 31, no. 3, pp. 357–374, 2012.

[211] H. H. Jazi, H. Gonzalez, N. Stakhanova, and A. A. Ghorbani, "Detecting HTTP-based application layer DoS attacks on web servers in the presence of sampling," *Comput. Netw.*, vol. 121, pp. 25–36, Jul. 2017.

[212] M. Wieling, J. Rawee, and G. van Noord, "Reproducibility in computational linguistics: Are we willing to share?" *Comput. Linguistics*, vol. 44, no. 4, pp. 641–649, 2018.

[213] G. Engelen, V. Rimmer, and W. Joosen, "Troubleshooting an intrusion detection dataset: The CICIDS2017 case study," in *Proc. IEEE Secur. Privacy Workshops (SPW)*, May 2021, pp. 7–12.

[214] D. Olszewski, A. Lu, C. Stillman, K. Warren, C. Kitroser, A. Pascual, D. Ukirde, K. Butler, and P. Traynor, "'Get in researchers; we're measuring reproducibility': A reproducibility study of machine learning papers in tier 1 security conferences," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2023, pp. 3433–3459.

[215] F. B. Kokulu, A. Soneji, T. Bao, Y. Shoshitaishvili, Z. Zhao, A. Doupé, and G.-J. Ahn, "Matched and mismatched SOCs: A qualitative study on security operations center issues," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2019, pp. 1955–1970.

[216] T. J. Hastie, "Generalized additive models," in *Statistical Models in S*. Evanston, IL, USA: Routledge, 2017, pp. 249–307.

[217] B. Abdollahi and O. Nasraoui, "Explainable restricted Boltzmann machines for collaborative filtering," 2016, *arXiv:1606.07129*.

[218] S. Hart, "Shapley value," in *Game Theory*. London, U.K.: Palgrave Macmillan, 1989, pp. 210–216.

[219] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?' Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 1135–1144.

[220] Z. Shi Li, C. Werner, N. Ernst, and D. Damian, "GDPR compliance in the context of continuous integration," 2020, *arXiv:2002.06830*.

[221] R. Brown and K. Nickels, "SANS 2023 CTI survey: Keeping up with a changing threat landscape," SANS Inst., Rockville, MD, USA, Tech. Rep., Jul. 2023. Accessed: Jun. 3, 2024. [Online]. Available: https://www.sans.org/white-papers/2023-cti-survey-keeping-up-changing-threat-landscape/

KATHARINA DIETZ (Student Member, IEEE) received the master's degree in computer science from the University of Würzburg, Germany, in 2020, where she is currently pursuing the Ph.D. degree with the Chair of Communication Networks. She is also a Research Assistant with the Chair of Communication Networks, University of Würzburg. Her research interests include managing and securing communication networks with machine learning-based approaches, ranging from performance prediction to anomaly detection.

MICHAEL MÜHLHAUSER received the degree in information systems from the University of Bamberg. He is currently a Research Assistant with the Chair for Privacy and Security in Information Systems, University of Bamberg, Germany. His research interests include privacy, including the analysis of network traffic to uncover potential privacy problems.

JOCHEN KÖGEL received the Diploma degree in electrical engineering and information technology, the master's degree in information technology, and the Ph.D. degree in electrical engineering and information technology from the University of Stuttgart, Germany. He works in research and development with IsarNet Software Solutions GmbH, where he is responsible for research activities and system/software architecture of the IsarFlow network monitoring software. His research interests include large-scale data analysis for achieving a better understanding of complex communication networks for networking professionals.

STEPHAN SCHWINGER received the Diploma degree in technomathematics from the Chemnitz University of Technology and the Ph.D. degree in mathematics from Technical University Berlin. He is currently a Data Scientist at genua GmbH, Kirchheim near Munich. His research interests include machine learning, big data analysis, and network security.

MARLEEN SICHERMANN received the master's degree in computer science from the Chair of Communication Networks, University of Würzburg, Germany, in 2023, where she is currently pursuing the Ph.D. degree. She is also a Research Assistant with the Chair of Communication Networks, University of Würzburg. Her research interests include performance evaluation and modeling of communication systems, along with traffic measurements and modeling.

MICHAEL SEUFERT (Senior Member, IEEE) received the Diploma and Ph.D. degrees in computer science, the bachelor's degree in economathematics, and the Habilitation degree in computer science from the University of Würzburg, Germany, in 2011, 2017, 2018, and 2023, respectively. He also received the First State Examination degree in mathematics, computer science, and education for teaching in secondary schools, in 2011. He is a Full Professor with the University of Augsburg, Germany, heading the Chair of Networked Embedded Systems and Communication Systems. His research focuses on user-centric communication networks, including QoE of internet applications, AI/ML for QoE-aware network management, as well as group-based communications.

DOMINIK HERRMANN received the Diploma degree (Hons.) in management information systems from the University of Regensburg, in 2008, and the Ph.D. degree in computer science from the University of Hamburg, in 2014. From 2015 to 2017, he was a Temporary Professor with the University of Siegen. Since October 2017, he has been a Full Professor of Privacy and Security in Information Systems with the University of Bamberg, Germany.

TOBIAS HOẞFELD (Senior Member, IEEE) was the Head of the Chair of Modeling of Adaptive Systems, University of Duisburg–Essen, Germany, from 2014 to 2018. Since 2018, he has been a Full Professor and the Head of the Chair of Communication Networks, University of Würzburg, Germany. He is a member of the editorial board of IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, ACM SIGMM Records, and *Quality and User Experience* (Springer).