

NRC Publications Archive Archives des publications du CNRC

An exploratory study on domain knowledge infusion in deep learning for automated threat defense

Khazadeh, Sourena; Pinto Neto, Euclides Carlos; Iqbal, Shahrear; Alalfi, Manar; Buffett, Scott

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. / La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

For the publisher's version, please access the DOI link below. / Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

Publisher's version / Version de l'éditeur:

<https://doi.org/10.1007/s10207-025-00987-4>

International Journal of Information Security, 24, pp. 1-19, 2025-01-28

NRC Publications Archive Record / Notice des Archives des publications du CNRC :

<https://nrc-publications.canada.ca/eng/view/object/?id=2f0c54c5-9f3a-42aa-ba58-3c03a1c4873a>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=2f0c54c5-9f3a-42aa-ba58-3c03a1c4873a>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.



An exploratory study on domain knowledge infusion in deep learning for automated threat defense

Sourena Khanzadeh^{1,2} · Euclides Carlos Pinto Neto¹ · Shahrear Iqbal¹ · Manar Alalfi² · Scott Buffett¹

© Crown 2025

Abstract

The wide adoption of interconnected services leads to the creation of supportive solutions and business opportunities. Conversely, this new paradigm is targeted by malicious activities, aiming to compromise systems' confidentiality, integrity, and availability. However, advanced methods lack contextual awareness, which prevents their deployment to real-world systems. Considering that the process of making informed decisions stems from the expertise of analysts based on their experience, the use of cybersecurity domain knowledge has the potential to improve Deep Learning and Deep Reinforcement Learning operations in real scenarios. Therefore, the main goal of this research is to study and evaluate the use of Knowledge Infused Learning in the context of automated threat defense. We define how cybersecurity domain knowledge can be infused into Deep Learning and Reinforcement Learning, highlighting the main challenges and benefits. Besides, we present a roadmap to apply domain knowledge for red and blue teaming activities and discuss the implications of Knowledge Infused Learning in explainability, and actionable reporting. Finally, we list the open challenges to guide the development of next-generation security solutions.

Keywords Cybersecurity · Knowledge Infusion · Explainable Artificial Intelligence

1 Introduction

In the contemporary era, the vast number of technological advancements fosters a cohesive environment for sharing information, services, and conducting business. Conversely, this new paradigm is targeted by malicious actors seeking to compromise systems' confidentiality, integrity, and availability [1, 2]. The growth of technology directly correlates with the ever-evolving nature of cyber attacks, resulting

in increasingly compromised systems. Traditional methods, while effective in the past, are no longer sufficient to combat these sophisticated cyberattacks. They are being outmaneuvered, highlighting the need for more dynamic and adaptive solutions.

Automated Threat Defense (ATD) has emerged as a crucial approach to address these challenges. ATD involves using advanced technologies and systems to automatically detect, analyze, and respond to cybersecurity threats in real-time, minimizing the impact of security incidents by reducing the time between threat detection and response.

Employing Artificial Intelligence (AI), specifically Deep Learning (DL) and deep reinforcement learning (DRL) [3], is essential for addressing modern ADT systems. These methods leverage vast amounts of data to address high-dimensional, non-linear problems, enabling rapid data processing, anomaly detection, and the automation of defensive responses, thus enhancing cybersecurity against sophisticated attackers. However, DL and DRL currently lack sufficient contextual awareness, which limits their deployment in real-world systems.

To address the limitations of traditional AI models, domain knowledge can be integrated using a technique

✉ Shahrear Iqbal
Shahrear.Iqbal@nrc-cnrc.gc.ca

Sourena Khanzadeh
sourena.khanzadeh@torontomu.ca

Euclides Carlos Pinto Neto
EuclidesCarlos.PintoNeto@nrc-cnrc.gc.ca

Manar Alalfi
manar.alalfi@torontomu.ca

Scott Buffett
Scott.Buffett@nrc-cnrc.gc.ca

¹ National Research Council Canada, Fredericton, NB, Canada

² Department of Computer Science, Toronto Metropolitan University (TMU), Toronto, ON, Canada

called knowledge-infused learning. This approach enhances AI models by incorporating structured knowledge, such as rules, ontologies, and symbolic representations. These elements enable AI models to better understand and interpret data within specific contexts, allowing for more informed decision-making and improved performance in complex and dynamic environments.

The main goal of this research is to study the effective application of knowledge-infused learning to improve AI models for automated threat defense. We investigate solutions available in the literature and discuss how they can be leveraged to solve existing complex problems. Additionally, we present a roadmap for integrating domain knowledge in red and blue teaming activities and discuss the implications of knowledge-infused learning on explainability, actionable reporting, and the use of Large Language Models (LLMs) in knowledge management. Finally, we identify the open challenges that must be addressed to guide the development of next-generation security solutions. Our contributions are as follows:

- We provide insights into how Knowledge-Infused Learning (KIL) can enhance Deep Learning (DL) and Reinforcement Learning (RL) models by improving accuracy, explainability, and adaptability.
- We critically examine existing methodologies and identify key opportunities and challenges in leveraging expert knowledge to bolster cybersecurity strategies against increasingly sophisticated threats.
- We offer a comprehensive roadmap for future research and practical solutions to advance the adoption of Knowledge-Infused Learning in cybersecurity.

This paper is organized as follows:

Section 2 offers a literature review of related work. Section 3 provides a comprehensive yet concise background on KIL, introducing DL, DRL, and Cybersecurity. Section 4 explores domain knowledge infusion, explaining what domain knowledge is, how it can be infused into machine learning models, and providing more details on KIL in terms of DL and RL. Section 5 discusses how KIL can improve automated threat defense, including application areas. Section 6 identifies research gaps and potential future directions. Finally, Sect. 7 draws the overall conclusions of this study.

2 Related work

Knowledge infusion plays a pivotal role in cybersecurity, allowing for the integration of diverse sources of information to enhance detection, prevention, and response mechanisms. This section provides an overview of existing literature on knowledge infusion in cybersecurity, focusing on its impact

on threat intelligence, anomaly detection, and automated response systems.

2.1 Techniques of KIL in the literature

In literature, infusing knowledge into RL has a few approaches, and one of the methods to perform this task is to merge prior knowledge represented as Cybersecurity Knowledge Graph (CKGs) to guide the exploration of an RL agent for malware detection [4].

A supervised RL approach proposed by Moreno et al. [5] utilizes external knowledge and validates it in specific tasks, such as “wall following” in the case of mobile robotics. One direct way is to infuse knowledge in Deep Neural Networks (DNNs), proposed by multiple research efforts [6–8], approaches shallow, semi-deep, and deep infusion methods. Knowledge injection in DNN is a powerful paradigm that includes external knowledge into neural network architectures, allowing higher reasoning in downstream tasks. This approach eases the learning process and grants neural networks the ability to take decisive measures to reason for the improvement of downstream tasks.

2.2 Surveys about KIL

Li et al. [9] survey the use of knowledge graphs in industrial products and services, including their application in DRL models. Graph convolutional networks, which can process graph data, have shown promise for infusing expert knowledge into DRL models. Hu et al. [10] propose a Petri-net-based dynamic scheduling system that uses a DRL model with a graph convolutional network. Inoue et al. [11] also propose a DRL model for high-precision assembly tasks, while Schoettler et al. [12] use meta-reinforcement learning for robotic industrial insertion tasks.

However, these techniques face challenges, such as the lack of appropriately structured datasets and the need for effective knowledge representation. Zhou et al. [13] presents an approach where leveraging Knowledge Graphs (KG) for Reinforcement Learning in interactive recommender systems to address limitations of existing DRL methods.

Several surveys have explored the idea of machine learning and artificial intelligence in cybersecurity, where knowledge infusion is central to improving accuracy and reducing false positives.

For instance, Kim et al. [14] presents a detailed survey on knowledge integration in deep learning, albeit in a mechanical engineering context. A notable concept from the paper is the use of informed deep learning, where prior knowledge is integrated at different stages of the learning process, leading to more robust models. Three methods of knowledge integration: feature engineering, designing, and regularizing. These methods are used to enhance traditional deep learn-

ing by infusing it with additional domain knowledge. In the cybersecurity domain, feature engineering can involve integrating specific threat intelligence or system configuration data to enhance model performance. Designing new neural network architectures to incorporate domain-specific knowledge is another method discussed in the paper, which can be adapted to cybersecurity by incorporating threat models or knowledge graphs. Regularizing techniques help constrain models to adhere to known rules or patterns, which can be useful in cybersecurity to prevent false positives or adhere to known threat behavior patterns. The paper also discusses different deep learning architectures such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), emphasizing the flexibility of these models to incorporate prior knowledge.

Giuseppe Futia and Antonio Vetrò [15] focus on explainable artificial intelligence and proposes a neural-symbolic way to bridge the gap between DL and symbolic AI. The paper highlights the critical role that Knowledge Graphs (KG's) play in making AI systems more transparent and understandable, particularly in contexts where AI impacts human life, such as recruitment, medical diagnoses, and judicial decisions. KG's with their semantic networks of entities and relationships, can serve as a bridge between complex deep learning models and human-understandable explanations.

Dash et al. [16] presents a comprehensive survey of various approaches to infuse domain knowledge into deep learning models. This paper categorizes these techniques into three broad categories: transforming input data, modifying loss functions, and altering model structures.

Incorporating domain knowledge can start at the data input stage by altering the dataset to include more domain-specific features. This can be achieved through:

Transforming complex relational data into simpler, more structured formats. This technique uses methods such as Inductive Logic Programming (ILP) to convert complex relationships into features suitable for deep neural networks. For instance, propositionalization might flatten a set of logical relations into a feature-based dataset suitable for a multilayer perceptron (MLP). Knowledge Graphs: Knowledge graphs represent relational domain knowledge, typically as labeled graphs where nodes represent entities and edges represent relationships. These graphs can be integrated with deep learning models to enrich the information used during training. Knowledge-infused learning can be achieved by incorporating these graphs into the training process of various deep network architectures, leading to improved explainability and accuracy.

A common approach to infusing domain knowledge involves altering the loss function used to train deep learning models. This can be done by: Semantic Loss: Introducing additional penalty terms into the loss function based

on domain-specific constraints. These penalties can ensure that the model's predictions align with logical or numerical domain rules. Regularisation: Regularising embedding from declarative knowledge encoded in first-order logic, or adding terms to the loss function based on domain-specific relationships, helps guide the model toward domain-compliant solutions.

Another way to infuse domain knowledge is by designing deep learning models with specific architectures that inherently incorporate this knowledge. This involves:

Specialized Structures: Constructing network architectures that reflect domain-specific rules. For example, Knowledge-Based Artificial Neural Networks (KBANNs) create network structures from propositional rules, ensuring that domain knowledge is integrated into the model's architecture. Bayesian Formulations: Incorporating domain knowledge as prior distributions over model parameters. This approach allows constraints based on domain knowledge to guide model learning in a Bayesian framework.

Xiaozheng Xie et al. [17] delve into methods to improve deep learning models by integrating medical domain knowledge. The key to integrating domain knowledge into deep learning is to use established information from experts to guide or constrain models to better align with domain-specific expectations. This survey explores techniques from the medical domain, which can be adapted to other fields such as cybersecurity.

The survey provides several methods for integrating domain knowledge into deep learning. Feature-level integration involves combining domain-specific features with those extracted by deep learning models. In medical image analysis, this is achieved through concatenating features or merging handcrafted features with deep learning outputs. In cybersecurity, feature-level integration could involve combining network traffic statistics with other cybersecurity data. The survey discusses how certain deep learning architectures can incorporate medical knowledge, e.g., simulating the diagnostic patterns of medical professionals. In cybersecurity, knowledge-based architectures can mirror typical security analyst workflows or threat detection patterns. Medical datasets often suffer from limited data availability, so transfer learning from larger natural image datasets is common. In cybersecurity, transfer learning could be used to incorporate broader data sources, while multi-task learning allows for learning from multiple tasks, potentially increasing the generalization ability of the models.

Finally, relevant correlated research works focus on different issues. For example, the authors in [18] and [19] focus on analyzing security and risk management in IoT systems. Similarly, the authors in [20] review solutions for security automation focusing on specific threats. Conversely, these works do not comprehensively explore domain knowledge to improve security automation. A comprehensive evalua-

Table 1 Overview of literature review with focus on cybersecurity aspects

Reference	Knowledge Graph Integration	Domain Knowledge Infusion	Cybersecurity Application Direct	Limitation in Scope for Cybersecurity
Survey				
Kim et al. [14]	✗	✓	✗	✗
Futia and Vetrò [15]	✓	✓	✓	●
Dash et al. dash2022review	✗	✓	●	✗
Xiaozheng Xie et al. [17]	✗	✓	✗	✓
Ours	●	✓	✓	●
Techniques				
Piplai et al. [4]	✓	✓	✓	●
Moreno et al. [5]	✗	✓	✗	✗
Kursuncu et al. [6]	✗	✓	✗	✗
Sheth et al. [7]	✗	✓	✗	✗
Gaur et al. [8]	✗	✓	✗	✗
Li et al. [9]	✓	✓	✓	●
Hu et al. [10]	✓	✓	✓	●
Inoue et al. [11]	✓	✓	✓	●
Schoettler et al. [12]	✓	✓	✓	●
Zhou et al. [13]	✓	✓	✓	●

- ✓(Check Mark): Indicates that the aspect is fully addressed or applicable
- ✗(Cross Mark): Indicates that the aspect is not addressed or not applicable
- ●(Half Mark): Indicates partial relevance or consideration but not fully developed

tion of KIL in the context of automated threat defense is presented in the following sections.

In Table 1, we compare our work with related surveys. The legend and guide on how to read the table is as follows:

- **Knowledge Graph Integration** Does the work integrate knowledge graphs effectively?
- **Domain Knowledge Infusion** Does the work effectively infuse domain-specific knowledge?
- **Cybersecurity Application Direct** Is there a direct application to cybersecurity?
- **Limitation in Scope for Cybersecurity** Are there limitations that reduce its applicability to cybersecurity?

3 Background

This section outlines the fundamental concepts essential for understanding Knowledge-Infused Learning. It is divided into four main subsections: deep learning, reinforcement learning, deep reinforcement learning, and a brief overview of cybersecurity.

3.1 Deep learning (DL)

Deep learning [21], a subset of machine learning, leads numerous applications requiring human-like decision mak-

ing. It utilizes Deep Neural Networks (DNNs), which calculate neurons (nodes or cells within each layer) in a forward pass [22]. The parameters, mainly composed of weights and biases, are then trained by back-propagating after calculating the loss function. A simple example of a loss function is the mean squared error (MSE), written as:

$$MSE = \frac{1}{N} \sum_{i=1}^n (x_i - y_i)^2$$

where N is the size of the data frame y_i is the predicted output and x_i is the ground truth value.

Moreover, a very basic DNN consists of one input neuron connected to one output neuron, as more complex problems require more complex structure of the DNN model, there becomes an immense reliance on what is called hidden layers. Hidden layers are intermediary layers responsible for learning intricate structures of data and finding patterns, they are always in between inputs and outputs. In a dense neural network, the output of the input neurons is connected to the input of the first hidden layer neurons. The output of these hidden layer neurons is then connected to the input of the subsequent hidden layers. Finally, the output of the last hidden layer neurons is connected to the output of the DNN.

3.2 Deep reinforcement learning (DRL)

DRL [23] is a subset of Reinforcement Learning (RL) [24] where deep neural networks are used to approximate complex functions involved in RL processes. There are various approaches in DRL:

- **Deep Q-Networks (DQN)** [25] This technique uses neural networks to approximate the Q-values in Q-learning, a common RL algorithm. The neural network predicts the expected cumulative reward for each action in a given state.
- **Policy Gradient Methods** [26] These methods directly learn a parameterized policy $\pi_{\theta}(a|s)$, where θ represents the neural network's parameters. Examples include REINFORCE [27], Advantage Actor-Critic (A2C) [28], and Proximal Policy Optimization (PPO) [29].
- **Actor-Critic Architectures** [30] These architectures consist of two neural networks: one for the policy (actor) and one for the value function (critic). The actor chooses actions, while the critic evaluates them to guide the actor's learning.
- **Model-based DRL** [31] This approach learns a model of the environment to plan and make decisions. It uses neural networks to predict the environment's dynamics and outcomes of actions, aiding in decision-making and planning.

DRL has been used in various applications, from playing complex games (e.g., Go and Chess) to autonomous driving and robotics, demonstrating its versatility and effectiveness in solving intricate problems in uncertain and dynamic environments.

3.3 Automated threat defense (ATD)

ATD has emerged as a vital strategy in modern cybersecurity, addressing the limitations of traditional static and signature-based systems that struggle against sophisticated and dynamic cyber threats. ATD leverages advanced technologies, including AI, Machine Learning (ML), and specifically DL and DRL, to create systems capable of detecting, analyzing, and responding to threats in real-time, thereby reducing the potential impact of security incidents [32, 33]. By employing AI and ML algorithms, ATD systems can identify patterns and anomalies within vast datasets, enabling the detection of both known and unknown threats. These systems conduct comprehensive threat analysis by correlating data from multiple sources and subsequently initiate automated responses such as isolating affected systems or blocking malicious IP addresses [3]. Integration with other security tools such as firewalls, SIEM, and SOAR platforms ensures a coordinated approach, while continuous learning capabilities

allow these systems to adapt and refine their algorithms over time [34]. Despite these advancements, challenges remain, particularly in the area of contextual awareness, where integrating domain knowledge through KIL is being explored to enhance AI models' decision-making capabilities. Through ongoing research and development, ATD holds the promise of becoming a cornerstone of next-generation cybersecurity strategies, providing a proactive and intelligent approach to managing modern cyber threats [32].

4 Domain knowledge infusion

This section introduces the core concepts of knowledge-infused learning and is divided into three main subsections. Sect. 4.1 defines and explains the significance of domain knowledge in the context of AI. Section 4.2 explores various methods for integrating domain knowledge into machine learning models. Section 4.4 focuses on metrics and comparisons, providing an in-depth analysis of the performance and benefits of KIL in DL and RL.

4.1 Domain knowledge

In the context of AI and ML, domain knowledge encompasses the rules, principles, and insights that experts in a field use to make informed decisions and solve problems [35]. This knowledge is typically acquired through extensive experience, education, and practice within specific domains such as healthcare, finance, engineering, and cybersecurity.

For instance, in cybersecurity, domain knowledge includes an understanding of various types of cyber threats, attack vectors, defense mechanisms, and the specific protocols and systems used to secure networks and data. This knowledge is crucial for developing AI models and algorithms that can effectively detect and extenuate security breaches. Domain knowledge can be represented in structured forms such as databases, ontologies, taxonomies, and graphs, or unstructured forms such as expert opinions, research papers, and case studies. By infusing domain knowledge into AI models, it becomes possible to enhance their performance, accuracy, and explainability.

Integrating domain knowledge helps AI models understand and interpret data within specific contexts more accurately. It allows the models to make more informed decisions by leveraging expert-curated knowledge, leading to better performance in complex and dynamic environments. For example, in cybersecurity, incorporating knowledge about common attack patterns and known vulnerabilities can help AI systems more effectively identify and respond to threats. This enhanced contextual awareness makes the models more robust and capable of handling real-world scenarios where

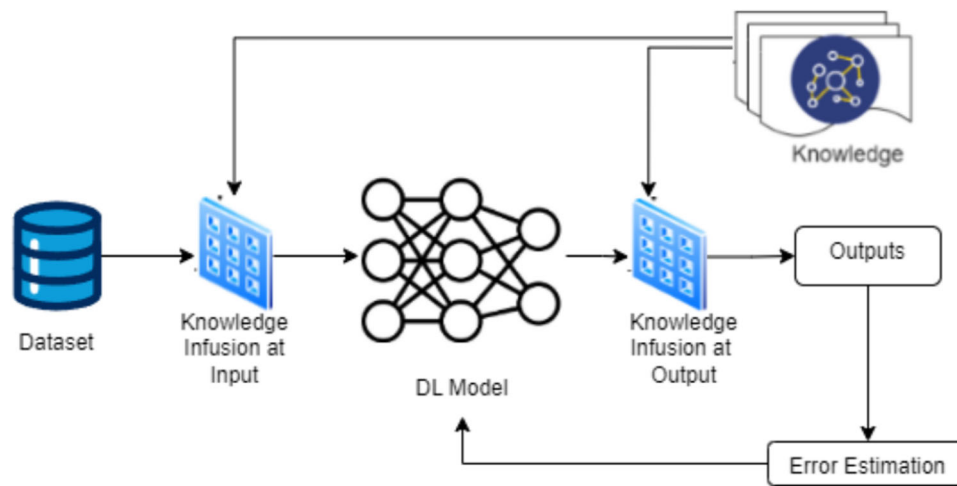


Fig. 1 Shallow infusion

understanding the broader context is essential for accurate decision-making.

4.2 Knowledge infusion into machine learning models

The infusion of expert knowledge in DL models can enhance their performance, accuracy, and explainability. Seth et al. [7] present three techniques for infusing knowledge into DL models: Shallow Infusion, Semi-Deep Infusion, and Deep Infusion.

Shallow Infusion Represents the simplest form of infusion. The infusion of knowledge in this method incorporates external knowledge for DL models without significantly altering its architecture. External knowledge can be incorporated as input (or output) to deep learning models or using a regular term that includes training knowledge. One example of a shallow infusion being applied to a field is Natural Language Processing (NLP). NLP is tasked to capture information about statistical data of words or phrases to enhance the overall performance. To apply shallow infusion in NLP we can commence by adding domain-specific knowledge as input features, such as named entities, part-of-speech tags, or syntactic dependencies.

These features enable a model to comprehend more nuances of the given text and drastically improve the performance of activities such as sentiment analysis, text classification, or question answering. In computer vision, shallow infusion can be assembled by incorporating expert-defined rules, such as object detection or segmentation rules, into the model. For example, a model trained to detect objects in images can be better captured by incorporating the shape, size, and color of objects in the training data. Another method to utilize shallow infusion is to inject external knowledge sources, such as Knowledge Graphs (KGs), into the DL

model architecture. Figure 1 represents the knowledge being infused at either input or output in order to improve the performance of the model.

Semi-Deep Infusion Similar to shallow infusion, semi-deep infusion incorporates expert knowledge with the difference that it inserts it into the hyperparameters of the DL model. The ultimate goal of this method is to leverage domain knowledge to adapt the process of learning [7]. This can be accomplished by adopting knowledge into attention mechanisms, learnable constraints, regularization functions, loss functions, activation functions, and optimization methods. In NLP, Semi-deep infusion may also entail integrating semantic constraints or linguistic rules into the model structure to help create coherent and contextually suitable material. Semi-deep infusion allows DL models to learn from expert knowledge without overhauling the model's ability to learn from data. This strikes a balance between leveraging domain-specific knowledge and allowing the model to be in sync and learn from the training data. Figure 2 showcases the learning constraints of the DL model depending on the expert knowledge infused into the system.

Deep Infusion Being at the extreme end, deep infusion is the most complex method of knowledge infusion. This method aims to fuse the feature representation generated by the original DL process with new representations extracted from the knowledge base. There are several ways of extracting feature representation from a knowledge graph, which enables the knowledge infusion to happen in multiple representation layers of DL architectures [7].

Deep infusion combines knowledge sources at different stratified levels of abstraction to be transferred across different layers of the deep learning model. In our own words, deep infusion allows for multi-layer representation, granting the model the ability to learn complex data representations in low-level and high-level concepts for the task at hand. It

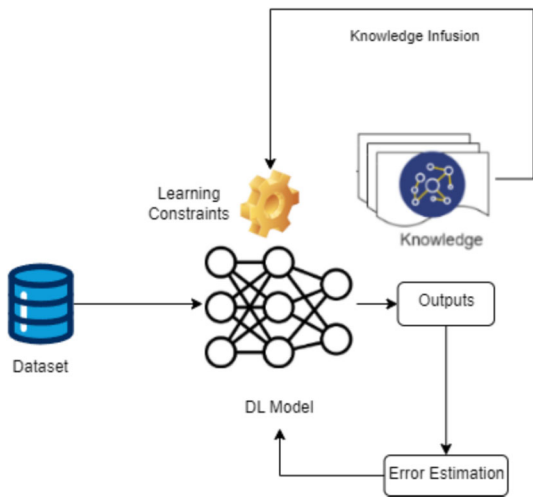


Fig. 2 Semi-deep infusion

is important to note that the concept of deep infusion is not well-defined in the literature and requires further research. Figure 3 represents deep infusion at its core, combining a DL representation with a knowledge graph to enhance the performance of DL models.

Similar to DL, RL offers various methods to leverage domain knowledge for improved performance and contextual awareness. In addition to the traditional DL infusion methods used in the internal DL model of DRL, three other infusion methods can be considered in the learning process: Reward infusion, State infusion, and Policy infusion.

Reward Infusion Domain knowledge can make an RL agent’s rewards more relevant and useful for a given task.

Although reward signals are essential to RL agents’ efficiency and early convergence, they are difficult to estimate for several reasons. In many cases, engineering reward functions that can precisely measure the quality of a given state become difficult when the environment is complex. In the case of cybersecurity, the estimation of rewards can be misleading in many ways, e.g., the system may present an unknown vulnerability, which produces a misleading sense of security. Incorporating domain knowledge into the calculation of reward signals enhances the understanding of the task at hand and the context in which actions are taken, compared to using a generic reward function. This may entail modifying the reward in accordance with particular requirements or results that, within the context of the domain, are preferred. The RL model can better comprehend the intricacies of the world it operates in by integrating context awareness into the reward function.

Besides, by eliminating the need for complex, manually built reward adjustments, situational awareness may be included in reward systems, which makes the process of building reward functions easier. A context-aware reward system has the ability to dynamically modify the reward in response to the current circumstances, eliminating the need to create intricate reward functions that attempt to account for every case. Reward infusion is illustrated in Fig. 4, where the reward is the input of knowledge infusion and subsequently RL agent.

State Infusion In RL, a state represents the environment that captures all the relevant information the agent is supposed to make decisions. It is the collection of all the observable variables accessible to the agent at any point in

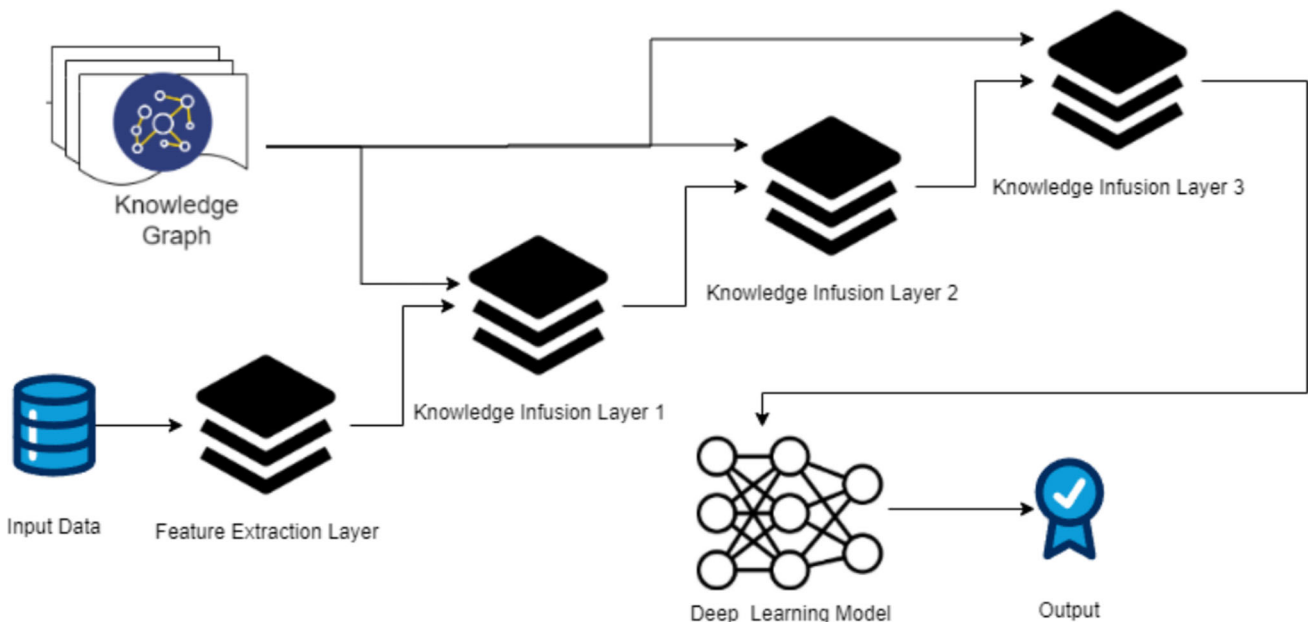


Fig. 3 Deep infusion

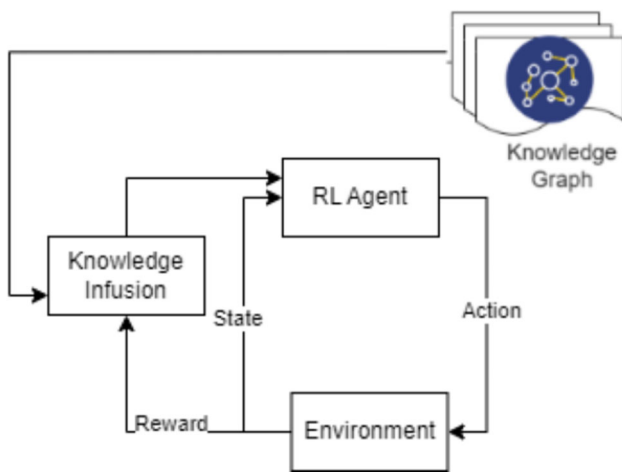


Fig. 4 Reward infusion

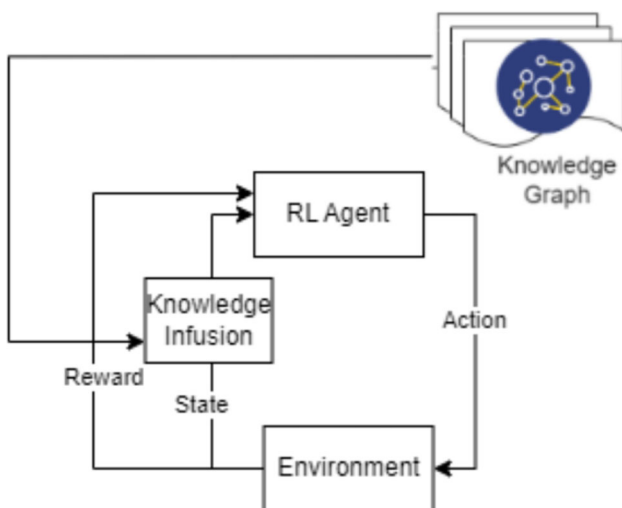


Fig. 5 State infusion

time T . The state can be thought of as a snapshot of the environment at a given moment represented using efficient data structures.

Infusing knowledge into RL can benefit the state representation in several ways. First, knowledge can be used to augment the state with additional information that is not directly observable. For example, in a game of chess [36], the agent might not be able to see the entire board, but through the infusion of knowledge, it can learn about potential threats, favorable positions, or strategies that have been used in previous games. Infusing knowledge into the learning process can help disambiguate between these states, allowing agents to make more informed and accurate decisions. State infusion is illustrated in Fig. 5, where infusion occurs in between state and RL agent.

Policy Infusion The internal RL process of changing the agent's policy can be optimized by infusing external domain knowledge. This can affect the choice of actions in a way that

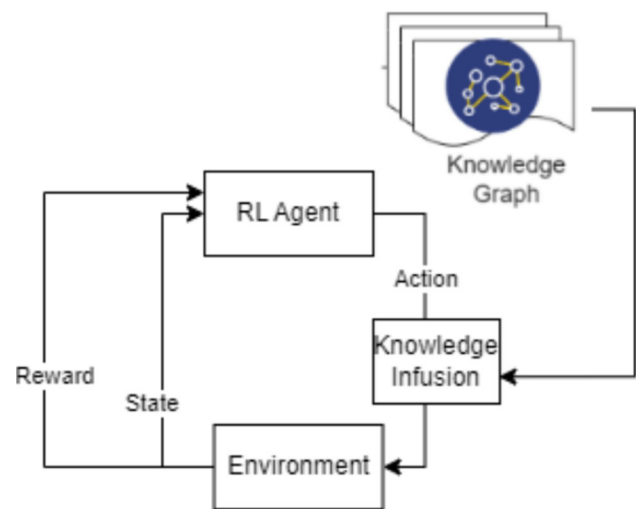


Fig. 6 Policy infusion

takes into account factors other than the environment's immediate reward input. This kind of infusion serves as an overlay or filter on the policies that the agent would otherwise pick up from its interactions with the outside world. For example, infusing knowledge into the policy function enables agents to avoid actions that oppose the organization's strategy and values. When policy infusion occurs, limitations or guidelines based on outside information are applied to the actions that the RL agent's policy suggests. This may involve prioritizing actions that are in line with domain-specific objectives or compliance standards, or it may necessitate eliminating some actions that are thought to be risky or unsatisfactory. The agent is successfully guided by external knowledge to avoid bad actions and to engage in actions that are viewed as more advantageous or acceptable in the larger context. In these situations, policy injection can stop the agent from acting in a way that could result in unfavorable or dangerous effects. Policy infusion is illustrated in Fig. 6, intermediate representation of knowledge infusion from action and environment.

4.3 Benefits of infusing knowledge into DL models

Moreover, there are several benefits of infusing knowledge into DL models. Traditional approaches such as ML and DL have shown remarkable strides in fields such as computer vision and NLP. However, there are some issues in the adoption of such techniques. For example, there is a lack of interpretability, usability, robustness, biases toward data, complexity, and high computation cost. Thus, infusing knowledge into Deep Learning (DL) models is an effective solution to overcome these challenges [6].

Improving Generalizability Knowledge-infused DL models have the potential to improve generalizability by leverag-

ing external knowledge representations, such as knowledge graphs. Therefore, allowing models to learn from input data and structured knowledge leads to enhanced generalization capabilities across diverse domains and applications.

Reduction of Bias and False Alarms By incorporating external knowledge, knowledge-infused DL models can reduce bias and false alarms. The integration of structured knowledge helps disambiguate important concepts and redefine the emphasis of essential and irrelevant terms, leading to more reliable and accurate predictions.

Enhanced Explainability Knowledge infusion in DL models can enable interpretation by providing a basis for model predictions and decisions. The inclusion of structured knowledge provides a clearer understanding of how the model uses external knowledge to make inferences and predictions.

Reliability and Stability Knowledge-infused DL models have the potential to demonstrate improved reliability and robustness. By using external knowledge bases, models benefit from a broader understanding of the domain, resulting in reliable and robust performance in real-world applications.

Reducing Big Data Reliance Traditional DL models typically require more data for training. Knowledge infusion DL models can potentially reduce reliance on big data by using external knowledge to improve the learning process. This can be especially useful in areas where big data is not readily available.

Enhanced Coverage and Disambiguation The infusion of external knowledge within latent layers of DL models can enhance the coverage of learned information. Additionally, it can aid in disambiguating important concepts defined in knowledge graphs, leading to more comprehensive and accurate representations of the underlying domain.

Improved Interpretability Knowledge-infused DL models can provide more interpretable results than traditional DL models. By merging external knowledge, the models can provide interpretations for their predictions and decisions, making them more transparent and understandable to users.

Better Handling of Rare Events Traditional DL models will struggle to determine outliers. Knowledge-infused DL models can weigh external knowledge to better comprehend and handle these rare events, leading to more accurate predictions and decisions.

More Efficient Learning Knowledge-infused DL models can learn more efficiently than traditional DL models. By consolidating external knowledge, the models can benefit from fewer training time steps, leading to faster and more efficient learning.

Improved Transfer Learning Transfer learning is where one model's learning gets transferred to another. Knowledge-infused DL models can improve transfer learning by leveraging external knowledge to transfer knowledge across related tasks more effectively.

Better Handling of Concept Drift Concept drift refers to the phenomenon where underlying information gets lost over time. Knowledge-infused DL models can better handle concept drift by leveraging external knowledge to adapt to changing concepts and maintain accuracy over time.

Improved Data Integration Knowledge-infused DL models can integrate data from multiple sources more effectively than traditional DL models. By incorporating external knowledge, the models can better understand the relationships between different data sources, leading to more accurate predictions and decisions.

Reduced Overfitting Knowledge-infused DL models may display reduced overfitting, due to the models generalize better.

Enhanced Feature Representation The infusion of knowledge can lead to enhanced feature representation, allowing the models to capture more nuanced and domain-specific information, which traditional DL models may not effectively capture.

4.4 Benefits of infusing knowledge into DRL/RL models

Infusing knowledge into RL models can significantly reduce the amount of training batch size and iteration needed to achieve high rewards, showing better results for more complex environments where decision-making by the agent requires more knowledge to adapt [37]. The experiment revealed that adding knowledge to the agent's learning process improves sample efficiency, and the benefits increase with the complexity of the environment. Adding knowledge to the agent's learning process can help it generalize and rationalize better. Traditional RL relies solely on state and reward information, meaning that the agent has to learn from scratch in every situation it encounters. This is time-consuming in the training phase regardless of the RL algorithm in use. By contrast, infusing knowledge provides additional context that can be effective in various situations. This can help the agent to make more accurate predictions and make better choices, leading to better performance in a wide range of situations.

Additionally, knowledge infusion can also help agents to learn more robustly. Traditional RL can be very sensitive to slight variations in the environment or the task. Infusing knowledge can help agents adapt more easily to environmental changes, reducing the risk of overfitting and making them more robust to unexpected situations.

4.5 Evaluation metrics for knowledge-infusion Deep Learning

Specific metrics need to be considered to evaluate the performance of DL models when knowledge is infused. There are a few possible metrics for comparing knowledge infu-

Table 2 Pros and Cons of Different Metrics in Deep Learning Models

Metric	Pros	Cons
Model Accuracy	Direct measure of performance Easy to understand and calculate	Can be misleading for imbalanced datasets Does not account for the type of errors (false positives/negatives)
Computational Efficiency	Effective for balanced datasets Saves time and resources	More efficient models may sacrifice accuracy or complexity
Interpretability	Essential for deploying models in resource-limited environments Enhances trust and usability Important for sensitive applications (e.g., healthcare) Helps in debugging and improving models	May not be suitable for complex tasks requiring deeper models Can be challenging to achieve in complex models May require additional techniques or tools

sion techniques, including model accuracy, computational efficiency, generalization to new data, interpretability, and flexibility to different types of knowledge. Another approach to evaluating the effectiveness of knowledge-infused DL is to use traditional DL architectures without expert knowledge to compare the performance of each technique side by side.

Model Accuracy This metric measures the correctness of the model's predictive outcome or the trueness of a classification. The higher the score of this metric, the better the model's performance. Suppose we want to make a spam or ham (not spam) classification DL model for emails. The email text is taken as an input to classify whether a text is spam or ham. Assume we have a dataset of 1000 labeled emails, with 700 ham and 300 spam classifications. We split the dataset into a training set of 800 emails and a test set of 200 emails. We train or DL to maximize the accuracy of the spam classification. After the training, we obtained these results: Out of 200 emails in the test set, the model correctly classified 180 emails as ham and 50 as spam. The overall accuracy of the model is $(180+50)/200 = 0.65$ or 65%.

Computational Efficiency It is mostly refers to the measurement of time and efficiency of the computing resources but guided by the time it takes to train and run the model. It is evaluated by measuring the training time, memory usage, and processing speed of the model. Lower consumption of computational resources equates to a more efficient model. Assuming we choose a DL problem to solve using neural networks, we can employ two models trying to solve the same problem. Model A is a deep neural network with 10 layers and 6 million trainable parameters. Model B is a shallow neural network with 2 layers and 200,000 trainable parameters. Model A takes 20h to train with a maximum memory usage of 12 GB. Model B takes 1h to train with a maximum memory usage of 1 GB. This shows that Model B, while it

provides 200k parameters, performs better as far as computational efficiency goes.

Interpretability This metric refers to the ability to understand the decision-making process of the model, particularly how it incorporates and uses external knowledge. Suppose we have trained a deep learning model to classify different types of skin diseases based on the skin lesion images. To enhance the interpretability of this model and make it more transparent, [38] employs a technique called "Grad-CAM" to visualize and highlight important regions of the input images responsible for the model's prediction. Grad-CAM produces results such as making comparisons on why the model is classified as benign given the shape of the mole.

In DL, evaluating a model's performance is more than just accuracy. It is crucial to consider a spectrum of metrics, each offering a solution for finding the applicability and functionality of the model. The aforementioned metrics have pros and cons, Table 2 considers some of the different aspects of the metrics. Furthermore, we will compare shallow, semi-deep, and deep infusion on a high level and discuss the pros and cons of these infusion methods in Table 4.

Similarly to the DL approach, it is important to consider specific metrics to evaluate the performance of DRL models when knowledge is infused. In this context, important metrics that can be used include:

Cumulative Reward Indicates the total reward gained by the agent over time. Knowledge infusion should ideally lead to an increase in cumulative reward and decision-making.

Convergence Rate The velocity at which the RL agent's policy converges to an optimal or near-optimal solution. Knowledge infusion should ideally reduce the time or number of episodes required for convergence.

Sample Efficiency This measures how effectively the RL agent can learn from a limited number of interactions with the

Table 3 Comparison of Metrics for Knowledge Infusion Techniques in Reinforcement Learning

Metric	Pros	Cons
Cumulative Reward	<ul style="list-style-type: none"> Direct measure of agent performance Easy to quantify and compare Aligns well with the goal-oriented nature of RL 	<ul style="list-style-type: none"> May not reflect short-term learning dynamics Can be influenced by outliers or extreme rewards
Convergence Rate	<ul style="list-style-type: none"> Reflects efficiency of learning process Important for time-sensitive applications Can indicate stability and reliability of learning 	<ul style="list-style-type: none"> Harder to measure in environments with high variance Does not always correlate with optimal policy performance
Sample Efficiency	<ul style="list-style-type: none"> Critical in environments where data is scarce or expensive to obtain Reflects practical usability in real-world scenarios Can lead to faster deployment of RL models 	<ul style="list-style-type: none"> May encourage overfitting to limited data Difficult to evaluate in complex environments
Policy Robustness	<ul style="list-style-type: none"> Essential for reliability in dynamic or uncertain environments Indicates adaptability to new situations Enhances trust in automated decision-making 	<ul style="list-style-type: none"> Challenging to quantify and test May require extensive simulation or testing in varied scenarios
Exploration vs. Exploitation Balance	<ul style="list-style-type: none"> Critical for comprehensive learning in RL Ensures both short-term gains and long-term knowledge acquisition Can lead to more adaptive strategies 	<ul style="list-style-type: none"> Difficult to find the right balance May depend heavily on the specific environment and task
Computational Efficiency	<ul style="list-style-type: none"> Important for scalability and practical application Reduces resource usage and costs Enables application in resource-constrained scenarios 	<ul style="list-style-type: none"> High computational efficiency may compromise learning quality Not always a primary concern in theoretical research
Interpretability of the Policy	<ul style="list-style-type: none"> Crucial for transparency and trust in AI systems Facilitates debugging and improvement of RL models Important in regulated industries 	<ul style="list-style-type: none"> Can be at odds with model complexity and performance Challenging to achieve in deep learning-based RL
Scalability	<ul style="list-style-type: none"> Essential for applying RL to real-world, complex problems Demonstrates the adaptability of the knowledge infusion technique Allows for broader application and generalization 	<ul style="list-style-type: none"> Scalability can be limited by computational resources More complex environments increase the risk of overfitting

Table 4 Comparison of Shallow, Semi-Deep, and Deep Infusion Learning

Aspect	Shallow Infusion	Semi-Deep Infusion	Deep Infusion
Integration Level	Surface-level integration, often limited to data preprocessing or basic feature engineering steps	Intermediate level integration, typically at specific layers or modules within the model, combining both data and knowledge to enhance learning	Fundamental integration within the core architecture of the model, deeply embedding knowledge into the learning process at multiple stages
Model Architecture	Standard machine learning or deep learning models with no changes to their original structure	Modified architecture that includes specific enhancements or modules designed to integrate external knowledge	Highly specialized and complex architectures designed to seamlessly integrate and utilize external knowledge throughout the entire learning process
Learning Process	Primarily driven by data, with external knowledge used to improve feature quality or provide additional context	Combines learning from both data and strategically placed knowledge, allowing for enhanced learning at certain points	A unified learning process where data and deeply integrated knowledge work together continuously, leading to a more holistic understanding and improved performance
Pros	Easy to implement, minimal changes to existing workflows, low complexity, and quick to deploy	Balanced approach offering a good compromise between performance improvement and added complexity, with noticeable enhancements in model performance	Yields highly sophisticated models that provide deep insights, superior performance, and a more thorough understanding of the data and the problem domain
Cons	Offers limited enhancement capabilities, often leading to underutilization of available knowledge, may not significantly improve complex tasks	Requires careful and thoughtful design to effectively integrate knowledge without overwhelming the model, adding a moderate level of complexity	Very high complexity, demanding significant expertise in both model architecture and knowledge integration, can be resource-intensive in terms of computational power and time

environment. Effective knowledge infusion allows an agent to learn using fewer samples.

Policy Robustness This assesses how well the learned policy performs under varying conditions and uncertainties in the environment. Knowledge infusion should ideally lead to policies that are more robust to changes and noise.

Exploration vs. Exploitation Tradeoff Infused knowledge can impact how an agent explores the environment versus exploiting known good strategies. This metric evaluates the effectiveness of this balance, which is crucial for efficient learning in RL.

Computational Efficiency This includes the computational resources required (e.g., time and memory) for the learning process. Effective knowledge infusion techniques should not significantly increase the computational burden.

Interpretability of the Policy The degree to which the decision-making process of the RL agent is understandable, especially when knowledge is infused. This is increasingly important in applications where understanding the basis of decisions is crucial.

Scalability How well the knowledge infusion technique can be applied to larger or more complex problems. This is particularly relevant in real-world applications where state and action spaces can be very large.

Table 3 presents the high-level pros and cons for the metrics provided in the previous sections.

5 Knowledge-infused learning and automated threat defense

In this section, we introduce the application areas of KIL in AI models, particularly within the cybersecurity domain. We explore how infusing domain knowledge can enhance explainability, making AI systems more transparent and understandable. Additionally, we delve into the use of LLMs in knowledge-infused cybersecurity learning, highlighting their potential to improve decision-making and operational efficiency.

5.1 Application areas

The infusion of knowledge can support RL and DL methods in multiple ways. This Section presents a discussion regarding the use of domain knowledge to solve Cybersecurity problems. Will delve into applications of knowledge infused learning in the domain of Cybersecurity such as Red Teaming and Blue Teaming.

5.1.1 Red teaming

RL plays an important role in the automation of cybersecurity operations. Red team activities concern penetration testing, a common method for assessing the security of the system,

Table 5 Summary of ML methodologies in cybersecurity applications categorized by Red Teaming and Blue Teaming

Citation	ML Methodology	Application/Goal	Technique
Red Teaming			
Nhu et al. [39], Zhang et al. [40]	Deep RL (A3C), Deep RL (Adv. Actor-Critic)	Automated penetration testing, Network security via IP address mutation	Reconnaissance
Zhong et al. [41], Arif et al. [42]	GAN, GAN + Deep RL	Generating adversarial malware, Evasion of malware detectors	Resource Development
Venkatesan et al. [43], Kujanpaa et al. [44]	RL, Deep RL	Detecting and mitigating botnets, Automated privilege escalation	Initial Access, Privilege Escalation
Maeda et al. [45], Wang et al. [46]	Deep RL (A2C), RL	Post-exploitation automation, Software obfuscation	Discovery, Command & Control
Cody et al. [47], Rishu et al. [48]	RL	Optimal data exfiltration paths, Realistic emulation	Exfiltration
Blue Teaming			
Hore et al. [49], Sheng et al. [50]	Deep RL + Integer Programming, Deep RL (Pointer Network)	Cyber vulnerability management, Vulnerability prioritization	Identify
Sewak et al. [51], Liang et al. [52]	Deep RL, Deep RL	Threat detection and endpoint protection, Detection of IP watermarking	Protect
An et al. [53], Arshad et al. [54]	Deep RL (Deep-Q-Network), Deep RL	Defense against data integrity attacks, Anomaly detection	Detect
Maliais et al. [55], Hughes et al. [56]	RL, Deep RL	Optimizing intrusion detection, Customizing intrusion response systems	Respond
Prasad et al. [57], Wei et al. [58]	RL, Deep RL	Recovery in cyber-manufacturing attacks, Optimal recovery of transmission lines	Recover

program, or network. Although there are many tools to automate the process, penetration testing is often done manually and relies on the experience of the ethical hacker. The areas of RL application in red teaming include:

Reconnaissance In any attack against any system, program, network, or infrastructure, the attacker must gather intel on how to break into the system. In the reconnaissance stage, the attacker collects information about the victim to prepare and set the way for the penetration of the target. In Nhu et al. [39], the authors present an automated penetration testing approach that utilizes deep reinforcement learning to automate the process. This effort includes techniques such as reconnaissance and exploitation and trains the RL agent following the Asynchronous Advantage Actor Critic (A3C) model to learn to execute exploits automatically and accurately identify vulnerabilities in the environment. A benchmark is created to compete against real-world vulnerabilities in its experimental environment. This contribution introduces a tool that performs information gathering, vulnerability exploitation, and reporting tasks. Furthermore, the tool can accumulate learned results from previous environments to successfully exploit vulnerabilities for the next exploit in different environments. Zhang et al. [40] propose an Intelligence-Driven Host Address Mutation (ID-HAM) scheme to improve network security. ID-HAM works by mutating host IP addresses in response to adversarial scanning in a way that optimizes the security and survivability of connections. This research applies deep reinforcement learning to develop an advantage actor-critic algorithm for HAM that models the mutation process as a Markov decision process. The security analysis and extensive simulations show that ID-HAM is more effective than existing solutions in reducing scanning hits without significantly affecting communication.

Resource Development After conducting target reconnaissance and gathering all necessary data, attackers move to the next phase, which relies on executing a more targeted attack. The Resource Development (RD) tactic entails designing attacks with consideration for the targeted system's constraints. Zhong et al. [41] discuss the use of RL for generating adversarial examples of malware that can bypass black-box detectors. The authors propose a novel framework that uses a Generative Adversarial Network (GAN) to generate such examples. They evaluate their framework using three commonly used detection systems and report high success rates in generating false negatives that can evade detection. The authors also discuss the limitations and future directions of their work, including the need for more extensive evaluation using different types of malware and detection systems and the importance of exploring multi-agent reinforcement learning for generating adversarial examples that can overcome more robust detection systems. Overall, the paper provides a valuable contribution to adversarial machine

learning and cybersecurity. The effort presented in [42] describes an evasion framework named IF-MalEvade, which utilizes Generative Adversarial Network (GAN) and Deep Reinforcement Learning (DRL) to generate fully working malware samples with several effective perturbations such as header section manipulation and benign bytes insertion. The DRL framework selects a few suitable action sequences to change malicious samples, thus allowing the malware samples to bypass various black-box ML-based malware detectors and the detection search engines of VirusTotal while maintaining the executability and malicious behavior of the original malware samples.

Initial Access (IA), Execution (Ex), and Persistence (Ps) Using the available resources to gain unauthorized access to the victim's systems is the foundation of Initial Access (IA). This strategy uses weaknesses based on many strategies that are limited by the system. The Execution (Ex) technique then represents the attack phase. Attacks can target many technologies and elements of secure operation. Subsequently, the Persistence (Ps) strategy determines the capabilities for prolonged system access. A proposed approach [43] for detecting and mitigating the risk of stealthy botnets in resource-constrained environments. The approach involves using reinforcement learning to optimally deploy defensive mechanisms, such as honeypots and network-based detectors, in the target network to reduce the lifetime of stealthy botnets by maximizing the number of bots identified and taken down. The authors provide a proof-of-concept for this approach and study its performance in a simulated environment.

Privilege Escalation (PE), Defense Evasion (DE), and Credential Access (CA) Kujanpaa et al. [44], the authors discuss the potential threat of using deep reinforcement learning to train automated agents for performing fully automated attacks. The authors present an agent that can escalate privileges in a Windows 7 environment and can perform a wide variety of different techniques.

Discovery (Ds), Lateral Movement (LM), and Collection (CI) Maeda and Mimura [45] discuss a method for automating post-exploitation in information systems using deep reinforcement learning and the PowerShell Empire. The method involves reinforcement learning agents selecting modules as actions and the state of the agents being defined by parameters such as the type of account that was compromised. The learning progress of three reinforcement learning models is compared, and the A2C is shown to be the most efficient. The trained agent using A2C can obtain administrative privileges to the domain controller.

Command and Control (C2) Wang et al. [46] discusses the challenge of effectively obfuscating software to prevent reverse engineering and other forms of attack. The authors propose a technique for generating an optimal program obfuscation scheme through reinforcement learning.

The method involves a reinforcement learning model that selects a sequence of obfuscation transformations oriented toward producing an optimal obfuscation result while retaining reasonable instrumentation overhead. The evaluation demonstrates that the models can effectively obfuscate executable files at a low cost.

Exfiltration (Ef) and Impact (Im) Cody et al. [47] explore how reinforcement learning, in combination with attack graphs and cyber terrain, can be used to identify optimal paths for data exfiltration from enterprise networks. The authors describe how exfiltration can be a challenging issue for enterprises due to the difficult detection of encapsulated data and the targeted network paths preferred by adversaries for data theft reduction. The document outlines how RL can be combined with attack graphs to develop a reward function capable of detecting paths with minimal detection. Additionally, the paper outlines the potential for the technique in large network environments. A new method for discovering exfiltration paths within computer networks using reinforcement learning is proposed in [48]. The authors aim to account for nuances in adversarial behavior and include communication payload and protocol into the Markov decision process to emulate attributes of network-based exfiltration events more realistically. By doing so, practitioners can detect expected adversary behavior under various payload and protocol assumptions.

Discussion for domain knowledge-infused learning for red teaming

Integrating domain knowledge into RL and DL for red teaming in cybersecurity significantly enhances the accuracy, efficiency, and effectiveness of these simulations. Red teaming involves simulating cyber attacks to identify and address vulnerabilities, and incorporating domain knowledge offers substantial benefits. By embedding knowledge about common attack patterns, known vulnerabilities, and typical defense mechanisms, RL agents can be trained to perform more realistic and sophisticated attacks. This enables more thorough testing and helps organizations better prepare for actual cyber threats. In the reconnaissance phase, domain knowledge can provide RL agents with detailed information about potential targets, enhancing information gathering and identifying entry points more accurately. During resource development, domain knowledge informs the creation of more effective attack tools and techniques by understanding the specific constraints and defenses of a target system. For initial access, execution, and persistence, domain knowledge guides RL models in developing more effective and stealthy attack vectors by understanding common entry points, execution techniques, and persistence mechanisms. Additionally, domain knowledge assists RL agents in escalating privileges, evading defenses, and accessing credentials by providing

insights into common strategies and weaknesses. Overall, the infusion of domain knowledge into RL and DL models for red teaming enhances the ability to simulate realistic cyber attacks, ultimately improving the cybersecurity posture of organizations.

5.1.2 Blue teaming

Moreover, Blue team activities refer to the defensive cybersecurity measures an organization undertakes to protect its assets, including networks, systems, and data. Blue team employs various tools and techniques to prevent and detect cyber attacks, such as Security Information and Event Management (SIEM), Intrusion Detection Systems (IDS), firewalls, and vulnerability scanners. Blue team also conducts regular security assessments, risk analysis, and incident response drills to ensure that they are prepared to respond to any potential security incidents. RL can be used in the automation of Blue teaming activities in multiple ways, such as:

Identify Hore et al. [49], the authors propose a framework called “Deep VULMAN” for cyber vulnerability management in a cybersecurity operations center (CSOC). The framework uses a deep reinforcement learning agent and an integer programming method to prioritize and select vulnerabilities for mitigation, considering future uncertainties and adjusting the response to fluctuations in vulnerability arrivals. The authors in [50] discuss various factors regarding vulnerability and recognize that the essence of vulnerability prioritization is a combinatorial optimization problem. When organizations consider economic concerns, business objectives, security, and other factors, vulnerability prioritization can become a multi-objective optimization problem. Although corporate risk estimation procedures differ depending on the needs, we may still utilize a pointer neural network to produce a prioritized vulnerability remediation plan if the input format and reward function are appropriately configured. Since its inception, the pointer network has been extremely successful in combinatorial optimization and natural language processing. In particular, it outperforms several other neural networks in the NLP summary challenge. Thus, employing pointer networks for vulnerability prioritization is a huge step in the right direction, and researchers will continue to supplement and improve VPnet in future work. VPnet, a vulnerability prioritization approach, employs a pointer network and deep reinforcement learning to produce almost optimum solutions in seconds under diverse scale scenarios and limitations, resulting in a 22.8% performance gain in a practical example. The authors also introduce a new training strategy that combines imitation and autonomous learning to increase model training performance. They also use a simple vulnerability risk calculation model considering severity, danger, impact, and asset criticality. As measuring

risk is vital in this context, this paper recognizes that risk is a combination of the “threat” faced by the target system (the threat actor’s ability and intention), the system’s “vulnerability” (weakness or exposure), and the “impact” of successful exploitation of vulnerability on the organization.

Protect Sewak et al. [51] review the applications of deep reinforcement learning in cybersecurity, specifically in threat detection and endpoint protection. The research paper focuses on the different applications of Deep Reinforcement Learning (DRL) in various aspects of threat detection and protection, particularly in MDR, IDS, and EDR systems. It also provides a brief introduction to the diverse types of DRL techniques. The authors also discuss the need for robust threat detection and response systems and introduce DRL and how it has emerged in cybersecurity. Liang et al. [52] propose a fast deep-reinforcement-learning (DRL)-based detection algorithm for virtual IP watermarks by combining the technologies of mapping function and DRL to preprocess the ownership information of the IP circuit resource. The proposed algorithm aims to detect copyright infringement of intellectual property (IP) circuit resources of electronic devices. In [53], a deep-Q-network detection (DQND) scheme is proposed to defend against data integrity attacks in AC power systems and avoid the problem of curse of dimension in conventional reinforcement learning schemes.

Detect Arshad et al. [54] conduct a systematic literature review on the use of deep reinforcement learning techniques for anomaly detection, including various DRL models and their performance comparison with alternative techniques, applications of anomaly detection, and representative anomaly datasets used in research articles from 2017–2022. Similarly, Sewak et al. [59] presents a thorough examination of the various uses of deep reinforcement learning in advanced cybersecurity threat detection and mitigation. This investigation considers the overly complicated cybersecurity threat landscape and the necessity to transition from traditional to sophisticated deep and machine learning defense mechanisms for threat detection and protection. The research then delves into Deep Reinforcement Learning (DRL), which has shown considerable promise in building AI solutions for domains that previously needed advanced human cognition. The study describes numerous strategies and algorithms under DRL that have produced great results in various sectors. The study explains how DRL is a more diversified and empowering technique than supervised and deep learning. In the introduction, the authors also mention the need for a comprehensive review of DRL applications in advanced cybersecurity threat detection and protection, which they aim to fill in this paper. The paper then discusses various DRL applications in the field of cybersecurity, such as Network IDS, Endpoint detection, Advanced threat protection, IoT defense, and 5G jamming. Finally, the paper highlights the

importance and benefits of using DRL in advanced cybersecurity threat detection and protection.

Respond Malialis and Kudenko [55] focus on the scalability of intrusion detection systems with machine learning. The authors propose a novel design to address the “curse of dimensionality” issue when scaling up. The paper discusses the application of distributed reinforcement learning in optimizing communication costs of information exchange among agents, as well as various intrusion response techniques such as Probabilistic Packet Marking and replication techniques. Overall, the paper presents several contributions and solutions related to intrusion detection and response. Hughes et al. [56] discusses how Intrusion Response Systems (IRS) are an active area of research that seeks to automatically respond to alerts generated by Intrusion Detection Systems on computer networks. The article proposes utilizing a Deep Reinforcement Learning approach to facilitate the creation of Response Profiles that can align with a company’s Incident Response Policies. These Response Profiles allow for the customization of Reward Functions to suit differing Incident Response Policies, such as in Cyber-Physical networks where different areas have various levels of policy, for example, more stringent policies may be required to preserve physical processes.

Recover Prasad et al. [57] propose a four-layer recovery architecture to ensure manufacturing systems achieve operational goals during a cyber-manufacturing attack. The area consists of a systems layer, attack identification layer, data auditing and detection layer, and RL-based recovery layer. The simulation of the conveyor system is constructed inside the Unity3D physics platform. Wei et al. [58] propose a recovery strategy for reclosing tripped transmission lines at an optimal reclosing time to minimize the cyber-attack impact on power systems. The proposed strategy is based on a deep Reinforcement Learning (RL) framework, established to simulate the power system dynamics during the attack-recovery process and generate the training data.

Discussion for domain knowledge-infused learning for blue teaming

Integrating domain knowledge into reinforcement learning (RL) and deep learning (DL) significantly enhances blue team activities in cybersecurity by improving the prevention, detection, response, and recovery from cyber attacks. Domain knowledge provides detailed insights into common attack patterns, known vulnerabilities, and typical defense mechanisms, enabling RL and DL models to perform more effectively. In the identification phase, domain knowledge helps prioritize and select vulnerabilities for mitigation, as demonstrated by the “Deep VULMAN” and VPnet frameworks, which optimize vulnerability management using deep RL and pointer networks. For protection, domain knowledge

improves threat detection and endpoint protection, exemplified by applications in IP watermarking and data integrity defense. It enhances anomaly detection by guiding RL agents to recognize and address unusual patterns accurately. In the response phase, domain knowledge optimizes intrusion response systems by creating customized response profiles aligned with incident response policies. Lastly, in the recovery phase, domain knowledge supports developing effective recovery strategies to minimize the impact of cyber-attacks, such as RL-based recovery architectures for manufacturing systems and power grids. Overall, embedding domain knowledge into RL and DL models makes blue team defense mechanisms more robust, context-aware, and adaptive to evolving threats, thereby significantly improving an organization's cybersecurity posture.

6 Open challenges

Despite the potential of KIL to revolutionize ATD, several research gaps and open challenges must be addressed to ensure its effective implementation. These include the integration with legacy systems, scalability and efficiency issues, and privacy and ethical considerations. Addressing these challenges is crucial for advancing KIL's application in ATD and enhancing its robustness, transparency, and ethical use.

6.1 Integration with legacy systems and technologies

When KIL and ATD integrate together into legacy systems, challenges will arise due to the outdated nature of these infrastructures. These systems often lack the degree of flexibility to support contemporary AI-driven solutions. It necessitates costly and time-consuming upgrades or modernization. Such systems' prototypes and data formats may be incompatible with today's AI methodologies, offering comprehensive adjustments to guarantee compatibility and robust security. In addition, data integration poses major challenges, as legacy systems manage data in formats not easily compatible with KIL and ATD algorithms. Despite these challenges, integrating KIL and ATD can strengthen cybersecurity by leveraging advanced AI and domain knowledge, although it requires substantial investment and effort to overcome technical and financial barriers.

6.2 Scalability and efficiency of knowledge infusion

Scalability and efficiency are crucial challenges for KIL in ATD. As cyber threats evolve, KIL systems must scale to handle larger datasets while maintaining processing speed and efficiency. Managing increasing data volumes from diverse

sources requires robust data management and scalable computing resources. Ensuring computational efficiency involves optimizing algorithms, using parallel computing, and leveraging cloud technologies. Infrastructure limitations and the need for continuous system monitoring and fail-safes are also significant concerns. Effective resource management and cost control, often through elastic cloud services, are essential to ensure that KIL systems can scale dynamically and efficiently.

6.3 Privacy and ethical considerations

Implementing KIL in cybersecurity raises significant privacy and ethical concerns. Ensuring compliance with data protection regulations such as GDPR [60] and HIPAA [61] is crucial, especially when handling sensitive personal information. KIL systems must be designed with robust privacy protections to prevent unauthorized access. Ethical frameworks are needed to guide the use of AI in cybersecurity, addressing issues such as algorithmic transparency, accountability, and avoiding unjustified surveillance. Ensuring fairness and eliminating bias in KIL systems require continuous attention, including designing objective algorithms and training AI on diverse datasets.

7 Conclusion

The current diversity in network technologies and large-scale architectures enable sharing information, services, and conducting business to meet society's demands. The complexity and profound impact cyberattacks have on business operations establish a major need for the development of advanced countermeasures to automate cybersecurity activities. Although the adoption of DL and DRL has demonstrated remarkable success in protecting cyber assets, these approaches present critical limitations that prevent their wide deployment. This paper presented a study regarding the use of Knowledge Infused Learning (KIL) in the context of automated threat defense. We investigated how cybersecurity domain knowledge can be infused into DL and RL and presented the main obstacles and benefits. This analysis comprises red and blue teaming activities. The state-of-the-art solutions introduce insights into how to develop automated threat defense solutions but still unveil critical research lines to be considered in the next few years. Finally, the future directions of this research include the numerical analysis of each infusion approach in DL and DRL for different cybersecurity problems. By addressing these challenges we can overcome more adaptive cyberattacks in the context of automated threat defense systems.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Selby, N., Vescent, H.: The cyber attack survival manual: tools for surviving everything from identity theft to the digital apocalypse. Weldon Owen International (2017)
- Wilson, D.C.: Cybersecurity. MIT Press (2021)
- Li, J.-H.: Cyber security meets artificial intelligence: a survey. *Front. Inf. Technol. Electron. Eng.* **19**(12), 1462–1474 (2018)
- Piplai, A., Ranade, P., Kotal, A., Mittal, S., Narayanan, S.N., Joshi, A.: Using knowledge graphs and reinforcement learning for malware analysis. In: 2020 IEEE International Conference on Big Data (Big Data), pp. 2626–2633. IEEE (2020)
- Moreno, D.L., Regueiro, C.V., Iglesias, R., Barro, S.: Using prior knowledge to improve reinforcement learning in mobile robotics. In: Proceedings of the Towards Autonomous Robotics Systems. Univ. of Essex, UK, 33 (2004)
- Kursuncu, U., Gaur, M., Sheth, A.: Knowledge infused learning (k-il): Towards deep incorporation of knowledge in deep learning. arXiv preprint [arXiv:1912.00512](https://arxiv.org/abs/1912.00512) (2019)
- Sheth, A., Gaur, M., Kursuncu, U., Wickramarachchi, R.: Shades of knowledge-infused learning for enhancing deep learning. *IEEE Internet Comput.* **23**(6), 54–63 (2019)
- Gaur, M., Gunaratna, K., Bhatt, S., Sheth, A.: Knowledge-infused learning: A sweet spot in neuro-symbolic AI. *IEEE Internet Comput.* **26**(4), 5–11 (2022)
- Li, X., Lyu, M., Wang, Z., Chen, C.-H., Zheng, P.: Exploiting knowledge graphs in industrial products and services: a survey of key aspects, challenges, and future perspectives. *Comput. Ind.* **129**, 103449 (2021). <https://doi.org/10.1016/j.compind.2021.103449>
- Hu, L., Liu, Z., Hu, W., Wang, Y., Tan, J., Wu, F.: Petri-net-based dynamic scheduling of flexible manufacturing system via deep reinforcement learning with graph convolutional network. *J. Manuf. Syst.* **55**, 1–14 (2020). <https://doi.org/10.1016/j.jmsy.2020.02.004>
- Inoue, T., De Magistris, G., Munawar, A., Yokoya, T., Tachibana, R.: Deep reinforcement learning for high precision assembly tasks. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 819–825 (2017). <https://doi.org/10.1109/IROS.2017.8202244>
- Schoettler, G., Nair, A., Ojea, J.A., Levine, S., Solowjow, E.: Meta-reinforcement learning for robotic industrial insertion tasks. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 9728–9735 (2020). <https://doi.org/10.1109/IROS45743.2020.9340848>
- Zhou, S., Dai, X., Chen, H., Zhang, W., Ren, K., Tang, R., He, X., Yu, Y.: Interactive recommender system via knowledge graph-enhanced reinforcement learning. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 179–188 (2020)
- Kim, S.W., Kim, I., Lee, J., Lee, S.: Knowledge integration into deep learning in dynamical systems: an overview and taxonomy. *J. Mech. Sci. Technol.* **35**, 1331–1342 (2021)
- Futia, G., Vetrò, A.: On the integration of knowledge graphs into deep learning models for a more comprehensible ai-three challenges for future research. *Information* **11**(2), 122 (2020)
- Dash, T., Chitlangia, S., Ahuja, A., Srinivasan, A.: A review of some techniques for inclusion of domain-knowledge into deep neural networks. *Sci. Rep.* **12**(1), 1040 (2022)
- Xie, X., Niu, J., Liu, X., Chen, Z., Tang, S., Yu, S.: A survey on incorporating domain knowledge into deep learning for medical image analysis. *Med. Image Anal.* **69**, 101985 (2021)
- Radanliev, P., De Roure, D., Maple, C., Nurse, J.R., Nicolescu, R., Ani, U.: Ai security and cyber risk in IoT systems. *Front. Big Data* **7**, 1402745 (2024)
- Wang, Z., Liu, D., Sun, Y., Pang, X., Sun, P., Lin, F., Lui, J.C., Ren, K.: A survey on IoT-enabled home automation systems: attacks and defenses. *IEEE Commun. Surv. Tutorials* **24**(4), 2292–2328 (2022)
- Qabajeh, I., Thabtah, F., Chiclana, F.: A recent review of conventional vs. automated cybersecurity anti-phishing techniques. *Comput. Sci. Rev.* **29**, 44–55 (2018)
- Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press (2016)
- Canziani, A., Paszke, A., Culurciello, E.: An analysis of deep neural network models for practical applications. arXiv preprint [arXiv:1605.07678](https://arxiv.org/abs/1605.07678) (2016)
- Li, Y.: Deep Reinforcement Learning: An Overview. arXiv preprint [arXiv:1701.07274](https://arxiv.org/abs/1701.07274) (2017)
- Wiering, M.A., Van Otterlo, M.: Reinforcement learning. *Adapt. Learn. Optim.* **12**(3), 729 (2012)
- Huang, Y.: Deep q-networks. *Deep Reinforcement Learning: Fundamentals, Research and Applications*, 135–160 (2020)
- Sutton, R.S., McAllester, D., Singh, S., Mansour, Y.: Policy gradient methods for reinforcement learning with function approximation. *Adv. Neural Inf. Process. Syst.* **12** (1999)
- Zhang, J., Kim, J., O'Donoghue, B., Boyd, S.: Sample efficient reinforcement learning with reinforce. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, pp. 10887–10895 (2021)
- Xin, X., Karatzoglou, A., Arapakis, I., Jose, J.M.: Supervised advantage actor-critic for recommender systems. In: Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, pp. 1186–1196 (2022)
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O.: Proximal policy optimization algorithms. arXiv preprint [arXiv:1707.06347](https://arxiv.org/abs/1707.06347) (2017)
- Barto, A.G., Sutton, R.S., Anderson, C.W.: Looking back on the actor-critic architecture. *IEEE Trans. Syst. Man Cybern.: Syst.* **51**(1), 40–50 (2020)
- Pong, V., Gu, S., Dalal, M., Levine, S.: Temporal difference models: Model-free deep rl for model-based control. arXiv preprint [arXiv:1802.09081](https://arxiv.org/abs/1802.09081) (2018)
- Huang, K., Madnick, S., Johnson, S.: Interactions between cybersecurity and international trade: a systematic framework (2018)
- Caverty, M.D.: The Politics of Cyber-Security. Taylor & Francis (2024)
- Li, C., Qiu, M.: Reinforcement learning for cyber-physical systems: with cybersecurity case studies. Chapman and Hall/CRC (2019)
- Ahlemeyer-Stubbe, A., Müller, A.: Why domain knowledge is essential for data scientists in marketing. *Appl. Market. Anal.* **7**(4), 362–373 (2022)
- Xenou, K., Chalkiadakis, G., Afantenos, S.: Deep reinforcement learning in strategic board game environments. In: Multi-Agent Systems: 16th European Conference, EUMAS 2018, Bergen,

- Norway, December 6–7, 2018, Revised Selected Papers 16, pp. 233–248. Springer (2019)
37. Wardenga, R., Kovriguina, L., Pliukhin, D., Radyush, D., Smoliakov, I., Xue, Y., Müller, H., Pismerov, A., Mouromtsev, D., Kudenko, D.: Knowledge graph injection for reinforcement learning (2023)
 38. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-cam: Visual explanations from deep networks via gradient-based localization. *Int. J. Comput. Vision* **128**(2), 336–359 (2019). <https://doi.org/10.1007/s11263-019-01228-7>
 39. Nhu, N.X., Nghia, T.T., Quyen, N.H., Pham, V.-H., Duy, P.T., et al.: Leveraging deep reinforcement learning for automating penetration testing in reconnaissance and exploitation phase. In: 2022 RIVF International Conference on Computing and Communication Technologies (RIVF), pp. 41–46 (2022). IEEE
 40. Zhang, T., Xu, C., Shen, J., Kuang, X., Grieco, L.A.: How to disturb network reconnaissance: A moving target defense approach based on deep reinforcement learning. *IEEE Trans. Inf. Forens. Secur.* (2023)
 41. Zhong, F., Hu, P., Zhang, G., Li, H., Cheng, X.: Reinforcement learning based adversarial malware example generation against black-box detectors. *Comput. Secur.* **121**, 102869 (2022)
 42. Arif, R.M., Aslam, M., Al-Otaibi, S., Martinez-Enriquez, A.M., Saba, T., Bahaj, S.A., Rehman, A.: A deep reinforcement learning framework to evade black-box machine learning based IoT malware detectors using GAN-generated influential features. *IEEE Access* **11**, 133717–133729 (2023)
 43. Venkatesan, S., Albanese, M., Shah, A., Ganesan, R., Jajodia, S.: Detecting stealthy botnets in a resource-constrained environment using reinforcement learning. In: Proceedings of the 2017 Workshop on Moving Target Defense, pp. 75–85 (2017)
 44. Kujanpää, K., Victor, W., Ilin, A.: Automating privilege escalation with deep reinforcement learning. In: Proceedings of the 14th ACM Workshop on Artificial Intelligence and Security, pp. 157–168 (2021)
 45. Maeda, R., Mimura, M.: Automating post-exploitation with deep reinforcement learning. *Comput. Secur.* **100**, 102108 (2021)
 46. Wang, H., Wang, S., Xu, D., Zhang, X., Liu, X.: Generating effective software obfuscation sequences with reinforcement learning. *IEEE Trans. Dependable Secure Comput.* **19**(3), 1900–1917 (2020)
 47. Cody, T., Rahman, A., Redino, C., Huang, L., Clark, R., Kakkar, A., Kushwaha, D., Park, P., Beling, P., Bowen, E.: Discovering exfiltration paths using reinforcement learning with attack graphs. In: 2022 IEEE Conference on Dependable and Secure Computing (DSC), pp. 1–8 (2022). IEEE
 48. Rishu, R., Kakkar, A., Wang, C., Rahman, A., Redino, C., Nandakumar, D., Cody, T., Clark, R., Radke, D., Bowen, E.: Enhancing exfiltration path analysis using reinforcement learning. *arXiv preprint arXiv:2310.03667* (2023)
 49. Hore, S., Shah, A., Bastian, N.D.: Deep vulman: A deep reinforcement learning-enabled cyber vulnerability management framework. *Expert Syst. Appl.* **221**, 119734 (2023)
 50. Sheng, Z., Yu, B., Liang, C., Zhang, Y.: Vpnet: A vulnerability prioritization approach using pointer network and deep reinforcement learning. In: International Conference on Digital Forensics and Cyber Crime, pp. 307–325. Springer (2022)
 51. Sewak, M., Sahay, S.K., Rathore, H.: Deep reinforcement learning for cybersecurity threat detection and protection: A review. In: International Conference On Secure Knowledge Management In Artificial Intelligence Era, pp. 51–72. Springer (2021)
 52. Liang, W., Huang, W., Long, J., Zhang, K., Li, K.-C., Zhang, D.: Deep reinforcement learning for resource protection and real-time detection in iot environment. *IEEE Internet Things J.* **7**(7), 6392–6401 (2020)
 53. An, D., Yang, Q., Liu, W., Zhang, Y.: Defending against data integrity attacks in smart grid: A deep reinforcement learning-based approach. *IEEE Access* **7**, 110835–110845 (2019)
 54. Arshad, K., Ali, R.F., Muneer, A., Aziz, I.A., Naseer, S., Khan, N.S., Taib, S.M.: Deep reinforcement learning for anomaly detection: A systematic review. *IEEE Access* (2022)
 55. Malialis, K., Kudenko, D.: Distributed response to network intrusions using multiagent reinforcement learning. *Eng. Appl. Artif. Intell.* **41**, 270–284 (2015)
 56. Hughes, K., McLaughlin, K., Sezer, S.: Policy-based profiles for network intrusion response systems. In: 2022 IEEE International Conference on Cyber Security and Resilience (CSR), pp. 279–286 (2022). IEEE
 57. Prasad, R., Mehr, S.A.Z., Moon, Y.: Recovery systems architecture for cyber-manufacturing systems against cyber-manufacturing attacks: Reinforcement learning approach. *Manuf. Lett.* **35**, 851–860 (2023)
 58. Wei, F., Wan, Z., He, H.: Cyber-attack recovery strategy for smart grid based on deep reinforcement learning. *IEEE Trans. Smart Grid* **11**(3), 2476–2486 (2019)
 59. Sewak, M., Sahay, S.K., Rathore, H.: Deep reinforcement learning in the advanced cybersecurity threat detection and protection. *Inf. Syst. Front.* **25**(2), 589–611 (2023)
 60. Voigt, P., Bussche, A.: The eu general data protection regulation (gdpr). *A Practical Guide*, 1st Ed. Springer, Cham 10(3152676), 10–5555 (2017)
 61. Gostin, L.O., Levit, L.A., Nass, S.J.: Beyond the hipaa privacy rule: enhancing privacy, improving health through research (2009)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.