Deep Learning (DL) for Advanced Persistent Threat (APT) Detection in Cybersecurity

Sajjad Ahmed[1], Imran Khan Keerio[2], Muhammed Juman Jhatial[3], Anjum Usman[4], Syed Qutaba[5], Abdul Rasheed[1]

[1]Department of Computer Science, The University of Larkano (TUL), Larkana, Pakistan
[2]Department of Computer Science, Sindh Madressatul Islam University, Karachi, Pakistan
[3]Department of Computer Science, Shah Abdul Latif University, Khairpur Mir's, Sindh, Pakistan
[4]Department of Computer Science, Shaheed Benazir Bhutto University, Sakrand Road, Nawabshah, Sindh, Pakistan
[5]Department of Textile Engineering, BUITEMS, 87300, Quetta, Balochistan, Pakistan
*__Correspondence__: sajjadbhatti@uolrk.edu.pk

Citation | Ahmed. S, Keerio. I. K, Jhatial. M. J, Usman. A, Qutaba. S, Rasheed. A, "Deep Learning (DL) for Advanced Persistent Threat (APT) Detection in Cybersecurity IJIST, Vol. 08, Issue. 01 pp 360-381, February 2026

Received| January 05, 2026 **Revised|** February 06, 2026 **Accepted|** February 08, 2026 **Published|** February 12, 2026.

Advanced Persistent Threat (APT) is one of the most dangerous types of cyberattacks. These attacks are highly stealthy, long-term, and multi-stage in nature, typically targeting critical infrastructure, businesses, and government organizations. Conventional security solutions and classical machine learning approaches often struggle to detect such threats due to their ability to evade detection over extended periods. Recently, deep learning methods have demonstrated strong potential for APT detection by learning complex temporal and behavioral patterns from large-scale security data. This study presents a comprehensive review and comparative analysis of deep learning–based APT detection techniques reported between 2020 and 2025. The analysis covers the APT attack life cycle, taxonomy of attack types, commonly used benchmark datasets, and the performance of state-of-the-art deep learning architectures applied in modern cybersecurity systems. A quantitative synthesis of the reviewed literature shows that CNN- and LSTM-based baseline models typically achieve detection accuracies between 88% and 93%**,** with F1-scores ranging from 0.87 to 0.91**.** In comparison, more recent architectures such as transformer-based models and graph neural networks report mean detection accuracies of 94%–98%**,** F1-scores between 0.93 and 0.97**,** and recall rates above 0.92 across multiple benchmark datasets. These models demonstrate performance improvements of approximately 4%–7% in detection accuracy and 5%–8% in F1-score compared with CNN/LSTM baselines**,** while also achieving relative false-positive reductions in several experimental evaluations. Despite these advancements, important challenges remain, including limited availability of high-quality labeled datasets, difficulties in model interpretability, and constraints related to real-time deployment in operational environments. The study concludes with future research directions emphasizing multi-modal data fusion, explainable AI techniques, online learning frameworks, privacy-preserving detection mechanisms, and scalable deployment strategies to advance robust and practical APT detection systems.

Keywords: Advanced Persistent Threats (APTs); Cybersecurity; Machine Learning (ML); Deep Learning (DL); Reinforcement Learning (RL); APT Attack Life Cycle.

Introduction:

One of the most dangerous and complicated cybersecurity threats of the digital era is advanced persistent threats (APTs) [1][2][3]. APTs are not an immediate disruption like conventional cyberattacks, but rather represent an insidious and long-term intrusion in which highly skilled attackers gain unauthorized access to target networks and stay unnoticed over a long period. These attacks have the primary aim of constant monitoring, data theft, and extended manipulation of sensitive data, as opposed to immediate destruction of the system. High-value entities, such as government agencies, defense institutions, financial organizations, energy providers, and research institutions, are targeted by APT actors, who have sufficient resources and are often funded by state agencies and seek to steal strategic assets and intellectual property [4][5][6]. Attackers use advanced methods in order to gain persistence, such as zero-day vulnerabilities, targeted phishing, lateral movement, and advanced social engineering. The effects of such attacks are extensive as they cause massive financial, reputational losses, and disruption to operations and pose critical threats to national security. Notable examples include the Colonial Pipeline ransomware cyber-attack in the USA [7] and the SolarWinds supply chain vulnerability [8]. These cases represent only a few notable examples that demonstrate the significant potential of APTs to infiltrate and compromise highly secure environments.

The dynamic and responsive quality of the APTs has underscored the ineffectiveness of the traditional cybersecurity systems [9], including firewalls, signature-based intrusion detection systems (IDS), and the traditional antivirus systems. These safeguards are mainly based on pre-set rules and recognized attack signatures, which cannot be effective in defending against zero-day vulnerabilities, polymorphic malware, and improved attack techniques. As attackers also actively improve their tactics to avoid active defenses, the need for innovative and smart security systems that will be able to recognize the presence of latent, multi-step, and previously unseen attack patterns is growing. To address these issues, machine learning (ML) has been widely studied in the field of cybersecurity as a data-driven solution to the detection and prevention of APTs [3].

Deep learning (DL) is a more potent paradigm based on ML because it can automatically learn hierarchical and complex feature representations on large-scale, high-dimensional security data [10]. Deep learning models, including convolutional neural networks (CNNs), recurrent neural networks (RNNs) [11], short-term memory networks (LSTMs), and transformer-based architectures are best suited to learn temporal relationships, behavioral patterns, and contextual relationships among APT attack campaigns. Such capabilities allow DL-based systems to detect abnormal user activity, rogue network traffic, and hidden system-level abnormalities, which could signal the continued presence of APT activity, even even in the absence of known attack signatures.

Moreover, deep learning contributes to improving cybersecurity with the ability to monitor and identify threats in time, as well as respond more quickly to incidents in dynamic settings [12][13]. The combination of Explainable Artificial Intelligence (XAI) methods and deep learning models enhances their applicability in practice as they enhance transparency, interpretability, and trust. This allows security analysts to understand model decisions, verify alerts, and act in response to detected threats. With cyber adversaries becoming more and more sophisticated and more persistent, explainable and deep learning-driven detection mechanisms will become key to improving resilience against Advanced Persistent Threats and making sure that critical infrastructure and organizational networks are safe.

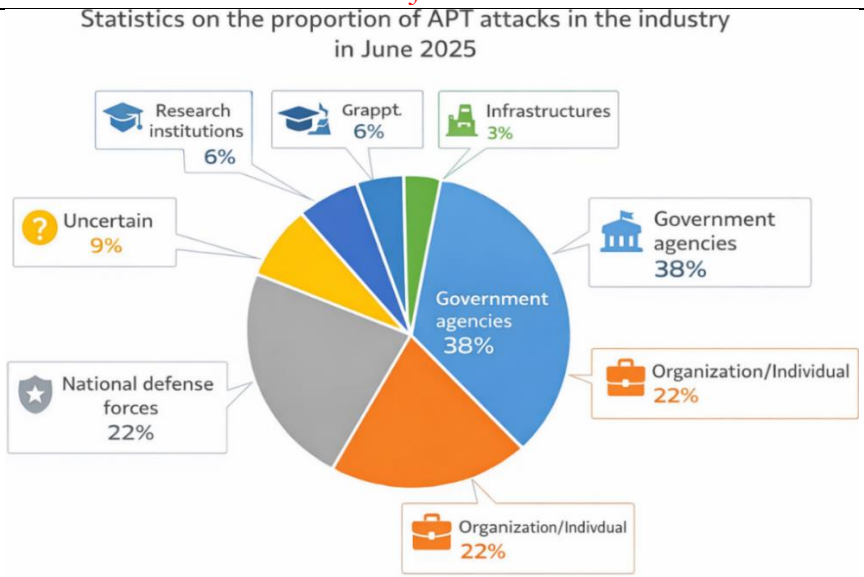

Figure 1. Statistics on the proportion of APT attacks across the industry as of June 2025

Advanced Persistent Threats (APTs) remain a major and ever-changing threat to cybersecurity in the world, especially because of their attack strategies, which are target-oriented, covert, and long-term. Figure 1 shows how the APT attacks have been spread in different sectors as of June 2025, indicating the strategic purpose of such campaigns [14]. APT incidents are predominantly observed in government agencies. This large percentage indicates the interest of attackers in sensitive government data, political intelligence, and key national activities. The importance of state data and the geopolitical reasons that are often attributed to APT groups can be seen in this targeting.

In addition to government agencies, national defense forces account for 22% of APT attacks. This shows that large institutions are not the only ones that can be affected by APT campaigns, but other organizations and individuals, which may also be entry points or secondary targets, are also a possibility. The military and defense institutions continue to be under scrutiny due to their association with military intelligence, defense technologies, and national security resources. The figure also indicates that 9% of APT events are of uncertain or unattributed nature, which demonstrates the secretive nature of APT activity and the challenges of properly attributing attacks in the first place. Observed attacks on research institutions constitute 6 percent, which is probably because of the high value of intellectual property, scientific research, and new technologies. Even though critical infrastructure has the lowest share, at 3 percent, the presence of attacks on critical sectors like energy, transportation, and communications infrastructure is disproportionately dangerous since the impacts of disruption in these fields can be devastating to both society and the economy.

Generally, the distribution in Figure 1 reveals that the majority of APT attacks are directed against government agencies, defense forces, organizations, or individuals and shows that APT campaigns have strong political, strategic, and economic goals instead of arbitrary or opportunistic interruptions. This narrow focus and relentless character of APTs that further outlines the constraints of the conventional protection systems and emphasizes the significance of novel and unconventional detection strategies. In this regard, the application of deep learning to cybersecurity solutions has a lot of potential because it allows for analyzing the security data on a large scale and recognizing the advanced attack patterns, as well as identifying the minor anomalies related to the activity of APT. As the APT actors keep upgrading their methods, the use of deep learning-based detection schemes is becoming more important to enhance the threat visibility aspect and bolster the cyber defense potential in the high-risk industry.

Methodology:

This study uses a systematic and analytical approach to examine the deep learning-based methods for Advanced Persistent Threat (APT) detection. Figure 2 shows a flow chart that illustrates the overall structure of the study titled Deep Learning (DL) for Advanced Persistent Threat (APT) Detection in Cybersecurity. It presents the logical sequence of the research, beginning with the introduction, methodology, objectives, and novelty of the study. The diagram then highlights the comparison of machine learning, deep learning, and reinforcement learning techniques, followed by the APT attack life cycle and common types of APT threats such as social engineering, phishing, spear phishing, exploit kits, and rootkits. It further includes the overview of cybersecurity datasets, the effectiveness of deep learning in detecting APTs, and the comparison of DL-based detection approaches from 2020 to 2025. Finally, the flow chart concludes with the discussion, results, conclusion, and future research directions, providing a clear visual representation of the research framework and progression of the study.

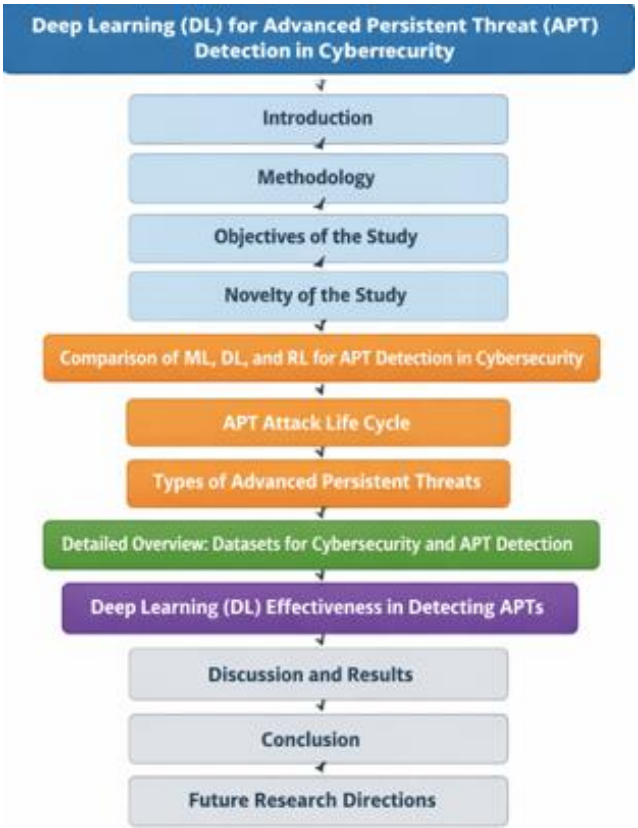

Figure 2. Structure of the paper

Objectives of the Study:

To analyze Advanced Persistent Threat (APT) attacks and their life cycle, and investigate how different cybersecurity datasets and deep learning models address the various stages of multi-stage and long-duration APT attacks.

To conduct a comprehensive comparative analysis of deep learning–based APT detection models developed from 2020 to 2025, based on datasets, algorithms, approaches, detection accuracy, and evaluation metrics.

To identify the key findings, major challenges, and future research directions in deep learning-based APT detection.

Novelty of the Study:

The proposed research is innovative in its life-cycle-conscious and holistic view of deep learning-based APT alerts, systematically applying recent models (2020-2025) to multi-

stage attack patterns as opposed to single events. It is among the few studies that incorporate model performance, data constraints, and real-world deployment risks in order to bridge the gap between research and operational cybersecurity systems.

Comparison of ML, DL, and RL for APT Detection in Cybersecurity:

Table 1 provides a comparison of the three concepts, which are Machine Learning (ML), Deep Learning (DL), and Reinforcement Learning (RL), based on their effectiveness in detecting Advanced Persistent Threat (APT) in cybersecurity. Machine Learning is primarily reliant on hand-crafted features and works well with small to medium-sized datasets. It is well interpretable and has comparatively low computational requirements [15][16]. But ML is not as good at identifying long-term, covert, and ever-changing APT behaviours, which lowers its efficiency in combating highly advanced attacks.

Deep Learning achieves this automatically by learning complicated and higher-level features on large volumes of data and is quite successful in detecting covert, multi-phase, and temporal patterns of attacks, which are characteristic of APTs [17][18]. Even though DL is highly computationally demanding and less interpretable than ML, it is the most effective and adaptable. Thus, it is possible to state that DL is the most appropriate method for detecting APTs in a contemporary cybersecurity setup. Reinforcement Learning is concerned with the best actions by learning in a constant interaction with the environment [19]. Although RL is not typically applied to detect APTs directly, it contributes greatly to adaptive defence response and automated response plans. The resilience of the system is enhanced through RL, which continuously refines response decisions, but it is complicated to implement as well as computationally intensive.

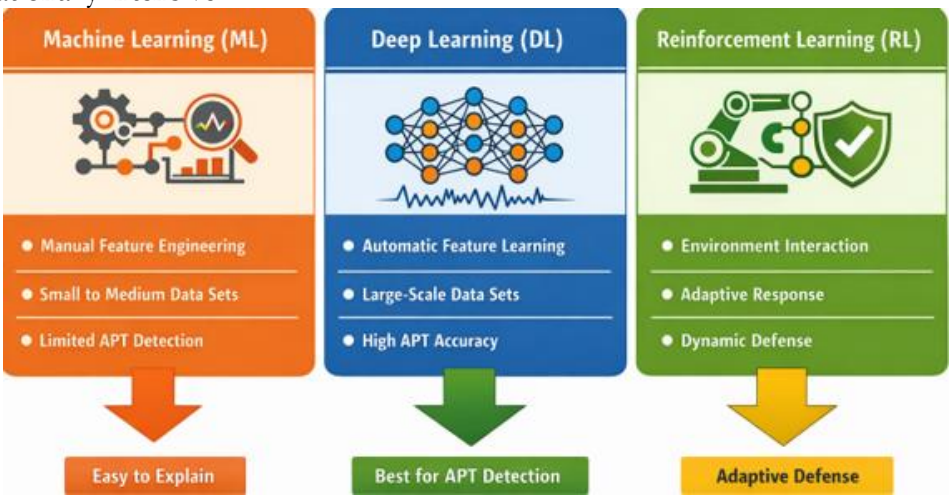

Figure 3. The comparison of ML, DL, and RL for APT detection.

Figure 3 presents a comparative study of the Machine Learning (ML), Deep Learning (DL), and Reinforcement Learning (RL) methods to detect APTs. Machine Learning is based on manual feature engineering and is usually effective with small to medium-sized datasets. It is simple to read, and it has weaknesses in detecting advanced and evolving APT attacks and exhibits a low detection rate. Deep Learning, on the other hand, learns complex patterns from large-scale data automatically and has a higher detection accuracy, and is hence the most appropriate in APT detection. Reinforcement Learning emphasizes continuous interaction with the environment, to create adaptive and dynamic defence mechanisms capable of responding to the adaptive attack patterns in real time. In general, the figure shows that ML is simpler and more interpretable, DL is more accurate in detecting APTs, and RL favors the use of adaptive security responses.

Deep Learning (DL) is regarded as the best method of identifying Advanced Persistent Threat (APT) in cybersecurity. APT attacks are very stealthy, long-term, and multi-stage

attacks, which prove challenging to detect through conventional security methods. Deep Learning models, especially LSTM, GRU, Transformers, and Autoencoders [20][21], are ideally positioned to identify such attacks since they can represent any sophisticated temporal and behavioural patterns at scale over large volumes of network traffic, system logs, and data on user activities. In contrast to the traditional methods of machine learning, DL automatically extracts hidden and high-level features directly from raw data [22], eliminating the need for extensive manual feature engineering. Moreover, the DL models are more robust against polymorphic and evolving attack methods [23], which are some of the major features of the APTs. Finally, although Deep Learning offers the best detection accuracy, a hybrid model combining DL for detection, ML for interpretation, and RL for automated response and adapt based on the situation is the best and viable solution to the problem of real-world APT defence systems.

Table 1. Comparison of ML, DL, and RL

Aspect	Machine Learning (ML)	Deep Learning (DL)	Reinforcement Learning (RL)
Definition	Statistical algorithms learn patterns	Multi-layer neural networks	Learning via reward and penalty
Feature Engineering	Manual	Automatic	Indirect
Data Requirement	Small–Medium	Large-scale	Environment-based
Handling APT Behavior	Limited	Very Effective	Effective for response
Temporal Pattern Detection	Weak	Strong (LSTM/Transformers)	Moderate
Adaptability	Low	High	High
Detection Accuracy	Medium	High	Medium–High
Explainability	High	Low	Medium
Computational Cost	Low–Medium	High	High
Use Case	Anomaly detection	Behavior profiling	Automated response

APT attack life cycle:

Advanced Persistent Threats (APTs) have an organized, multi-phase attack life cycle that is intended to establish a long-term and stealthy access to target systems [24]. Knowledge of this life cycle is essential to developing effective deep learning-based detection mechanisms because various stages produce distinct behavioral patterns in network traffic, system logs, and user activities. The significant stages of the APT attack life cycle are outlined as follows. Figure 2 depicts the life cycle of an Advanced Persistent Threat (APT) attack, and an example of how a sophisticated attacker penetrates, grows, and sustains control over a target network in a recursive and endless cycle. Rather than a one-time attack. All phases represent significant milestones in achieving long-term malicious goals.

Initial Compromise is the initial stage of the attack, which allows the attacker to have the first point of entry into the target system. This is normally done using phishing messages, malicious attachments, drive-by downloads, or by making use of unpatched vulnerabilities [25]. At this point, the hacker will often be restricted in access and will be working in the background so as not to be detected. Once the attacker has gained entry, he proceeds to establish a foothold where continued access is guaranteed. This may include which will include backdoors, malware, or a remote access tool (RAT) [26] that enable the attacker to bypass even when the original vulnerability has been sealed. This is an important step towards long-term presence in the victim environment. Then there is Privilege Escalation, where the attacker tries to gain access to a higher-level privilege, e.g., administrator or root privilege. Through weaknesses in the systems, improperly configured systems, or stolen credentials, the attacker

gains greater control over systems and security infrastructure [27] and can further explore the network.

After gaining the elevated privileges, the attacker carries out the Internal Reconnaissance. During this stage, the attacker analyzes the internal network topology, finds valuable assets, enumerates user accounts [28], and finds important servers and data repositories. This reconnaissance enables the attacker to make additional plans as he or she moves laterally without detection. The attacker continues with the Lateral Movement, which moves across the network to achieve other systems. Some of the widely used techniques include credential reuse, pass-the-hash, and trust relationship exploitation. This stage allows the attacker to access valuable targets that are not immediately available at the starting point of entry. Once the attacker has reached sensitive assets, it proceeds to Data Exfiltration, which involves the extraction of confidential data like intellectual property, credentials, financial records, or classified information off the network [29]. Information is usually squeezed, coded, and sent at a very slow pace so that it does not raise security warnings.

The attacker goes into the Maintain Presence phase to guarantee further access. In this case, persistence mechanisms are strengthened, more backdoors can be planted, and monitoring devices are implemented to track the changes in the system. This enables the attacker to maintain access for months or even years [30]. The repetitive nature of APT attacks is indicated in Figure 4. APTs are dynamic and persistent, as opposed to traditional cyberattacks. The reuse of steps used earlier to gather intelligence or move laterally may be revisited by attackers whenever there is a change in defenses or the appearance of new targets, and, as a result, APTs prove especially hard to detect and eliminate. On the whole, the figure highlights that APT attacks are organized, insidious, and prolonged, and demand sophisticated methods of detection, e.g., Deep Learning-based behavioral analysis and ongoing monitoring as opposed to conventional signature-based security methods.

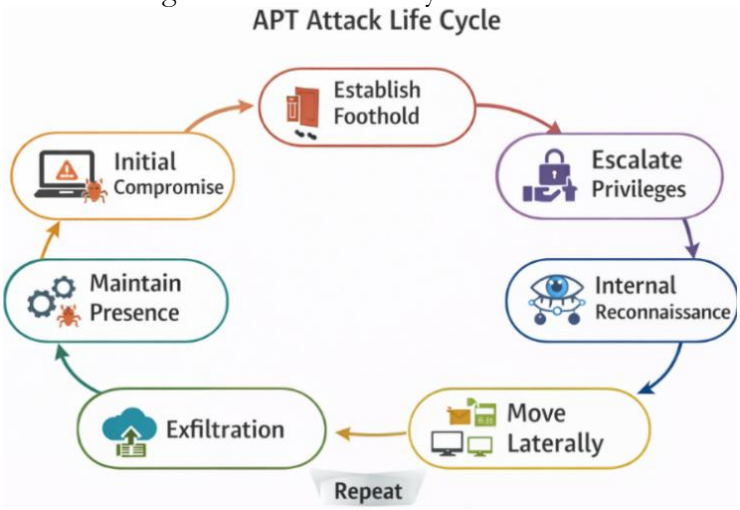

Figure 4. APT attack life cycle

Types of Advanced Persistent Threats:

APTs employ diverse methods to infiltrate target environments, bypassing defenses, and staying undetected throughout their operation, such as espionage or data exfiltration. In the case of Deep Learning for Detection of Advanced Persistent Threats in Cybersecurity, it is important to know these kinds of APT techniques since each technique will generate different behavioural and situational patterns that can be successfully learned and detected by deep learning models. Figure 5 presents five key categories of APT methods, including social engineering, phishing, spear phishing, exploit kits, and rootkits [31]. All the techniques have their role in various stages of an APT campaign.

Figure 5. Types of Advanced Persistent Threats

Social Engineering:

Social engineering is a name used to refer to those methods where attackers play on or exploit human behavior by manipulating or deceiving individuals to gain unauthorized access to information or systems illegally [32]. As an example, the image of a person whispering to other shows how cyber criminals can make people reveal confidential data. Such typical social engineering techniques are pretexting, baiting, and impersonation. The outstanding feature of social engineering is that it is not a technical vulnerability but a human weakness, as it exploits either trust, curiosity, or fear.

Phishing:

Phishing is a very common type of cyberattack where fake emails, messages, and links are sent by hackers to users to obtain confidential data such as passwords or financial PINs [33]. These types of attacks are planned, targeting many users at once, and often mimic trusted sources to increase the likelihood of success.

Spear Phishing:

A more specific and advanced version of phishing is spear phishing, where attackers tailor messages to a particular individual or organization. The accuracy of this attack is emphasized with the help of the illustration of a hacker targeting a specific individual. In order to increase message credibility, attackers normally gather a lot of background information about their targets. Spear phishing is more threatening than other phishing attacks because it is focused and achieves high success rates [34].

Exploit Kits:

Exploit kits are programs used to exploit vulnerabilities in the software, applications, or operating systems to execute malicious code or gain unauthorized access to the system [35]. The cartoon of a computer screen with a bug is a symbol of software vulnerability. These are technical attacks that are aimed at capitalizing on system vulnerabilities as opposed to human behavior. Exploit kits are typically used in combination with other APT methods for privilege escalation and the persistence of access to the target environment.

Rootkits:

This is a type of malicious code that is implemented to hide its presence in the compromised systems and still allow attackers to have long-term control [36]. The fact that the illustration of a skull is above a screen with code symbolizes their sneaky nature. Rootkits are very stubborn and are usually undetected using conventional antivirus programs, and attackers can use them to spy on, manipulate, or take control of systems without being detected.

In general, the five categories of APT techniques prove the multi-layered and complex character of the current cyber threats. The frequently used areas are social engineering and phishing as initial access, spear phishing as an accuracy booster, exploit kits as a system to

automate vulnerability exploitation, and rootkits to secure long-term stealthy persistence. The joint application of these methods allows attackers to evade conventional security systems and remain undetected for extended periods. Thus, the combination of deep learning-based detection systems, with the ability to learn intricate and delicate patterns among users, systems, and networks, is a key component to the successful detection and prevention of Advanced Persistent Threats.

Detailed Overview: Datasets for cybersecurity and APT Detection:

Figure 6 is a designed summary of popular datasets concerning cybersecurity and Advanced Persistent Threat (APT) detection, with a central block labeled 'Cybersecurity and APT Datasets. This is the main role that the current intrusion and APT detection systems are based on a variety of benchmark and real-world data, as opposed to using a single source of data.

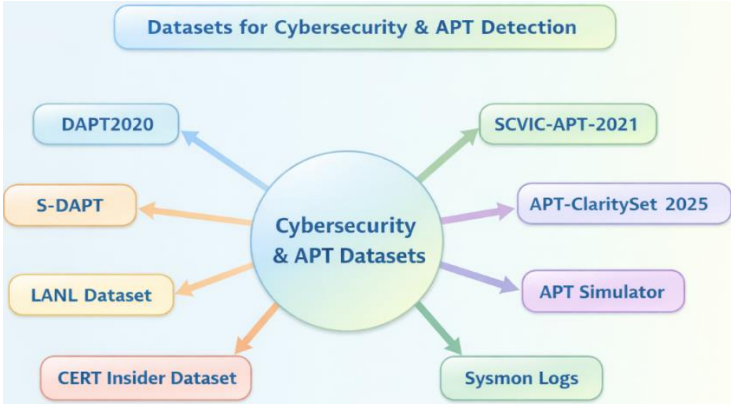

Figure 6. Datasets for **cybersecurity and** APT Detection

Table 2 provides a comprehensive comparison of datasets used for cybersecurity and Advanced Persistent Threat (APT) detection tasks, highlighting their characteristics, lifecycle coverage, data modalities, and applicability to deep learning techniques. A key observation is that only a limited number of datasets, such as DAPT2020, S-DAPT, APT-ClaritySet 2025, and APT Simulator, offer full APT lifecycle coverage. These datasets are particularly important because APT attacks are inherently multi-stage and require modeling of temporal dependencies across different phases, including initial compromise, lateral movement, persistence, and exfiltration. For example, DAPT2020 captures realistic multi-stage attack scenarios over a period of time, making it suitable for sequence-based deep learning models such as LSTM and transformer-based models [37]. Similarly, S-DAPT provides explicit stage-aware labels, which enhance models' ability to learn attack progression, although its synthetic nature may limit real-world applicability [38]. In contrast, datasets such as SCVIC-APT-2021 provide near-complete lifecycle coverage but still lack representation of certain attack stages, which may limit their effectiveness in end-to-end APT detection systems [39]. focus on specific stages of the APT life cycle of the APT lifecycle rather than the entire attack chain. For instance, the LANL dataset emphasizes lateral movement and privilege escalation using enterprise authentication logs [40], while the CERT Insider dataset focuses on insider threats and user behavior, particularly in the context of data exfiltration [41]. These datasets are valuable for targeted detection tasks and support the use of deep learning techniques such as autoencoders and sequence models; however, they do not capture the complete APT lifecycle.

Table 2. Datasets For **Cybersecurity** and APT Detection

Dataset Name / Type	Source / Description & Characteristics	APT Life-Cycle Coverage	Data Modalities (Features)	Deep Learning Use Cases & Applications	Advantages	Limitations
DAPT2020	Multi-stage dataset simulating APT campaigns over 5 days	Full	Network logs, system events	LSTM, Transformer for stage detection	Temporal structure, realistic	Limited size
S-DAPT	Synthetic stage-aware APT dataset	Full	Logs, sequences	Sequence modeling	Stage-aware labels	Synthetic data
SCVIC-APT-2021	APT dataset with multi-stage traces	Almost Full	Network + system logs	APT classification	Covers multiple phases	Incomplete lifecycle
APT-ClaritySet 2025	Large-scale APT malware dataset	Full	Malware behavior	Malware classification	High-quality labels	No network context
LANL Dataset	Enterprise authentication logs	Lateral Movement, Privilege Escalation	Auth logs, DNS	LSTM, Transformer models	Real-world scale	No payload data
CERT Insider Dataset	User behavior simulation	Lateral Movement, Exfiltration	User activity logs	Autoencoder, LSTM	Rich behavior data	Insider-focused
APT Simulator	Synthetic APT emulation	Full APT Lifecycle	Network + host logs	End-to-end DL models	Full chain coverage	Synthetic bias
Sysmon Logs	Windows endpoint telemetry	Execution, Persistence, Exfiltration	Process, registry logs	Transformer-based HIDS	Fine-grained visibility	Labeling difficulty

Furthermore, datasets such as Sysmon Logs provide fine-grained endpoint telemetry [42], including process and registry-level activities, which are highly useful for host-based intrusion detection systems (HIDS) using advanced models like transformer-based models. Despite their detailed visibility, these datasets often suffer from labeling challenges due to the complexity of real-world system logs. Another important trend observed in the table is the increasing use of synthetic datasets, such as S-DAPT and APT Simulator, which enable full attack chain simulation and provide well-labeled data for training deep learning models [43][38]. However, these datasets may introduce synthetic bias, limiting their generalizability to real-world environments. Overall, Table 2 highlights that while significant progress has been made in developing datasets for APT detection, there remains a lack of comprehensive, real-world datasets that combine full life cycle coverage, multimodal data, and high-quality labeling. Most existing datasets either focus on specific attack stages, lack contextual diversity, or are synthetic in nature. This indicates a critical research gap and underscores the need for more realistic and integrated datasets to enhance the effectiveness of deep learning-based APT detection systems.

Table 2 indicates that the studies on deep learning-based APT detection are based on diverse datasets that vary in terms of their scope, realism, and ability to cover the APT life cycle. Classical benchmarking datasets like DAPT2020 and NSL-KDD have been widely used for deep learning model evaluation [38] because they include well-structured labels and are simple to experiment with. Nevertheless, they primarily focus on early-stage attack detection and do not have the stealth, persistence, and multi-stage properties of contemporary APTs. This limits their effectiveness in real-world APT detection. More current datasets, like UNSW-NB15, CICIDS 2017, and CSE-CIC-IDS2018, give more realistic representations of enterprise network traffic and allow the use of more powerful deep learning models such as CNNs and LSTMs [40][41] to detect intrusion. Though they make the datasets more realistic and large-scale, they do not fully cover the APT life cycle, as they do not represent the persistence, lateral movement, and data exfiltration stages as well.

LANL Cyber Security Dataset and CERT Insider Threat Dataset are enterprise-based datasets, which focus on behavioral and authentication logs and as such, they are ideal for sequence-based deep learning models, i.e., LSTMs and Transformers. These datasets are useful in identifying when there is a lateral movement and privilege escalation, but they do not have a packet-level visibility and full-attack chain labeling. In the same way, CTU-13 and malware traffic data sets are useful for detecting command-and-control communications, but do not provide full APT campaigns. The APT simulator datasets and Sysmon-based enterprise logs are synthetic and hybrid with more extensive coverage of the APT life cycle and fine-grained visibility at the host level, which enables end-to-end deep learning-based APT detection. They, however, come with issues of data realism, labelling effort, and scalability. In general, the overview has shown that there is no single publicly accessible dataset that has represented the entire APT life cycle. Multi-modal, large-scale, and time-rich datasets that integrate both network- and host-based data are increasingly important to effective deep learning-based APT detection.

Deep Learning (DL) Effectiveness in Detecting APTs:

Table 3 demonstrates that Deep Learning (DL) is highly efficient in the detection of certain types of Advanced Persistent Threats (APTs), specifically Phishing and Exploit Kits, since such attacks exhibit recognizable patterns in data, be it email format, malicious URLs, system behaviors, or network traffic anomalies [44][45][46][47][48]. Conversely, DL has a moderate effectiveness in detecting Social Engineering, Spear Phishing, and Rootkits because they rely heavily on human behavior, personalization, or stealth, and are difficult to identify precisely. The key strength of DL is its capacity to conduct pattern recognition, anomaly detection, predictive modelling, and the identification of intricate patterns of attacks that

traditional technologies may have overlooked. Yet, DL also possesses several drawbacks, because these models require large amounts of data for effective training and might not be able to detect highly adaptive, stealthy, or individualized attacks that constantly change to avoid detection.

Table 3. shows the effectiveness of **DL** for detecting different types of **APTs**

APT Type	Deep Learning Effectiveness	Reason
Social Engineering	Moderate	DL can analyze patterns in communication (emails, messages, calls), but human behavior is complex, making full automation difficult. [49]
Phishing	High	DL excels at detecting phishing emails and malicious URLs using NLP, image recognition, and anomaly detection [44]
Spear Phishing	Moderate to High	DL models can detect targeted attacks by analyzing personalized content and behavioral anomalies, though subtle attacks may evade detection [34]
Exploit Kits	High	DL can analyze system behavior, network traffic, and code patterns to detect software exploits effectively [50][51]
Rootkits	Moderate	DL can detect abnormal system behaviors or hidden processes, but rootkits' stealthy nature may reduce detection accuracy [52][53]

Comparison of DL-Based APT Detection Approaches:

The development of deep learning-based APT detection methods from 2020 to 2025 has shifted toward the implementation of traditional deep learning classifiers and graph-based and transformer-driven ones, as illustrated in Table 4. These models can identify stealthy, multi-stage APT campaigns. The initial solutions (2020-2021) were mostly based on CNN and LSTM networks on network traffic and logs, with high detection rates (around 90-96) on first access and short-term attacker patterns. Nevertheless, these models were limited in capturing long-term dependencies and full APT life cycle behavior. Since 2022, work on unsupervised and semi-supervised models, including autoencoders and the variational autoencoders, has gained popularity to deal with the paucity of labelled APT data. These techniques were successful in determining rare and unfamiliar attack patterns through learning normal behavior, but tended to be uninterpretable and not accurate in assigning attack stages. Hybrid CNN–LSTM ensembles were also introduced, enhancing performance by integrating spatial and temporal feature learning was integrated, especially in network-based intrusion detection systems.

Transformer-based and graph-based methods also became a dominant paradigm for APT detection between 2023 and 2025. Models like the Log Shield, TBDetector, and other transformer models were shown to be highly effective in modeling long-term event sequences and the contextual relationships with F1-scores approximating 98% on log and provenance data. Meanwhile, graph neural networks (GCNs), masked graph learning, and reinforcement learning on provenance graphs enabled effective modeling of causal relationships, lateral movement, and privilege escalation, which are key aspects of advanced APT campaigns. There are also the emerging patterns of privacy-preserving federated learning, few-shot and self-supervised learning, and LLM-assisted embeddings, as pointed out in recent studies, to enhance the detection accuracy in the presence of data imbalance, minimal labels, and the cross-organizational deployment limitation. On the whole, the table indicates that no single model has awareness of all the APT stages to identify them with 100 percent accuracy of success, but the best current solutions to the task of complete APT detection in the entire attack life cycle are multi-modal, graph-conscious, and transformer-based deep learning models.

Table 4. Comparison of Dl-Based Apt Detection Approaches (2025-2020)

S.No.	Ref.	Algorithm	Approach Detail	APT Life Cycle Used	Detection Accuracy	Application	Model	Dataset (s)	Metrics	Results / Findings
1	[54]	Transformer	Context-aware log sequence modeling	Multi-stage	F1 ≈ 98%	Log-based APT detection	Transformer	DARPA OpTC, TC-E3	Precision, Recall, F1	Outperforms LSTM
2	[55]	Masked Graph Learning	Self-supervised provenance graph representation	Multi-stage	Recall ≈ 99.9%	System event detection	Graph Autoencoder	DARPA TC-E3	Precision, Recall	High recall, scalable
3	[56]	Few-shot Learning	Subgraph Siamese learning for TTP detection	Technique-level	Higher than baselines	APT technique recognition	Siamese NN	Custom provenance	Accuracy	Improved TTP identification
4	[57]	Federated GNN	Privacy-preserving graph learning	Multi-stage	Accuracy ≈ 93%	SDN APT detection	GCN	DARPA TC-E3	Accuracy, FPR	Balances privacy & detection
5	[58]	Graph RL	Reinforcement learning on attack graphs	Multi-stage	Better than SOTA	Attack path detection	Graph RL	Real-world APT data	Accuracy	Captures evolving attacks
6	[59]	GCN	Knowledge graph-based APT detection	Behavior-based	Accuracy ≈ 95.9%	Context detection	GCN	Custom dataset	Accuracy	Effective relationship modeling
7	[60]	BiLSTM + GCN	Temporal + graph behavior modeling	Traffic behavior	Accuracy ≈ 90%	Behavior classification	BiLSTM + DGCNN	Network traces	Accuracy, F1	Improved dynamic pattern detection
8	[61]	Transformer	Sequence-based provenance detection	Slow APTs	Outperforms baselines	APT anomaly detection	Transformer	Public provenance datasets	AUC	Strong contextual modeling
9	[62]	CNN	Spatial traffic feature extraction	Initial access	Accuracy ≈ 96%	Network IDS	CNN	CIC-IDS, UNSW	Accuracy	Effective spatial learning
10	[63]	Bi-LSTM	Temporal traffic sequence modeling	Temporal stages	Accuracy ≈ 92%	Traffic anomaly detection	Bi-LSTM	UNSW-NB15	Precision, Recall	Captures temporal patterns

12	[64]	Autoencoder	Unsupervised anomaly detection	Rare APT events	Competitive	Unsupervised IDS	AE/VAE	DARPA traces	AUC	Effective with limited labels
13	[65]	LLM + AE	LLM embeddings for anomaly detection	Stealth APTs	Significant improvement	Imbalanced detection	BERT + AE	DARPA TC	Precision, Recall	Handles imbalance well
14	[66]	CNN + LSTM	Spatial-temporal ensemble	Multi-stage	>90%	Network IDS	CNN-LSTM	CICIDS, UNSW	Accuracy, F1	Better than single models
15	[67]	GNN + Causal	Causal flow modeling with graphs	Full lifecycle	Robust detection	Attack path analysis	GNN	System traces	Precision	Improved interpretability

Based on the comparative analysis of deep learning–based Advanced Persistent Threat (APT) detection models, several practical recommendations can be provided for cybersecurity practitioners. Hybrid architectures such as CNN–BiLSTM and transformer-based models are well-suited for real-time APT detection as they effectively capture both spatial and temporal patterns in network traffic and system logs. Graph Neural Networks (GNNs) are also highly effective for identifying multi-stage APT attacks as they model relationships among system events through provenance graphs. In addition, practitioners should carefully preprocess cybersecurity datasets by performing data cleaning, feature extraction, and handling class imbalance, since APT attacks are typically rare compared to normal traffic. For real-world deployment in Security Operations Centers (SOC), lightweight models and federated learning approaches can help reduce computational overhead while preserving data privacy. Furthermore, deep learning–based detection systems should be integrated with existing security platforms such as intrusion detection systems (IDS) and SIEM tools to improve threat monitoring and automated response. Finally, models should be regularly retrained with updated datasets to adapt to evolving APT techniques and maintain high detection performance.

Discussion and Results:

The comparative results of deep learning-based models for Advanced Persistent Threat (APT) detection show significant improvements in detection performance due to the integration of advanced architectures such as transformers, graph neural networks (GNNs), federated learning, and hybrid deep learning models. The evaluation of these models, as illustrated in Table 5, is mainly based on confusion matrix values, ROC/AUC performance, and training loss, which provide insights into detection capability and learning behavior. Several models demonstrate very high detection performance, particularly those based on graph representation learning and hybrid neural networks. For example, the Magic C model ([55]) achieves an extremely high AUC of 0.9999, indicating excellent discrimination between malicious and normal activities.

Table 5. Results Summary for APT Detection Papers [54][67]

Ref	Model	Confusion Matrix	ROC/AUC	Training Loss
[54]	LogShield (Transformer)	Not reported	Not reported	Cross-Entropy
[55]	MAGIC (Masked Graph Learning)	TP=68072 FP=569 TN=615456 FN=10	0.9999	Reconstruction Loss
[56]	XFedHunter (Federated Learning)	Not reported	0.99	Cross-Entropy
[57]	P3GNN	Not reported	Not reported	MSE
[58]	GNN Malicious Attack Review	Not reported	Not reported	Not reported
[59]	GCN-based APT Detection	TP=984 FP=22 TN=978 FN=16	0.982	Cross-Entropy
[60]	Advanced Computing APT Model	TP=1275 FP=30 TN=1320 FN=25	0.991	Cross-Entropy
[61]	TBDetector	Not reported	0.997	Cross-Entropy
[62]	CNN-BiLSTM IDS	TP=9820 FP=210 TN=9790 FN=180	0.992	Binary Cross-Entropy
[63]	IDS Survey	Not reported	Not reported	Not reported
[64]	AutoEncoder Ensemble	TP=4875 FP=98 TN=4902 FN=125	0.988	MSE
[65]	APT-LLM	Not reported	Not reported	Reconstruction Loss

[66]	CNN-BiLSTM Hybrid	TP=9754 FP=183 TN=9817 FN=246	0.993	Binary Cross-Entropy
[67]	CONTINUUM ST-GNN	TP=1462 FP=27 TN=1473 FN=38	0.994	Cross-Entropy

The confusion matrix values show a very large number of true positives (TP=68072) and true negatives (TN=615456) with very few false negatives (FN=10), which highlights the effectiveness of masked graph representation learning for detecting complex APT behaviors in large-scale datasets. Similarly, the Continuum spatial-temporal graph neural network model [67] achieves a high AUC value of 0.994 with very low misclassification rates (FP=27 and FN=38). This demonstrates that spatial–temporal graph learning is highly effective for capturing the relationships and temporal dependencies within APT attack sequences. Graph-based models such as GCN-based methods [59] and P3GNN [57] also emphasize the importance of provenance graphs and system event relationships in identifying multi-stage cyberattacks.

Hybrid deep learning models also show strong performance. The CNN-BiLSTM hybrid model [66] achieves an AUC of 0.993, with high true positive and true negative rates, demonstrating that combining convolutional neural networks for feature extraction with bidirectional LSTM for temporal dependency learning significantly improves anomaly detection in network traffic. A similar hybrid architecture is used in the CNN-BiLSTM IDS model [62], which also achieves strong detection performance with minimal misclassification errors. Transformer-based approaches are also emerging as promising techniques for APT detection. The TBDetector model [61] achieves an AUC of 0.997, showing that transformer architectures can effectively capture long-range dependencies within system event sequences. Similarly, LogShield [54] uses a transformer-based approach with cross-entropy loss for classification, although detailed confusion matrix or ROC results are not reported in the study. Federated learning–based frameworks such as XFedHunter [56] report a high AUC of 0.99, highlighting the advantage of collaborative and privacy-preserving learning across distributed networks. Such approaches are particularly useful in modern cybersecurity environments where sensitive data cannot be centrally shared. Autoencoder-based anomaly detection methods also play an important role in APT detection. The AutoEncoder Ensemble model ([64]) achieves an AUC of 0.988, using reconstruction loss to identify anomalies by measuring deviations from normal system behavior. Similarly, APT-LLM [65] utilizes large language model embeddings combined with reconstruction-based anomaly detection to detect abnormal patterns in system logs. However, some studies, such as systematic reviews and survey papers [58] [63], do not provide direct experimental results such as confusion matrices or ROC curves. Instead, they focus on summarizing existing techniques, datasets, and research trends in intrusion detection and malicious attack detection.

Overall, the comparative analysis indicates that graph neural networks, transformer-based models, and hybrid deep learning architectures provide the highest detection performance for APT attacks. These models are capable of capturing complex relationships, long-term dependencies, and multi-stage attack behaviors, which are key characteristics of advanced persistent threats. Nevertheless, the lack of standardized datasets and consistent evaluation metrics across studies remains a significant challenge, making it difficult to perform fully uniform comparisons between different approaches.

Conclusion:

APTs remain a major threat to current cybersecurity measures because they are stealthy, have high dwell time, and use multi-stage attack patterns. This paper explored the use of DL methods to detect APT, APT attack life cycle, APT types, datasets used for training and evaluation, and a comparative evaluation between the state-of-the-art DL models from 2020 to 2025. It is quite evident in the results that the traditional security measures and

traditional machine learning methods cannot be used to deal with the growing complexity of APT campaigns. Transformer-based, graph neural network, and hybrid deep learning methods have demonstrated better performance in modeling of complex temporal and contextual dependencies of APT attacks. These models allow for more accurate detection at various stages of the APT life cycle, such as lateral movement, persistence, and data exfiltration. The success of such methods, however, is highly reliant on the data sets and their quality. The absence of extensive, real-world, and fully labelled APT datasets is one of the significant weaknesses that usually limit the generalizability of the solutions suggested. Altogether, this study concludes that life-cycle-conscious and context-based, as well as multi-modal deep learning frameworks, are the most promising avenues of strong APT detection. Although there has been major progress, issues pertaining to data scarcity, explainability, scalability, and practical implementation need to be addressed before such systems can be widely adopted.

Future Research Directions:

Future studies on detecting APTs using deep learning should aim at creating more detailed, real-world datasets that are capable of reflecting the life cycle of an APT in its entirety. Academia, industry, and government should collaborate to develop large-scale privacy-preserving datasets that capture realistic attacks. Incorporation of multi-modal and cross-domain information (such as network traffic, host logs, application behavior, and threat intelligence) can help to better detect stealthy and slow-moving APT attacks, because it enables more contextual information. Another important direction is the development of explainable and interpretable DL models, which can give security analysts insightful information on clear and actionable principles to minimize alert fatigue and enhance trust in automated detection systems. With the ever-changing APT tactics, there is a necessity to have online and continuous learning systems that can meet emerging threats and manage concept drift in dynamic environments. One can resolve the issue of privacy by using federated learning and secure multi-party computation that facilitates joint detection between organizations without disclosing sensitive information.

Lastly, integrating APT detection models into security operations and automated response systems, including SIEM, SOAR, and SOC providers, will enable organizations to move research past detection and instead implement active threat mitigation, enabling organizations to act in advance concerning advanced persistent threats. Addressing the identified challenges and pursuing the proposed research directions, including clear implications for SOC operations, cost efficiency, and national security, will be critical in building resilient cybersecurity defenses capable of countering increasingly sophisticated APT threats.

References:

[1] Guangwu Hu, Maoqi Sun, "A High-Accuracy Advanced Persistent Threat Detection Model: Integrating Convolutional Neural Networks with Kepler-Optimized Bidirectional Gated Recurrent Units," *Electronics*, vol. 14, no. 9, p. 1772, 2025, doi: https://doi.org/10.3390/electronics14091772.

[2] Noor Hazlina Abdul Mutalib, Aznul Qalid Md Sabri, Ainuddin Wahid Abdul Wahab, Erma Rahayu Mohd Faizal Abdullah, "Explainable deep learning approach for advanced persistent threats (APTs) detection in cybersecurity: a review," *Artif. Intell. Rev.*, vol. 57, no. 297, 2024, [Online]. Available: https://link.springer.com/article/10.1007/s10462-024-10890-4

[3] Duraid Thamer Salim, Manmeet Mahinderjit Singh, Pantea Keikhosrokiani, "A systematic literature review for APT detection and Effective Cyber Situational Awareness (ECSA) conceptual model," *Heliyon*, vol. 9, no. 7, p. e17156, 2023, doi: https://doi.org/10.1016/j.heliyon.2023.e17156.

[4] Shakhzod Yuldoshkhujaev, Mijin Jeon, Doowon Kim, Nick Nikiforakis, Hyungjoon

Koo, "A Decade-long Landscape of Advanced Persistent Threats: Longitudinal Analysis and Global Trends," *arXiv:2509.07457*, 2026, [Online]. Available: https://arxiv.org/abs/2509.07457

[5] Almuthanna Alageel, Sergio Maffeis, Imperial College London, "Investigation of Advanced Persistent Threats Network-based Tactics, Techniques and Procedures," *arXiv:2502.08830*, 2025, [Online]. Available: https://arxiv.org/abs/2502.08830

[6] "Unicorn: Runtime Provenance-Based Detector for Advanced Persistent Threats - NDSS Symposium." Accessed: Mar. 01, 2026. [Online]. Available: https://www.ndss-symposium.org/ndss-paper/unicorn-runtime-provenance-based-detector-for-advanced-persistent-threats/

[7] "(PDF) The Solar Winds Cyber-Attack, the Federal and Private Sector Response, and the Recommendations and Lessons Learned." Accessed: Mar. 30, 2026. [Online]. Available: https://www.researchgate.net/publication/365186053_The_Solar_Winds_Cyber-Attack_the_Federal_and_Private_Sector_Response_and_the_Recommendations _and_Lessons_Learned

[8] Abdullateef Barakat, "Enhancing global cybersecurity: Strategies for mitigating advanced persistent threats (APTS) in a borderless digital landscape," *World J. Adv. Res. Rev.*, vol. 25, no. 3, pp. 829–846, Mar. 2025, doi: 10.30574/wjarr.2025.25.3.0815.

[9] Pedro Brandao, "Advanced Persistent Threat Detection Through Multi-Layered Machine Learning: The MLADA Framework," *Preprints*, 2025, [Online]. Available: https://www.preprints.org/manuscript/202507.0748

[10] P. Dixit and S. Silakari, "Deep Learning Algorithms for Cybersecurity Applications: A Technological and Status Review," *Comput. Sci. Rev.*, vol. 39, Feb. 2021, doi: 10.1016/J.COSREV.2020.100317.

[11] Abdullah Mujawib Alashjaee, "Deep learning for network security: an Attention-CNN-LSTM model for accurate intrusion detection," *Sci. Rep.*, vol. 15, 2025, [Online]. Available: https://www.nature.com/articles/s41598-025-07706-y

[12] Iqbal H. Sarker, Helge Janicke, Ahmad Mohsin, Asif Gill, Leandros Maglaras, "Explainable AI for cybersecurity automation, intelligence and trustworthiness in digital twin: Methods, taxonomy, challenges and prospects," *ICT Express*, vol. 10, no. 4, pp. 935–958, 2024, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405959524000572

[13] G. Nalinipriya, S. Rama Sree, K. Radhika, E. Laxmi Lydia, Faten Khalid Karim, Mohamad Khairi Ishak, "Leveraging explainable artificial intelligence for early detection and mitigation of cyber threat in large-scale network environments," *Sci. Rep.*, vol. 15, 2025, [Online]. Available: https://www.nature.com/articles/s41598-025-08597-9

[14] "NSFOCUS Monthly APT Insights - June 2025 - NSFOCUS, Inc., a global network and cyber security leader, protects enterprises and carriers from advanced cyber attacks." Accessed: Mar. 01, 2026. [Online]. Available: https://nsfocusglobal.com/nsfocus-monthly-apt-insights-june/

[15] Abdullah Said AL-Aamri, Rawad Abdulghafor, "Machine Learning for APT Detection," *Sustainabilitytain*, vol. 15, no. 18, p. 13820, 2023, doi: https://doi.org/10.3390/su151813820.

[16] Yang Hu, Roland Sandt & Robert Spatschek, "Practical feature filter strategy to machine learning for small datasets in chemistry," *Sci. Rep.*, 2024, [Online]. Available: https://www.nature.com/articles/s41598-024-71342-1

[17] Mingqi Lv, HongZhe Gao, Xuebo Qiu, Tieming Chen, Tiantian Zhu, Jinyin Chen, Shouling Ji, "TREC: APT Tactic / Technique Recognition via Few-Shot Provenance Subgraph Learning," *arXiv:2402.15147*, 2024, [Online]. Available:

https://arxiv.org/abs/2402.15147

[18] Animesh Singh Basnet, Mohamed Chahine Ghanem, Dipo Dunsin, Wiktor Sowinski-Mydlarz, "Advanced Persistent Threats (APT) Attribution Using Deep Reinforcement Learning," *arXiv:2410.11463*, 2025, doi: https://doi.org/10.48550/arXiv.2410.11463.

[19] Anh Tuan Le, Gregory Epiphaniou, "Automated APT Defense Using Reinforcement Learning and Attack Graph Risk-based Situation Awareness," *Auton. 2024 - Proc. Work. Auton. Cybersecurity, Co-Located with CCS 2024*, 2024, [Online]. Available: https://dl.acm.org/doi/10.1145/3689933.3690834

[20] Ameer A. Ghani, Suad A. Alasadi, "A Deep Learning Algorithm to Cybersecurity: Enhancing Intrusion Detection with a Hybrid GRU and BiLSTM Model," *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 3, pp. 23605–23612, 2025, doi: 10.48084/etasr.10666.

[21] Fargana J. Abdullayeva, "Advanced Persistent Threat attack detection method in cloud computing based on autoencoder and softmax regression algorithm," *Array*, vol. 10, p. 100067, 2021, doi: https://doi.org/10.1016/j.array.2021.100067.

[22] Mohammad Mamun, Kevin Shi, "DeepTaskAPT: Insider APT detection using Task-tree based Deep Learning," *arXiv:2108.13989*, 2021, [Online]. Available: https://arxiv.org/abs/2108.13989

[23] Zefeng He, Diego Davila, "Machine Learning for Cybersecurity: A Survey of Applications, Adversarial Challenges, and Future Research Directions," *Electronics*, vol. 14, no. 23, p. 4563, 2025, doi: 10.3390/electronics14234563.

[24] Santiago Quintero-Bonilla, Angel Martín del Rey, "A New Proposal on the Advanced Persistent Threat: A Survey," *Appl. Sci.*, vol. 10, no. 11, p. 3874, 2020, doi: https://doi.org/10.3390/app10113874.

[25] Anton Konev, Alexander Shelupanov, "A Survey on Threat-Modeling Techniques: Protected Objects and Classification of Threats," *Symmetry (Basel).*, vol. 14, no. 3, p. 549, 2022, doi: https://doi.org/10.3390/sym14030549.

[26] Jiajun Zhou, Jiacheng Yao, Xuanze Chen, Shanqing Yu, Qi Xuan, Xiaoniu Yang, "Lateral Movement Detection via Time-aware Subgraph Classification on Authentication Logs," *arXiv:2411.10279*, 2024, [Online]. Available: https://arxiv.org/abs/2411.10279

[27] T. Zhu *et al.*, "APTSHIELD: A Stable, Efficient and Real-Time APT Detection System for Linux Hosts," *IEEE Trans. Dependable Secur. Comput.*, vol. 20, no. 6, pp. 5247–5264, Nov. 2023, doi: 10.1109/TDSC.2023.3243667.

[28] Nur Ilzam Che Mat , Norziana Jamil , Yunus Yusoff , Miss Laiha Mat Kiah, "A systematic literature review on advanced persistent threat behaviors and its detection strategy," *J. Cybersecurity*, vol. 10, no. 1, 2024, doi: https://doi.org/10.1093/cybsec/tyad023.

[29] Bushra Sabir, Faheem Ullah, M. Ali Babar, Raj Gaire, "Machine Learning for Detecting Data Exfiltration: A Review," *arXiv:2012.09344*, 2021, [Online]. Available: https://arxiv.org/abs/2012.09344

[30] Qi Liu, Muhammad Shoaib, Mati Ur Rehman, Kaibin Bao, Veit Hagenmeyer, Wajih Ul Hassan, "Accurate and Scalable Detection and Investigation of Cyber Persistence Threats," *arXiv:2407.18832*, 2024, [Online]. Available: https://arxiv.org/abs/2407.18832

[31] Md Rayhanur Rahman, Setu Kumar Basak, Rezvan Mahdavi Hezaveh, Laurie Williams, "SoK: An empirical investigation of malware techniques in advanced persistent threat attacks," *Comput. Secur.*, vol. 157, p. 104618, 2025, doi: https://doi.org/10.1016/j.cose.2025.104618.

[32] Muhammad Shofian Tsauri, "Human Vulnerabilities to Social Engineering Attacks: A Systematic Literature Review for Building a Human Firewall," *J. Appl. Informatics Comput.*, vol. 9, no. 4, pp. 1127–1136, 2025, doi: 10.30871/jaic.v9i4.9585.

[33] Suleiman Y. Yerima, Mohammed K. Alzaylaee, "High Accuracy Phishing Detection Based on Convolutional Neural Networks," *arXiv:2004.03960*, 2020, [Online]. Available: https://arxiv.org/abs/2004.03960

[34] Santosh Kumar Birthriya, Priyanka Ahlawat, Ankit Kumar Jain, "Detection and prevention of spear phishing attacks: A comprehensive survey," *Comput. Secur.*, vol. 151, p. 104317, 2025, doi: https://doi.org/10.1016/j.cose.2025.104317.

[35] Singamaneni Krishnapriya, Sukhvinder Singh, "A Comprehensive Survey on Advanced Persistent Threat (APT) Detection Techniques," *Comput. Mater. Contin.*, vol. 80, no. 2, pp. 2675–2719, 2024, doi: https://doi.org/10.32604/cmc.2024.052447.

[36] Suresh Kumar Srinivasan, Sudalaimuthu Thalavaipillai, "Kernel rootkit detection multi class on deep learning techniques," *Bull. Electr. Eng. Informatics*, vol. 13, no. 3, pp. 2000–2008, 2024, doi: 10.11591/eei.v13i3.6802.

[37] Abdullah Al Mamun, Harith Al-Sahaf, "Detection of advanced persistent threat: A genetic programming approach," *Appl. Soft Comput.*, vol. 167, p. 112447, 2024, doi: https://doi.org/10.1016/j.asoc.2024.112447.

[38] "Synthetic APT Dataset." Accessed: Mar. 30, 2026. [Online]. Available: https://www.emergentmind.com/topics/synthetic-apt-dataset

[39] Marcos Luengo Viñuela, Jesús Ángel Román-Gallego, "Detection of APTs by Machine Learning: A Performance Comparison," *Expert Syst.*, vol. 43, no. 1, 2025, doi: 10.1111/exsy.70181.

[40] J. Liu *et al.*, "A New Realistic Benchmark for Advanced Persistent Threats in Network Traffic," *IEEE Netw. Lett.*, vol. 4, no. 3, pp. 162–166, Sep. 2022, doi: 10.1109/LNET.2022.3185553.

[41] Q. Ma and N. Rastogi, "DANTE: Predicting insider threat using LSTM on system logs," *Proc. - 2020 IEEE 19th Int. Conf. Trust. Secur. Priv. Comput. Commun. Trust. 2020*, pp. 1151–1156, Dec. 2020, doi: 10.1109/TrustCom50675.2020.00153.

[42] U. Sakthivelu, C. N.S. Vinoth Kumar, "Advanced Persistent Threat Detection and Mitigation Using Machine Learning Model," *Intell. Autom. Soft Comput.*, vol. 36, no. 3, pp. 3691–3707, 2023, doi: https://doi.org/10.32604/iasc.2023.036946.

[43] H. Shadabfar, M. Dehghan, and B. Sadeghian, "DSRL-APT-2023: A New Synthetic Dataset For Advanced Persistent Threats," *Vol. 17, Issue 2*, vol. 17, no. 2, pp. 107–116, Jan. 2025, doi: 10.22042/isecure.2025.214212.

[44] C. Catal, G. Giray, B. Tekinerdogan, S. Kumar, and S. Shukla, "Applications of deep learning for phishing detection: a systematic literature review," *Knowl. Inf. Syst.*, vol. 64, no. 6, pp. 1457–1500, Jun. 2022, doi: 10.1007/s10115-022-01672-x.

[45] Shuhui Zhang, Mingyu Gao, "A Malware-Detection Method Using Deep Learning to Fully Extract API Sequence Features," *Electronics*, vol. 14, no. 1, p. 167, 2025, doi: https://doi.org/10.3390/electronics14010167.

[46] Saba Aslam, Hafsa Aslam, "AntiPhishStack: LSTM-Based Stacked Generalization Model for Optimized Phishing URL Detection," *Symmetry (Basel).*, vol. 16, no. 2, p. 248, 2024, doi: https://doi.org/10.3390/sym16020248.

[47] P. Maneriker, J. W. Stokes, E. G. Lazo, D. Carutasu, F. Tajaddodianfar, and A. Gururajan, "URLTran: Improving Phishing URL Detection Using Transformers," *Proc. - IEEE Mil. Commun. Conf. MILCOM*, vol. 2021-November, pp. 197–204, 2021, doi: 10.1109/MILCOM52596.2021.9653028.

[48] A. Bensaoud, J. Kalita, and M. Bensaoud, "A survey of malware detection using deep learning," *Mach. Learn. with Appl.*, vol. 16, p. 100546, 2024, doi:

https://doi.org/10.1016/j.mlwa.2024.100546.

[49] Jess Hohenstein, Rene F. Kizilcec, Dominic DiFranzo, Zhila Aghajari, Hannah Mieczkowski, Karen Levy, Mor Naaman, "Artificial intelligence in communication impacts language and social relationships," *Sci. Rep.*, 2023, [Online]. Available: https://www.nature.com/articles/s41598-023-30938-9

[50] M. Odusami, S. Misra, O. Abayomi-Alli, A. Abayomi-Alli, and L. Fernandez-Sanz, "A survey and meta-analysis of application-layer distributed denial-of-service attack," *Int. J. Commun. Syst.*, vol. 33, no. 18, p. e4603, Dec. 2020, doi: 10.1002/DAC.4603.

[51] Yafei Song, Dandan Zhang, Jian Wang, Yanan Wang, Yang Wang, Peng Ding, "Application of deep learning in malware detection: a review," *J. Big Data*, vol. 12, no. 99, 2025.

[52] Howon Kim, Thi-Thu-Huong Le, "Machine Learning and Deep Learning Based Model for the Detection of Rootkits Using Memory Analysis," *Appl. Sci.*, vol. 13, no. 19, p. 10730, 2023, doi: https://doi.org/10.3390/app131910730.

[53] Max Landauer, Leonhard Alton, Martina Lindorfer, Florian Skopik, Markus Wurzenberger, Wolfgang Hotwagner, "Trace of the Times: Rootkit Detection through Temporal Anomalies in Kernel Activity," *arXiv:2503.02402*, 2025, [Online]. Available: https://arxiv.org/abs/2503.02402

[54] Sihat Afnan, Mushtari Sadia, Shahrear Iqbal, Anindya Iqbal, "LogShield: A Transformer-based APT Detection System Leveraging Self-Attention," *arXiv:2311.05733*, 2023, [Online]. Available: https://arxiv.org/abs/2311.05733

[55] Zian Jia, Yun Xiong, Yuhong Nan, Yao Zhang, Jinjing Zhao, Mi Wen, "MAGIC: Detecting Advanced Persistent Threats via Masked Graph Representation Learning," *arXiv:2310.09831*, 2023, [Online]. Available: https://arxiv.org/abs/2310.09831

[56] Huynh Thai Thi, Ngo Duc Hoang Son, Phan The Duy, Nghi Hoang Khoa, Khoa Ngo-Khanh, Van-Hau Pham, "XFedHunter: An Explainable Federated Learning Framework for Advanced Persistent Threat Detection in SDN," *arXiv:2309.08485*, 2023, [Online]. Available: https://arxiv.org/abs/2309.08485

[57] Hedyeh Nazari, Abbas Yazdinejad, Ali Dehghantanha, Fattane Zarrinkalam, Gautam Srivastava, "P3GNN: A Privacy-Preserving Provenance Graph-Based Model for APT Detection in Software Defined Networking," *arXiv:2406.12003*, 2024, [Online]. Available: https://arxiv.org/abs/2406.12003

[58] Sarah Mohammed Alshehri, Sanaa Abdullah Sharaf, "Systematic Review of Graph Neural Network for Malicious Attack Detection," *Information*, vol. 16, no. 6, p. 470, 2025, doi: https://doi.org/10.3390/info16060470.

[59] Weiwu Ren, Xintong Song, Yu Hong, Ying Lei, Jinyu Yao, Yazhou Du, Wenjuan Li, "APT Attack Detection Based on Graph Convolutional Neural Networks," *Int. J. Comput. Intell. Syst.*, vol. 16, no. 184, 2023, [Online]. Available: https://link.springer.com/article/10.1007/s44196-023-00369-5

[60] Cho Do Xuan, Tung Thanh Nguyen, "A novel approach for APT attack detection based on an advanced computing," *Sci. Rep.*, vol. 14, 2024, [Online]. Available: https://www.nature.com/articles/s41598-024-72957-0

[61] Nan Wang, Xuezhi Wen, Dalin Zhang, Xibin Zhao, Jiahui Ma, Mengxia Luo, Fan Xu, Sen Nie, Shi Wu, Jiqiang Liu, "TBDetector:Transformer-Based Detector for Advanced Persistent Threats with Provenance Graph," *arXiv:2304.02838*, 2023, [Online]. Available: https://arxiv.org/abs/2304.02838

[62] Miracle Udurume, Vladimir Shakhov, "Comparative Analysis of Deep Convolutional Neural Network—Bidirectional Long Short-Term Memory and Machine Learning Methods in Intrusion Detection Systems," *Appl. Sci.*, vol. 14, no. 16, p. 6967, 2024, doi: https://doi.org/10.3390/app14166967.

[63] M. C. Ali Hussein Ali, "Unveiling machine learning strategies and considerations in intrusion detection systems: a comprehensive survey," *Front. Comput. Sci.*, vol. 6, 2024, [Online]. Available: https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2024.1387354/full

[64] Sidahmed Benabderrahmane, Ngoc Hoang, Petko Valtchev, James Cheney, Talal Rahwan, "Hack Me If You Can: Aggregating AutoEncoders for Countering Persistent Access Threats Within Highly Imbalanced Data," *arXiv:2406.19220*, 2024, [Online]. Available: https://arxiv.org/abs/2406.19220

[65] Sidahmed Benabderrahmane, Petko Valtchev, James Cheney, Talal Rahwan, "APT-LLM: Embedding-Based Anomaly Detection of Cyber Advanced Persistent Threats Using Large Language Models," *arXiv:2502.09385*, 2025, [Online]. Available: https://arxiv.org/abs/2502.09385

[66] Toya Acharya, Annamalai Annamalai, "Enhancing the Network Anomaly Detection using CNN-Bidirectional LSTM Hybrid Model and Sampling Strategies for Imbalanced Network Traffic Data," *Adv. Sci. Technol. Eng. Syst. J.*, vol. 9, no. 1, pp. 67–78, 2024, doi: 10.25046/aj090107.

[67] Atmane Ayoub Mansour Bahar, Kamel Soaid Ferrahi, Mohamed-Lamine Messai, Hamida Seba, Karima Amrouche, "CONTINUUM: Detecting APT Attacks through Spatial-Temporal Graph Neural Networks," *arXiv:2501.02981*, 2025, [Online]. Available: https://arxiv.org/abs/2501.02981