electronics

MDPI

Article

Intelligent Detection of Cyber Attack Patterns in Industrial IoT Using Pretrained Language Models

Yifan Liu, Shancang Li * and Sarah Bin Hulayyil

School of Computer Science and Informatics, Cardiff University, Cardiff CF24 4AG, UK
* Correspondence: shancang.li@ieee.org

Abstract

Industrial Internet of Things (IIoT) systems are increasingly exposed to sophisticated and rapidly evolving cyber threats. In response, this work proposes a proactive threat detection framework that leverages pretrained transformer-based language models to identify emerging attack patterns within IIoT ecosystems. This work introduces a transformer-based framework that fine-tunes domain-specific pretrained models (SecBERT, SecRoBERTa, CyBERT), derives potential attack-path patterns from vulnerability–tactic mappings, and incorporates a retrieval-based fallback mechanism. The fallback not only improves robustness under uncertainty, but also provides a practical solution to the absence of labeled datasets linking ICS-specific MITRE ATT&CK tactics with vulnerabilities, thereby filling a key research gap. Experiments show that the fine-tuned models substantially outperform traditional machine learning baselines; SecBERT achieves the best balance while maintaining high inference efficiency. Overall, the framework advances vulnerability-driven threat modeling in IIoT and offers a foundation for the proactive identification of attack patterns.

Keywords: detection of cyber attack patterns; IIoT; language model; secure digital service ecosystems

Academic Editor: Aryya Gangopadhyay

Received: 5 September 2025
Revised: 5 October 2025
Accepted: 10 October 2025
Published: 18 October 2025

Citation: Liu, Y.; Li, S.; Hulayyil, S.B. Intelligent Detection of Cyber Attack Patterns in Industrial IoT Using Pretrained Language Models. *Electronics* **2025**, *14*, 4094. https://doi.org/10.3390/electronics14204094

1. Introduction

The Industrial Internet of Things (IIoT) has undergone significant advancement in recent years [1]. IIoT deployments typically consist of a large number of low-power sensors, controllers, and wireless communication modules that continuously monitor and interact with physical processes. The progress in low-power and low-complexity communication technologies has greatly facilitated scalable and energy-efficient connectivity, ensuring reliable operations in resource-constrained environments [2,3]. Enabled by these developments, IIoT has become a critical driver for the digitalization, automation, and intelligence of industrial control systems (ICS). However, the interconnection of heterogeneous devices also introduces new vulnerabilities, which make IIoT systems increasingly exposed to cybersecurity threats [4].

Cyber attacks have become increasingly sophisticated in recent years [5]. With the rapid development of industrial IoT, security threats at the network, software, and system layers have become increasingly prominent in ICS. Kaspersky indicated that 21.9% ICS computers were attacked in the first quarter of 2025 [6]. According to the report of ENISA (European Union Agency for Cybersecurity), exploits will increase significantly after the publication of the vulnerability [7]. According to Claroty, about 12% of operational technology (OT) devices were found to have known exploited vulnerabilities. Such weaknesses

represent a significant risk, as they may directly lead to the disruption of critical industrial processes [8].

Inferring attack patterns from known vulnerabilities is a common practice in threat analysis [9,10]. Mitre Att&CK provides a framework for potential attack technologies and corresponding tactics of malicious activity. Therefore, based on known vulnerabilities and the corresponding attack methods, potential attacks in the system could be predicted. Then, attack paths could be inferred. In addition, the Common Vulnerability Scoring System (CVSS) score and the Exploit Prediction Scoring System (EPSS) score from CVE can be used to evaluate the probability of an attack, allowing more targeted defensive measures or more efficient response and investigation of attack incidents [11,12].

As each CVE may involve multiple attack tactics, it is therefore a multi-label classification problem that identifies the tactics according to the description of the CVE [13]. Building on this idea, this paper creates a multilabel text classification model that accurately determines how a CVE might be exploited based on its description. The dataset was collected from the BRON database, which contains recent CVEs and records related to MITRE ATT&CK Tactics. The collected data was tokenized and processed using SecBERT [14], SecRoBERTa, and CyBERT [15], and then used to train the classifier. MITRE ATT&CK Tactics were used as a reference to predict potential attack paths or strategies.

Moreover, since ICS-related samples in existing datasets are sparse and highly imbalanced, the classifier may fail to provide accurate predictions for rare classes or tactics absent from the training set. To mitigate this limitation, prediction confidence is estimated using probability distribution and entropy. Samples with insufficient confidence are handled by a retrieval-based fallback module, which measures semantic similarity between the CVE description and tactic descriptions in the knowledge base.

The main contributions of this paper are summarized as follows:

(1) A CVE-to-MITRE ATT&CK tactic mapping method was proposed that jointly considers both ICS and Enterprise contexts, enabling more comprehensive coverage of vulnerabilities and attack tactics.
(2) A retrieval-based similarity search mechanism was incorporated to address the lack of labelled ICS-specific samples, providing a fallback solution that improves robustness under data scarcity.
(3) By leveraging vulnerability—MITRE ATT&CK mappings, a lightweight approach for detecting potential attack patterns was developed, supporting efficient identification of threats in industrial IoT environments.

By analyzing existing vulnerabilities using pretrained models, this work aims to identify and characterize attack tactics and patterns within industrial IoT (IIoT) scenarios. This proactive approach can facilitate the timely detection, response, and prevention of potential cyber attacks targeting critical infrastructure and operations.

2. Related Works

The integration of vulnerability management and threat management is critical for risk assessment. Feng et al. provided a systematic and comprehensive survey of vulnerabilities in IoT device firmware and their detection methods [16], which could be a feasible foundation for vulnerability-driven threat analysis. Understanding how adversaries exploit vulnerabilities can greatly assist defenders in establishing robust defensive measures. A key contribution in this topic is the methodology developed by the MITRE Center for mapping CVEs to the ATT&CK framework. This approach provides a structured way to contextualize the technical details of a vulnerability with the adversary behaviours outlined in ATT&CK. The methodology breaks down the exploitation process into three key categories, exploitation technique, primary impact, and secondary impact. By using

this template, security analysts can translate a CVE description into a narrative that aligns with known adversary tactics and techniques [17]. Tatam et al. discussed using the MITRE ATT&CK matrix combined with graph-based methods for modelling threats like Advanced Persistent Threats. They highlighted that in current research, the process of identifying threats based on vulnerabilities is still a manual process [18].

While this manual methodology provides a strong foundation for vulnerability and MITRE ATT&CK mapping, other related works have focused on different aspects of this field. Hemberg et al. proposed the BRON framework to trace relationships among records drawn from various vulnerability and attack information sources. An example illustrates relationships among CVE, CWE, CAPEC (Common Attack Pattern Enumeration and Classification), MITRE techniques, and MITRE tactics. This framework can serve as a foundation for constructing datasets to train language models that identify the corresponding MITRE tactics and techniques for each CVE. However, as this project relies on publicly available data, it may suffer from issues such as data measurement bias and update latency [19]. Some studies, on the other hand, emphasize the analysis of attack paths. Hankin et al. proposed a framework to assess potential risks based on system vulnerabilities. Their approach also relies on manually and incrementally establishing mappings among CVE, CWE, and CAPEC to infer potential attacks [5]. However, manual analysis requires specialized knowledge and usually comes with a high cost in time.

With the rapid advancement of AI technologies, machine learning (ML)-based methods have been introduced. Lakhdhar evaluated a set of algorithms (BinaryRelevance, LabelPowerset, ClassifierChains, MLKNN, BRKNN, RAkELd, NLSP, and Neural Networks), and acheived 99% accuracy in CVE classification task [20]. However, the ML methods highly depend on manual feature engineering, which is time consuming. Compared with traditional machine learning, the rise of NLP techniques has enabled end-to-end approaches. Grigorescu et al. presented a dataset comprising 1813 CVEs along with their corresponding attack techniques. They also introduced models mapping CVEs to techniques and employed LIME to enhance interpretability [21]. This work demonstrated the capabilities of language models in addressing the CVE classification. However, the limited dataset size and severe class imbalance significantly constrained model performance.

Language models begin to show improved performance when trained on sufficiently large datasets. Building upon previous work, Branescu et al. published a new dataset that extended the study by Grigorescu et al. [21]. They conducted experiments using transformer-based models, including CyBERT, SecBERT, SecRoBERTa, and TARS. SecRoBERTa achieved the best performance, with an F_1 score of 78.88%, 82.09% of precision, and 77.02% in recall. Moreover, GPT-4 was employed in this task but demonstrated relatively weaker performance [22]. However, due to the technical characteristics of certain attack tactics, the issue of sample imbalance persists. Existing ML-based methods still have room for improvement in terms of accuracy.

3. Proposed Methodology

As shown in Figure 1, the proposed method consists of two main components, CVE-Tactic mapping and potential attack pattern construction.

In the CVE-Tactic mapping module, the vulnerability descriptions were tokenized and processed using cybersecurity-oriented pretrained language models (CyBERT, SecBERT, and SecRoBERTa). The fine-tuned classifiers output probabilities for multiple MITRE ATT&CK tactics in a multi-label setting. To address the limitation that the training data only contains Enterprise tactics, an uncertainty estimation mechanism was introduced. When predictions are unreliable (low maximum probability or high entropy), the system triggers a retrieval-based fallback module, which compares CVE embeddings against both

Enterprise and ICS tactic representations to extend the prediction space and improve robustness in IIoT environments.

The predicted tactics are integrated with the combined MITRE ATT&CK Enterprise and ICS matrices, together with the network topology, to infer potential attack paths and patterns relevant to industrial IoT systems.

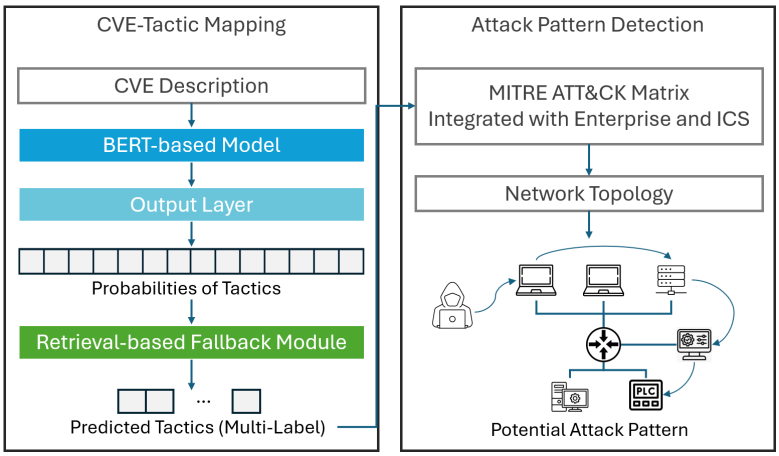


Figure 1. Attack pattern detection workflow.

3.1. Data Collection

The CVE-Tactic data is collected from the BRON database. It includes the CVE from 1999 to 2024. It does not provide a direct mapping between CVEs and tactics; instead, such mappings must be inferred through a chain of indirect dependencies. It links CVEs and tactics via a series of key-value identifiers, including mappings from CVEs to CWEs, CWEs to CAPECs, CAPECs to techniques, and techniques to tactics. Through this structure, an indirect mapping between CVEs and tactics can be constructed.

The processed dataset consists of three columns: *CVE-ID*, CVE-Description, and Tactics. For instance, CVE-2023-3519 is described as "Unauthenticated remote code execution" and involves four tactics: initial access, execution, privilege escalation, and lateral movement. After removing duplicates and invalid entries, the dataset contains 43,491 samples. However, the tactic exfiltration is entirely absent, while command and control and execution appear in only 9 and 95 samples, respectively, indicating a severe class imbalance.

To supplement missing labels and mitigate imbalance, a public dataset by Branescu et al. [22] was integrated with BRON data. This combined dataset incorporates records from ENISA, MITRE Engenuity, and manually labelled samples by Grigorescu et al. [21], followed by data cleaning to remove duplicates and invalid entries.

In the IIoT context, many IT techniques are adopted in ICS, with Enterprise technologies often forming their foundation [23]. Due to limited mapping data for the ICS matrix, available data primarily reflects the Enterprise matrix. Table 1 summarizes the distribution of tactics in the dataset.

Figure 2 presents the distribution of CVEs mapped to different MITRE ATT&CK tactics. The tactics with the highest number of CVEs are credential access, privilege escalation, and defense evasion, each with around 30,000 to 40,000 samples. discovery and lateral movement also show significant numbers of samples, whereas tactics such as exfiltration, command and control, and resource development have relatively few associated CVEs.

Table 1. Involved MITRE ATT&CK Tactics.

Attack Tactic	Number of Samples
Reconnaissance	5607
Resource Development	689
Initial Access	3306
Execution	2527
Persistence	30,666
Privilege Escalation	30,696
Defense Evasion	40,527
Credential Access	24,753
Discovery	13,315
Lateral Movement	13,869
Collection	7029
Command and Control	430
Exfiltration	170
Impact	3647

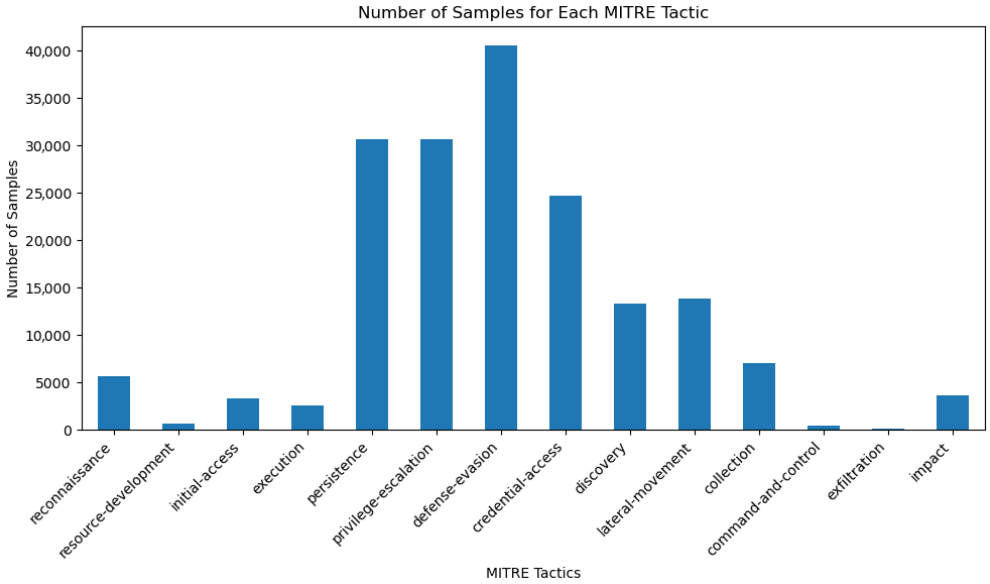


Figure 2. Tactic number distribution.

Meanwhile, Figure 3 shows the average token number of the samples and the word count of the CVE descriptions. Because of the hardware limitations of the experiment, the length of the description text was capped at 256 words. It can be observed that most descriptions contain fewer than 100 words, and descriptions around 50 words are the most common in CVE. The higher frequency of descriptions containing approximately 250 words is primarily due to the manual truncation of texts which exceeded the 256-word limit in the original document.

The average length of CVE descriptions associated with different attack tactics is similar, ranging between 60 and 80 words, as shown in Figure 4. However, the descriptions of CVEs related to the tactics' impact, execution, and exfiltration are noticeably longer compared to other categories. In particular, the number of tokens of CVE description related to execution and exfiltration is nearly one-third higher than those of other categories.

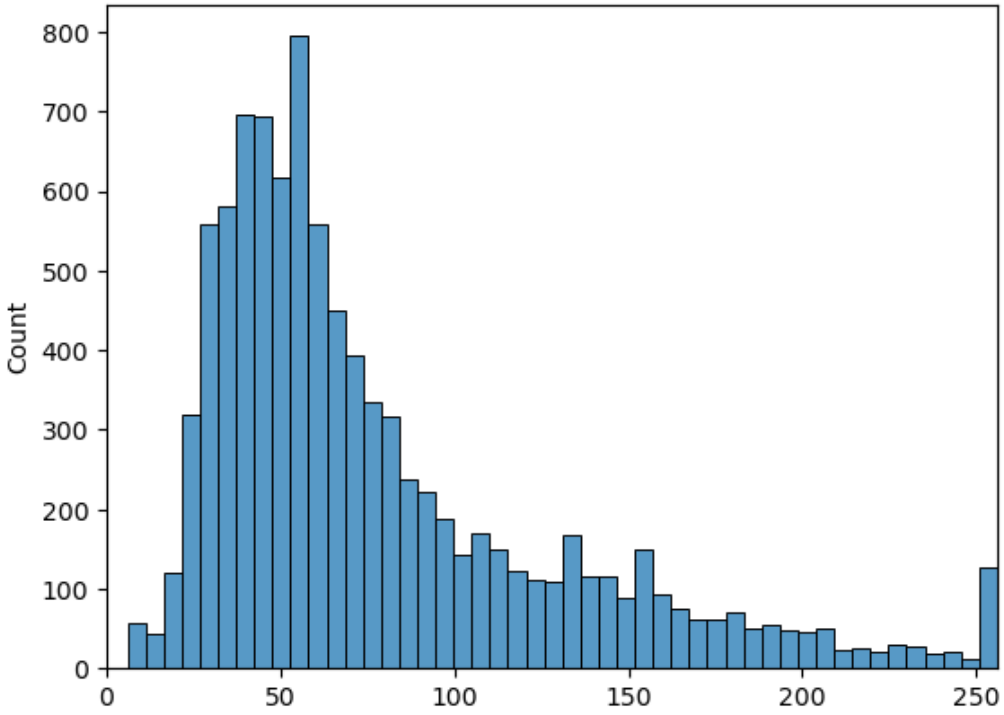


Figure 3. Token number distribution.

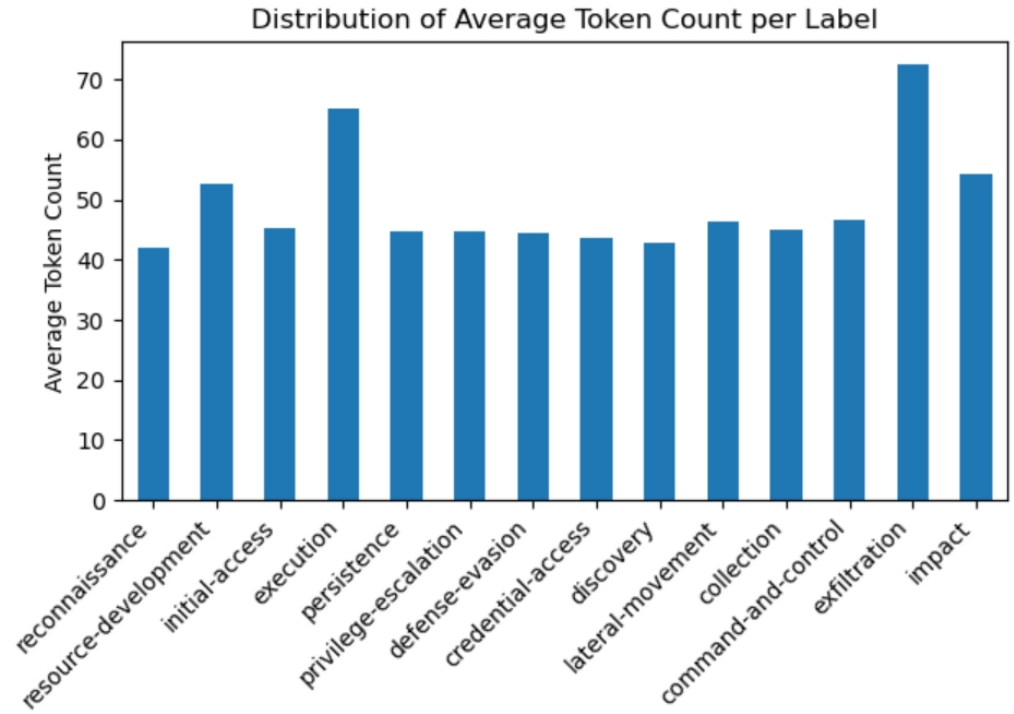


Figure 4. Average token numbers of labels.

3.2. CVE-Tactic Mapping

3.2.1. BERT Models

The BERT-based models could be regarded as a set of BERT layers which have the structure of a transformer. In this paper, there are six BERT layers in the model, as is shown in Figure 5. Each BERT layer contains 12 attention heads. These attention heads allow the model to focus on different parts of the input sequence simultaneously, capturing complex

relationships between tokens across the text. The multi-head architecture enables the model to effectively calculate the context and semantics information of a given text.

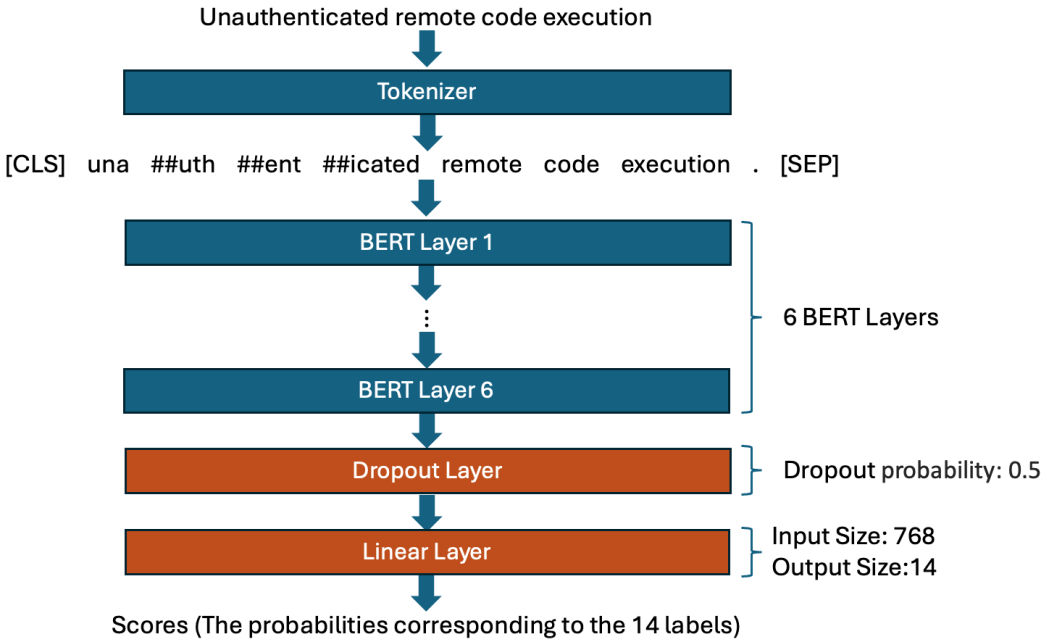


Figure 5. CVE Classification model structure.

The tokenization process was enabled by WordPiece tokenization. It will split sentence into smaller subword units. The tokenizer will attempt to match each word in the volcabulary; if a word is not found, it will split the unknown word into subword units, as shown in Figure 5.

To perform multi-label classification on our dataset, three pretrained transformer models, SecBERT, SecRoBERTa, and CyBERT were selected as base models. These models have been pretrained on large-scale cybersecurity corpora, including vulnerability descriptions and threat intelligence reports, enabling them to capture domain-specific terminology and semantics that are highly relevant to the task of mapping CVEs to the MITRE ATT&CK matrix. Moreover, they have been widely adopted in existing studies and have demonstrated strong potential in various cybersecurity-related applications.

The fine-tuning framework is constructed by extending pretrained transformer encoders with an additional classification layer. The pretrained models function as feature extractors that capture contextual embeddings from the input text. On top of these embeddings, a fully connected classification layer is applied to map the learned representations to the target classes. This architecture leverages domain-specific knowledge acquired during pretraining while adapting the model to the multi-label classification task.

Full fine-tuning updates all the parameters of the pretrained BERT model during training. This process involves adding a classifier layer on top of the pretrained architecture and then training the entire model end-to-end on the new dataset. This allows the model to adapt its internal representations to the specific nuances of the downstream task, leading to a higher level of performance and more accurate results [24].

3.2.2. Multi-Label Classifier

It can be seen from Figure 5 that two additional layers, a dropout layer and a linear layer, were employed after transformer layers, with the output layer consisting of 14 units. The dropout layer is used to mitigate the over fitting by ignoring some hidden nodes, though some information will be lost, which could potentially compromise the model's

performance. The dropout rate was set as 0.5; it will randomly drop out half of the input from the transformer during the training process. The linear layer is used to process the output of transformer layers, and generate scores for each label, namely, 14 different tactics. To predict each class independently, the final output goes through a sigmoid function, allowing each probability to range independency between 0 and 1. The prediction threshold of the sigmoid function was 0.5 to decide if a class is present or absent. During the training and evaluation, the code compares the threshold predictions with the true multiple label vectors to calculate the number of correct predictions for each class per sample.

3.3. Retrieval-Based Fallback

Since the training dataset only contains MITRE ATT&CK Enterprise tactics, the classifier is limited to predicting 14 categories. Although there is partial overlap with the ICS matrix, three ICS-specific tactics cannot be captured by the model. To address this limitation and enable the analysis of unlabeled data in the industrial IoT context, a retrieval-based similarity mechanism was incorporated. This approach complements the classifier by comparing CVE descriptions with both Enterprise and ICS tactic representations, thereby extending the prediction space and improving adaptability in IIoT environments.

Uncertainty estimation. For each input CVE description x, the classifier produces probability outputs for the 14 Enterprise tactics:

$$\mathbf{P}(x) = (p_1, p_2, \ldots, p_{14}), \quad p_i \in [0, 1]. \tag{1}$$

Two indicators are used to assess prediction reliability:

$$p_{\max}(x) = \max_i p_i \tag{2}$$

$$H(x) = \frac{1}{14} \sum_{i=1}^{14} \Big(- p_i \log p_i - (1 - p_i) \log(1 - p_i) \Big) \tag{3}$$

Equation (2) refers to the max probability of the model prediction output vector. This criterion detects cases where the classifier is not confident about any single tactic, i.e., all predicted probabilities are relatively low. A small $p_{\max}(x)$ indicates that the model does not strongly support any class.

Equation (3) refers to the mean entropy of prediction. It reflects the uncertainty of a Bernoulli distribution for each label. Averaging across all labels provides a global indicator of prediction dispersion. A high $H(x)$ suggests that the model assigns near-random probabilities (e.g., close to 0.5) to many classes simultaneously, meaning it is indecisive.

A sample is regarded as uncertain when

$$p_{\max}(x) < \theta_p \quad \text{or} \quad H(x) > \theta_H \tag{4}$$

In this case, the classifier output is discarded and a retrieval-based fallback triggered.

Similarity-based retrieval. For fallback, the CVE description is encoded into an embedding vector $\mathbf{e}(x)$ using a sentence transformer. Each tactic t_j (including the 17 tactics of Enterprise and ICS) is represented by an embedding $\mathbf{e}(t_j)$. The cosine similarity is computed as

$$s(x, t_j) = \frac{\mathbf{e}(x) \cdot \mathbf{e}(t_j)}{\|\mathbf{e}(x)\| \, \|\mathbf{e}(t_j)\|} \tag{5}$$

In addition, each technique k under tactic t_j is represented as $\mathbf{e}(k)$. Its similarity to the input is increased in the parent tactic by a factor λ (it was set to 0.1):

$$\tilde{s}(x, t_j) = \max\left(s(x, t_j), \max_{k \in \mathcal{K}(t_j)} \left(s(x, k) + \lambda\right)\right) \tag{6}$$

where $\mathcal{K}(t_j)$ is the set of techniques associated with tactic t_j.

Decision rule. The final prediction set under retrieval fallback is determined by thresholding:

$$\hat{Y}(x) = \{\, t_j \mid \tilde{s}(x, t_j) \geq \tau \,\} \tag{7}$$

where τ is a similarity threshold. For evaluation, only the 14 Enterprise tactics are retained, while the complete 17-dimensional retrieval scores are preserved for the analysis of ICS-specific behaviours.

3.4. Identify Attack Paths

The reconstruction attack path focuses on combining detected attacks with attack path prediction according to available vulnerability exploitation existing in the system. It has been indicated that exploits tend to increase after vulnerability publication. Hackers usually invade a system by exploiting known vulnerabilities. Moreover, the release of patch solutions corresponding to CVEs typically requires a certain amount of time. This interval allows attackers to exploit the vulnerability to inflict damage on the target system.

Figure 6 represents the process of the dependence inference framework. Firstly, the vulnerability scanning schemes could be employed to identify existing CVEs in each device of the system by analyzing the system information such as version, configuration, and installed software. Then mapping the recognized CVEs into correlated tactics can be conducted, therefore revealing the tactics available for each device. By converting the network topology of the system and the available tactics on each device into a matrix format, and using the MITRE ATT&CK Matrix as a reference, it is possible to calculate potential attack paths that could be executed within the system. This approach can generate an attack dependency inference model for target system.

Based on the known CVEs or vulnerabilities in the system, the potential attack path could be constructed by the following steps.

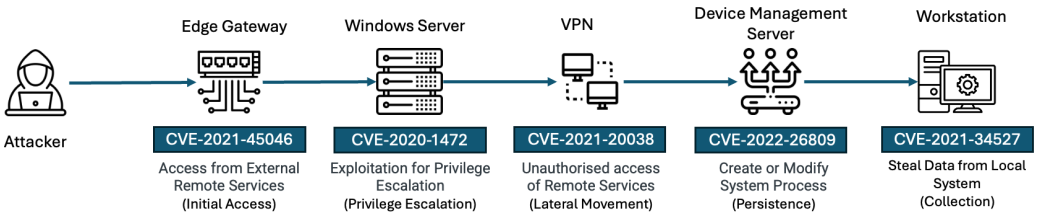


Figure 6. Attack path reconstruction in a smart home scenario.

In an IoT system containing $D = \{d_1, d_2, \ldots, d_n\}$ devices, d_n represents the device, and n refers to the number of devices. The connection statement of devices can be denoted as

$$C = \{(d_i, d_j) \| \; d_i \; is \; connected \; to \; d_j\} \tag{8}$$

Let $V = \{v_1, v_2, \ldots, v_m\}$ denote the CVE set, which includes m CVEs. For a device d_i, it may have multiple vulnerabilities v_j^i; a relationship set to describe the relationship between CVEs and devices could be represented as

$$R_{DV} = \{(d_i, v_j) | \; d_i \; has \; v_j\} \tag{9}$$

We use $T = \{t_1, t2, \ldots, t_k\}$ to present a tactic set, and then the relationship between vulnerability and tactic can be represented as

$$R_{VT} = \{(v_j, t_l) |\ v_j \textit{ is associated with } t_l\} \tag{10}$$

As a result, the possible tactics for device d_i can be derived based on the vulnerabilities associated with that device, and we have

$$R_{DT} = \{(d_i, t_l) |\ \exists\ v_j, (d_i, v_j) \in R_{DV} \& (v_j, t_l) \in R_{VT}\} \tag{11}$$

The potential attack path between d_{start} and d_{end} could be defined as

$$P_{start\_end} = \{(d_i, t_l) \cup P_{start\_end} | i = 1 \ldots n, l = 1 \ldots m\} \tag{12}$$

in which n is the number of devices involved, and m is the number of tactics that may be exploited for tactic t_l over device d_i. $(d_i, t_l) \in R_{DT}$ indicate device d_i may be exploited for tactic t_l. If there exists $(d_i, d_j) \in C$ and $(d_i, v_j^i) \in R_{DV}$ such that $(v_j^i) \in R_{VT}$, then a potential attack path exists from d_i to d_j.

Figure 6 shows an example. Assume that an IIoT system includes an edge gateway, Windows server, VPN gateway, device management server, and workstations, and for each device there exists a CVE. The edge gateway involves CVE-2021-45046, which can be exploited for initial access, allowing attackers to execute malicious code on the device using designed HTTP requests to gain access. The Windows server has CVE-2020-1472, which allows attackers to bypass authentication and log in with administrative privileges. This CVE is related to the Privilidge Escalation tactic in the MITRE ATT&CK Tactics.

The VPN gateway has CVE-2021-20038, a remote code execution vulnerability that lets attackers gain control of the device, enabling lateral movement across different networks. The device management server has CVE-2022-26809, which allows attackers to run malicious code when they gain sufficient privileges. It can be exploited to implant backdoor programmes and achieve persistence. The workstation has CVE-2021-34527, a vulnerability of Windows Print Spooler service, which can be utilized to execute code on the target system and access sensitive information. Attackers may use this CVE to steal or tamper with important data.

4. Experimental Results

The experiment is evaluated on an Ubuntu 22.04 server, with i7-12700k, 32 GB RAM, RTX3080 10 GB, CUDA 12.1.

In experiments, the improved models (SecBERT, SecRoBERTa, and CyBERT) with an additional dropout layer and linear classification layer achieved better performance in terms of precision, recall, and F_1-score compared with their original counterparts. A comparative analysis was conducted between BERT-based models and traditional ML approaches. As summarized in Table 2, BERT-based models substantially outperformed traditional classifiers in vulnerability-to-tactic mapping, with SecBERT providing the most balanced and robust results across all metrics. As shown in Figure 7, the improved SecBERT model converged around 15 epochs, with training loss steadily decreasing, while validation loss initially dropped slightly but later increased and plateaued, suggesting potential overfitting.

Figure 8 summarizes the inference latency and throughput of BERT-based models and traditional ML classifiers on the same hardware. The results show that SecBERT and CyBERT achieve the fastest inference, with an average latency of less than 1 ms per sample and throughput exceeding 1000 samples/s. SecRoBERTa is significantly slower, requiring on average 5 ms per sample (199 samples/s). Among the traditional classifiers, XGBoost

demonstrates competitive efficiency with 4.46 ms latency and 224 samples/s throughput, while Random Forest is slower (30.62 ms, 32.66 samples/s) and LightGBM shows the poorest efficiency (310 ms latency, only 3.23 samples/s). Overall, BERT-based models not only achieve higher classification accuracy, but also deliver superior or comparable inference efficiency, with SecBERT offering the best balance between accuracy and deployment efficiency. Moreover, the SecBERT-based fine-tuned SecBERT model requires around 5 GB memory during training and around 1.2 GB during inference, with an on-disc model size of 457 MB. Compared with large language models, this footprint is substantially smaller, resulting in lower compute and memory pressure when deployed on IIoT servers or workstations.

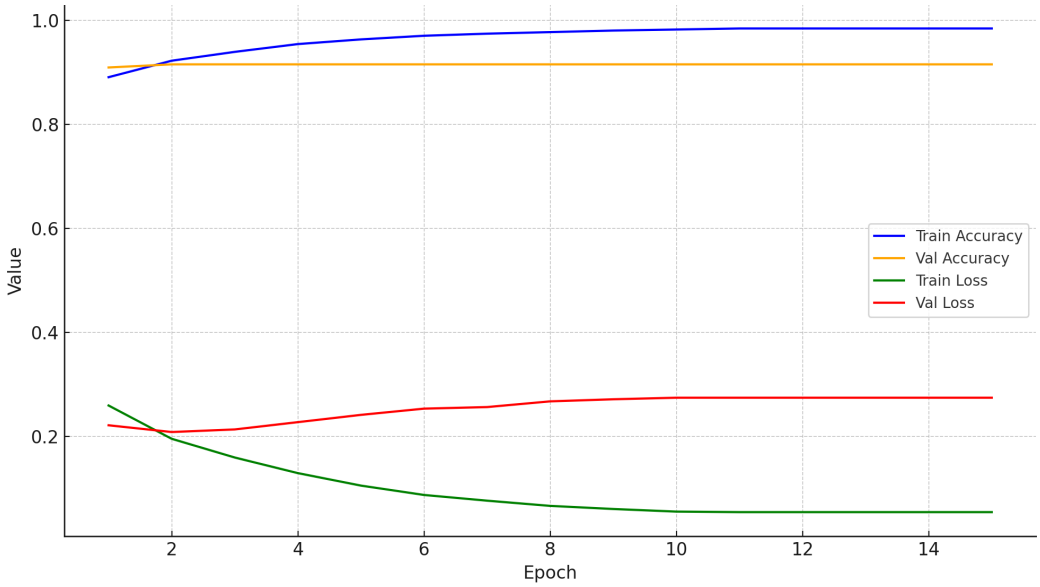
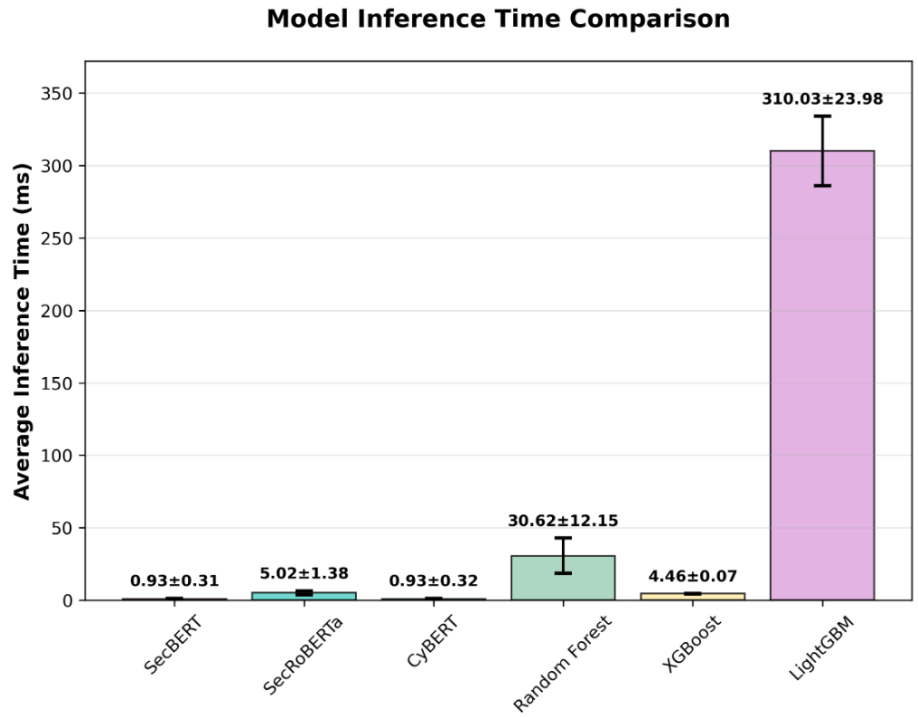


Figure 7. Accuracy and loss in the training process.

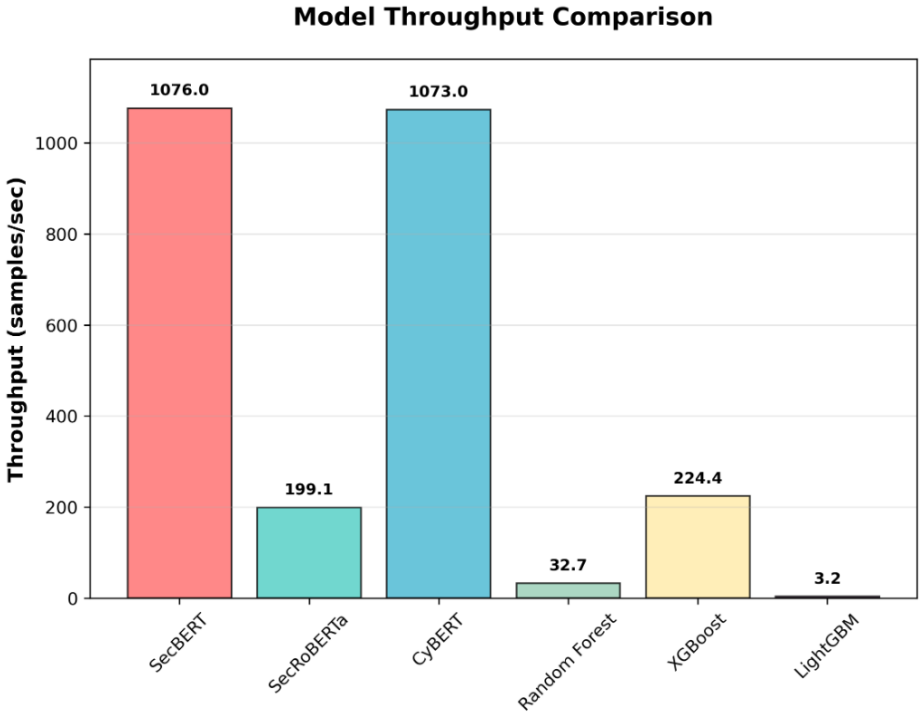
Figure 7 illustrates the training and validation accuracy of the improved models. Overall, the fine-tuned SecBERT achieved the best performance on this dataset, reaching 81.70% in micro precision, 84.91% in micro recall, and 83.27% in micro F_1-score. Table 3 presents the detailed performance per tactic. The model demonstrates strong precision across most tactics, particularly for persistence, privilege escalation, defense evasion, and credential access, where performance exceeded 90%. However, recall varied considerably among different tactics. The recall of command and control was only 25%, which is notably low. Many of these samples were misclassified into semantically related tactics such as credential access, persistence, defense evasion, and privilege escalation. This is likely due to class imbalance: the relatively small number of command and control samples limited the model's ability to capture distinctive contextual patterns, causing it to confuse them with tactics of higher textual similarity.

Table 2. Model performance comparison.

Model	Micro Precision	Micro Recall	Micro F_1
SecBERT	81.70%	84.91%	83.27%
SecRoBERTa	81.29%	84.50%	82.86%
CyBERT	81.31%	80.85%	81.08%
Random Forest	76.20%	71.52%	73.79%
XGBoost	78.38%	75.55%	76.94%
LightGBM	77.96%	75.36%	76.64%



(**a**) Model inference time comparison



(**b**) Model throughput comparison

Figure 8. Deployment performance comparison.

Table 3. Performance on each tactic prediction.

Tactic	Precision	Recall	F_1
reconnaissance	92.11%	90.28%	91.18%
resource-development	85.45%	63.51%	72.87%
initial-access	80.95%	74.16%	77.41%
execution	89.69%	72.80%	80.37%
persistence	94.43%	94.77%	94.60%
privilege-escalation	94.57%	94.30%	94.43%
defense-evasion	96.70%	97.90%	97.30%
credential-access	92.60%	93.46%	93.03%
discovery	91.95%	89.86%	90.89%
lateral-movement	91.94%	87.52%	89.67%
collection	90.05%	85.30%	87.61%
command-and-control	88.89%	25.00%	39.02%
exfiltration	92.31%	63.16%	75.00%
impact	91.21%	75.96%	82.89%

5. Conclusions

This work demonstrates the effectiveness of leveraging pretrained language models for mapping CVE descriptions to MITRE ATT&CK tactics, thereby enabling systematic analysis of potential attack vectors in industrial IoT environments. By constructing a cross-domain dataset that links CVEs with adversarial tactics and TTPs, we provided a solution for predictive analytics in IIoT threat modelling and proactive defense.

Through comprehensive experiments, the fine-tuned domain-specific models—SecBERT, SecRoBERTa, and CyBERT—significantly outperformed traditional ML baselines in vulnerability-to-tactic classification. Among them, SecBERT consistently delivered the best balance between accuracy and inference efficiency, achieving superior micro precision, recall, and F_1-score, while also maintaining sub-millisecond latency and high throughput during deployment. This work provided a potential solution for unlabeled vulnerability to MITRE ATT&CK tactic mapping, mitigating the requirement of manual analysis.

Nevertheless, the present study has certain limitations. The dataset suffers from class imbalance, which negatively impacts recall for tactics with fewer samples such as Command and Control. Moreover, some ICS-specific tactics remain underrepresented, limiting generalizability to broader IIoT contexts. In particular, there is still a notable gap in publicly available, large-scale datasets that explicitly map vulnerabilities (e.g., CVE/CWE) to ICS tactics/techniques. And terse or sparsely informative vulnerability descriptions provide limited context, which makes reliable tactic classification difficult. Due to operational safety, confidentiality, and under-reporting of incidents, it is difficult to obtain large volumes of real, high-fidelity ICS data.

Future work will address these limitations by constructing a knowledge base through the integration of large language models and threat intelligence reports to better classify unlabeled data, employing synthetic text generation and data augmentation techniques to alleviate class imbalance, and extending the framework to incorporate retrieval-based fallback mechanisms within real-time monitoring pipelines. These directions are expected to enhance both the robustness and adaptability of the proposed system in practical IIoT environments.

Author Contributions: Conceptualization, Y.L. and S.L.; methodology, Y.L. and S.L.; software, Y.L.; validation, Y.L., S.L. and S.B.H.; formal analysis, Y.L. and S.L.; writing, review and editing, Y.L., S.L. and S.B.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Dataset available on request from the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Khujamatov, H.; Reypnazarov, E.; Khasanov, D.; Akhmedov, N. IoT, IIoT, and cyber-physical systems integration. In *Emergence of Cyber Physical System and IoT in Smart Automation and Robotics: Computer Engineering in Automation*; Springer: Cham, Switzerland, 2021; pp. 31–50.
2. Yin, M.; Wei, C.; Takeda, K.; Jia, Y.; Xu, C.; Zhang, C.; Xu, H. Low-Complex Waveform, Modulation and Coding Designs for 3GPP Ambient IoT. *IEEE Commun. Stand. Mag.* **2025**. [CrossRef]
3. Ma, H.; Tao, Y.; Fang, Y.; Chen, P.; Li, Y. Multi-Carrier Initial-Condition-Index-aided DCSK Scheme: An Efficient Solution for Multipath Fading Channel. *IEEE Trans. Veh. Technol.* **2025**, *74*, 15743–15757. [CrossRef]
4. Mantravadi, S.; Schnyder, R.; Møller, C.; Brunoe, T.D. Securing IT/OT Links for Low Power IIoT Devices: Design Considerations for Industry 4.0. *IEEE Access* **2020**, *8*, 200305–200321. [CrossRef]
5. Sönmez, F.Ö.; Hankin, C.; Malacaria, P. Attack dynamics: An automatic attack graph generation framework based on system topology, CAPEC, CWE, and CVE databases. *Comput. Secur.* **2022**, *123*, 102938. [CrossRef]
6. Kaspersky. Threat Landscape for Industrial Automation Systems. Q1 2025. 2025. Available online: https://ics-cert.kaspersky.com/publications/reports/2025/05/15/threat-landscape-for-industrial-automation-systems-q1-2025/ (accessed on 28 September 2025).
7. Katos, V.; Rostami, S.; Bellonias, P.; Davies, N.; Kleszcz, A.; Faily, S.; Arnolnt, S.; Alexandros, P. *State of Vulnerabilities 2018/2019: Analysis of Events in the Life of Vulnerabilities*; Report/Study; European Union Agency for Cybersecurity (ENISA): Attiki, Greece, 2019.
8. Claroty. State of CPS Security 2025: OT Exposures. 2025. Available online: https://claroty.com/resources/reports/state-of-cps-security-ot-exposures-2025 (accessed on 28 September 2025).
9. Liu, Y.; Li, S. Hybrid cyber threats detection using explainable AI in Industrial IoT. In Proceedings of the 2023 International Conference on Human-Centered Cognitive Systems (HCCS), Cardiff, UK, 16–17 December 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6.
10. Lin, C.C.; Tsai, C.T.; Liu, Y.L.; Chang, T.T.; Chang, Y.S. Security and privacy in 5G-IIoT smart factories: Novel approaches, trends, and challenges. *Mob. Netw. Appl.* **2023**, *28*, 1043–1058. [CrossRef]
11. Li, S.; Liu, Y. Human-centric Artificial Intelligence enabled Digital Images and Videos Forensic Triage. In Proceedings of the 2023 International Conference on Human-Centered Cognitive Systems (HCCS), Cardiff, UK, 16–17 December 2023; IEEE: Piscataway, NJ, USA; 2023; pp. 1–5.
12. Wang, X.; Li, S.; Iqbal, M. Live Power Generation Predictions via AI-Driven Resilient Systems in Smart Microgrids. *IEEE Trans. Consum. Electron.* **2024**, *70*, 3875–3884. [CrossRef]
13. Wang, X.; Liu, Y.; Li, S. Deep Learning Enabled Keystroke Eavesdropping Attack Over Videoconferencing Platforms. In Proceedings of the IEEE INFOCOM 2023—IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Hoboken, NJ, USA, 20–20 May 2023; IEEE: Piscataway, NJ, USA; 2023; pp. 1–2.
14. Huang, H.; Wang, Y. SecBERT: Privacy-preserving pre-training based neural network inference system. *Neural Netw.* **2024**, *172*, 106135. [CrossRef] [PubMed]
15. Ranade, P.; Piplai, A.; Joshi, A.; Finin, T. CyBERT: Contextualized Embeddings for the Cybersecurity Domain. In Proceedings of the 2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, 15–18 December 2021; IEEE: Piscataway, NJ, USA; 2021; pp. 3334–3342.
16. Feng, X.; Zhu, X.; Han, Q.L.; Zhou, W.; Wen, S.; Xiang, Y. Detecting Vulnerability on IoT Device Firmware: A Survey. *IEEE/CAA J. Autom. Sin.* **2023**, *10*, 25–41. [CrossRef]
17. MITRE. Mapping MITRE ATT&CK to CVEs for Impact. 2021. Available online: https://github.com/center-for-threat-informed-defense/attack_to_cve/ (accessed on 28 September 2025).
18. Tatam, M.; Shanmugam, B.; Azam, S.; Kannoorpatti, K. A review of threat modelling approaches for APT-style attacks. *Heliyon* **2021**, *7*, e05969. [CrossRef] [PubMed]

19. Hemberg, E.; Kelly, J.; Shlapentokh-Rothman, M.; Reinstadler, B.; Xu, K.; Rutar, N.; O'Reilly, U.M. Linking threat tactics, techniques, and patterns with defensive weaknesses, vulnerabilities and affected platform configurations for cyber hunting. *arXiv* **2020**, arXiv:2010.00533.
20. Lakhdhar, Y.; Rekhis, S. Machine Learning Based Approach for the Automated Mapping of Discovered Vulnerabilities to Adversial Tactics. In Proceedings of the 2021 IEEE Security and Privacy Workshops (SPW), San Francisco, CA, USA, 24–27 May 2021; pp. 309–317.
21. Grigorescu, O.; Nica, A.; Dascalu, M.; Rughinis, R. Cve2att&ck: Bert-based mapping of cves to mitre att&ck techniques. *Algorithms* **2022**, *15*, 314.
22. Branescu, I.; Grigorescu, O.; Dascalu, M. Automated Mapping of Common Vulnerabilities and Exposures to MITRE ATT&CK Tactics. *Information* **2024**, *15*, 214. [CrossRef]
23. Zafra, D.K.; Lunden, K.; Alexander, O.; Brubaker, N.; Agboruche, G. In Pursuit of a Gestalt Visualization: Merging MITRE ATT&CK® for Enterprise and ICS to Communicate Adversary Behaviors. 2020. Available online: https://cloud.google.com/blog/topics/threat-intelligence/gestalt-mitre-attack-ics/ (accessed on 28 September 2025).
24. Shuttleworth, R.; Andreas, J.; Torralba, A.; Sharma, P. Lora vs full fine-tuning: An illusion of equivalence. *arXiv* **2024**, arXiv:2410.21228.