Novelty Assessment Report

Paper: Cyber-Zero: Training Cybersecurity Agents without Runtime
PDF URL: https://openreview.net/pdf?id=1gRTeAik4G
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-05

Abstract

Large Language Models (LLMs) have achieved remarkable success in software engineering tasks when trained with executable runtime environments, particularly in resolving GitHub issues. However, such runtime environments are often unavailable in other domains, especially cybersecurity, where challenge configurations and execution contexts are ephemeral or restricted. We present Cyber-Zero, the first runtime-free framework for synthesizing high-quality agent trajectories to train cybersecurity LLMs. Cyber-Zero leverages publicly available CTF writeups and employs persona-driven LLM simulation to reverse-engineer runtime behaviors and generate realistic, long-horizon interaction sequences without actual environments. Using trajectories synthesized by Cyber-Zero, we train LLM-based agents that achieve up to 13.1% absolute performance gains over baseline models on three prominent CTF benchmarks: InterCode-CTF, NYU CTF Bench, and Cybench. Our best model, Cyber-Zero-32B, establishes new state-of-the-art performance among open-weight models, matching the capabilities of proprietary systems like DeepSeek-V3-0324 and Claude-3.5-Sonnet while offering superior cost-effectiveness, and demonstrating that runtime-free trajectory synthesis can effectively democratize the development of state-of-the-art cybersecurity agents.

Core Task Landscape

This paper addresses: **Runtime-Free Trajectory Synthesis for Training Cybersecurity Agents**

A total of **11 papers** were analyzed and organized into a taxonomy with **12 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Offline Reinforcement Learning for Cybersecurity Defense**
- **Offensive Security Agent Training**
- **Reinforcement Learning Methodology**

Complete Taxonomy Tree

- Runtime-Free Trajectory Synthesis for Training Cybersecurity Agents Survey Taxonomy
- Offline Reinforcement Learning for Cybersecurity Defense
  - Network Defense and Access Control
  - Intrusion Detection via Sequence Modeling (1 papers)
    - [11] Real-time Network Intrusion Detection via Importance Sampled Decision Transformers (Hanhan Zhou, 2024) View paper
  - Access Control Policy Learning (1 papers)
    - [4] Dynamic Role-Based Access Control Policy for Smart Grid Applications: An Offline Deep Reinforcement Learning Approach (Georgios Fragkos, 2022) View paper
  - Autonomous Defense Orchestration (1 papers)
    - [9] Offline Reinforcement Learning for Autonomous Cyber Defense Agents (Alexander Wei, 2024) View paper
  - Domain-Specific Defensive Applications
  - Healthcare IoT Security (1 papers)
    - [1] Enhancing Cybersecurity in Healthcare IoT Systems Using Reinforcement Learning (Abid Mohamed Nadhir, 2025) View paper
  - Aviation Cyberphysical Security (1 papers)
    - [6] In Pursuit of Aviation Cybersecurity: Experiences and Lessons From a Competitive Approach (Martin Strohmeier, 2023) View paper
  - Human-Factor-Aware Defense Learning (1 papers)
  - [5] Adversarial Inverse Learning of Defense Policies Conditioned on Human Factor Models (Amirhossein Ravari, 2024) View paper
- Offensive Security Agent Training
  - Runtime-Free Trajectory Synthesis ★ (1 papers)
  - [0] Cyber-Zero: Training Cybersecurity Agents without Runtime (Anon et al., 2026) View paper
  - Runtime-Based Penetration Testing
  - Sequence Modeling for Penetration Testing (1 papers)
    - [2] Pentraformer: Learning agents for automated penetration testing via sequence modeling (Yunfei Wang, 2024) View paper
  - Reasoning-Optimized Penetration Testing (1 papers)
    - [7] Pentest-R1: Towards Autonomous Penetration Testing Reasoning Optimized via Two-Stage Reinforcement Learning (Kong He, 2025) View paper
  - AI System Red-Teaming (1 papers)
  - [3] ASTRA: Autonomous Spatial-Temporal Red-teaming for AI Software Assistants (Xu, 2025) View paper

- Reinforcement Learning Methodology
  - Offline-to-Online RL Stabilization (1 papers)
  - [10] TD3R: Stabilizing the Offline-to-Online Reinforcement Learning by Restricting Policy Updates (Biao Huang, 2025) View paper
  - Imitation Learning for Navigation (1 papers)
  - [8] Learning Your Way Without Map or Compass: Panoramic Target Driven Visual Navigation (Watkins-Valls, 2019) View paper

Narrative

Core task: runtime-free trajectory synthesis for training cybersecurity agents. The field addresses the challenge of training autonomous security agents without requiring expensive real-time interaction with live systems or simulators. The taxonomy reveals three main branches that capture distinct facets of this problem. Offline Reinforcement Learning for Cybersecurity Defense focuses on learning defensive policies from pre-collected datasets, enabling agents to protect systems against intrusions without online exploration—works like Offline Cyber Defense[9] and Healthcare IoT RL[1] exemplify this direction. Offensive Security Agent Training emphasizes the generation and use of synthetic attack trajectories to train penetration-testing agents, as seen in Pentraformer[2], ASTRA[3], and Pentest-R1[7]. The Reinforcement Learning Methodology branch encompasses foundational techniques—such as inverse learning approaches and policy optimization methods like TD3R[10]—that underpin both offensive and defensive settings.

A particularly active line of work centers on synthesizing realistic offensive trajectories without executing attacks in production environments. Cyber-Zero[0] sits squarely within the Offensive Security Agent Training branch and shares this runtime-free philosophy with Pentraformer[2] and ASTRA[3], yet it distinguishes itself by leveraging large-scale trajectory generation to bootstrap agent learning from scratch. In contrast, Pentest-R1[7] and related methods often rely on iterative refinement or hybrid simulation strategies. Meanwhile, the defensive side grapples with distribution shift and the scarcity of labeled attack data, prompting interest in offline methods like Offline Cyber Defense[9] that learn from historical logs. Across both branches, a central tension persists: how to ensure that synthetically trained agents generalize to real-world adversarial dynamics and novel attack vectors, without incurring the cost and risk of live deployment during training.

Related Works in Same Category

No sibling papers were found in the same taxonomy leaf. A taxonomy-subtopic-level comparison will be produced instead.

Taxonomy-Level Summary

Both subtopics involve training or evaluating AI agents in cybersecurity contexts, but target fundamentally different problem spaces. Runtime-Free Trajectory Synthesis focuses on generating training data for cybersecurity agents without requiring live execution environments, while AI System Red-Teaming applies automated agents to probe vulnerabilities in AI systems themselves (like coding assistants). The former addresses data generation for agent training; the latter addresses security evaluation of AI products.

Similarities: - Both involve AI agents operating in cybersecurity-related domains - Both aim to improve security outcomes through automated approaches - Both exclude traditional network penetration testing with runtime interaction

Differences: - Runtime-Free Trajectory Synthesis generates synthetic training trajectories from static data sources, while AI System Red-Teaming actively probes live AI systems for vulnerabilities - Runtime-Free focuses on training cybersecurity agents for general security tasks, while AI System Red-Teaming specifically targets AI system safety and security flaws - Runtime-Free explicitly avoids runtime environments as a design constraint, while AI System Red-Teaming requires interaction with functioning AI systems - The target of security analysis differs: Runtime-Free prepares agents to secure traditional systems, while AI System Red-Teaming secures AI systems themselves

Suggested Search Directions: - Hybrid approaches combining synthetic trajectory generation with AI system vulnerability discovery - Transfer learning from runtime-free synthesized trajectories to AI red-teaming tasks - Static analysis methods for identifying AI system vulnerabilities without runtime interaction

Sibling Subtopics

- **AI System Red-Teaming** (leaves: 1, papers: 1)
- Scope: Automated agent systems for discovering safety vulnerabilities in AI coding assistants and security guidance systems.
- Exclude: Excludes network penetration testing; see Runtime-Based Penetration Testing subcategory.

Contributions Analysis

Overall novelty summary. The paper introduces Cyber-Zero, a runtime-free framework that synthesizes agent trajectories from CTF writeups to train cybersecurity LLMs. According to the taxonomy, this work occupies the 'Runtime-Free Trajectory Synthesis' leaf under 'Offensive Security Agent Training', where it is currently the sole paper. This positioning suggests the paper addresses a relatively sparse research direction within the broader offensive security landscape, which includes more populated areas like runtime-based penetration testing with multiple sibling approaches.

The taxonomy reveals that Cyber-Zero's nearest neighbors are runtime-based methods in sibling leaves: 'Sequence Modeling for Penetration Testing' (Pentraformer) and 'Reasoning-Optimized Penetration Testing' (Pentest-R1). These approaches require executable environments or simulators, whereas Cyber-Zero explicitly avoids runtime interaction. The broader 'Offensive Security Agent Training' branch also includes 'AI System Red-Teaming', which targets AI safety vulnerabilities rather than traditional network penetration. The defensive counterpart branch ('Offline RL for Cybersecurity Defense') addresses policy learning from historical logs but focuses on protection rather than offensive trajectory synthesis.

Among 21 candidates examined, none clearly refute the three core contributions. The CYBER-ZERO framework itself was compared against 1 candidate with no refutation found. The synthesized trajectory dataset and ENIGMA+ agent scaffold each faced 10 candidates, with all classified as non-refutable or unclear. This limited search scope—covering top-K semantic matches and citation expansion—suggests that within the examined literature, no prior work directly overlaps with the combination of runtime-free synthesis, persona-driven simulation, and CTF writeup exploitation for cybersecurity agent training.

Based on the 21-candidate search, the work appears to occupy a novel position at the intersection of trajectory synthesis and offensive security training. However, the analysis does not cover exhaustive domain-specific venues or gray literature in cybersecurity competitions. The taxonomy structure indicates this is an emerging direction with sparse prior work, though the limited search scope means additional related efforts in specialized CTF or security conferences may exist beyond the examined set.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: CYBER-ZERO runtime-free trajectory synthesis framework

Description: The authors present CYBER-ZERO, a novel framework that synthesizes high-quality agent trajectories for training cybersecurity LLMs without requiring access to executable runtime environments. It uses persona-driven LLM simulation with dual models (CTF Player and Bash Terminal) to reverse-engineer behaviors from public CTF writeups and generate realistic multi-turn interaction sequences.

This contribution was assessed against **1 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Generative AI for Simulating Real World Dynamics Applications and Challenges

URL: View paper

Brief Assessment

GenAI Simulating Dynamics[32] focuses on using generative AI for simulating real-world dynamics in domains like autonomous driving and robotics, not on runtime-free trajectory synthesis for cybersecurity agent training. The candidate's approach of translating scenario descriptions into driving trajectories differs fundamentally from CYBER-ZERO's persona-driven LLM simulation for synthesizing cybersecurity agent trajectories from CTF writeups.

Contribution 2: Large-scale synthesized cybersecurity trajectory dataset

Description: The authors build a dataset of 6,188 high-quality CTF writeups spanning 4,610 unique challenges from 543 competitions across six task categories. These synthesized trajectories enable training of LLM agents for vulnerability discovery and exploitation tasks without requiring runtime environments.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Generative AI-Enhanced Cybersecurity Framework for Enterprise Data Privacy Management

URL: View paper

Brief Assessment

GenAI Privacy Framework[22] focuses on enterprise data privacy management using generative AI for synthetic dataset generation in anomaly detection contexts (financial, healthcare, smart city). It does not address CTF challenge trajectories, agent training for vulnerability discovery, or cybersecurity competition data synthesis.

2. A Multimodal Framework for Advanced Cybersecurity Threat Detection Using GAN-Driven Data Synthesis

URL: View paper

Brief Assessment

Multimodal GAN Cybersecurity[31] focuses on synthetic data generation for threat detection using GANs to augment network logs and malware binaries, not on synthesizing agent trajectories for training LLMs in vulnerability discovery tasks. The candidate addresses intrusion detection systems, while the original paper creates interactive CTF-solving trajectories.

3. Enhancing Large Language Models for Secure Code Generation: A Dataset-driven Study on Vulnerability Mitigation

URL: View paper

Brief Assessment

Secure Code LLMs[30] focuses on code generation security and vulnerability mitigation through a dataset of 180 samples targeting 21 vulnerability types. This is fundamentally different from the original paper's 6,188 CTF writeup trajectories for training agents in vulnerability discovery and exploitation tasks.

4. Evaluating Biased Synthetic Data Effects on Large Language Model-Based Software Vulnerability Detection

URL: View paper

Brief Assessment

Biased Synthetic Vulnerabilities[28] focuses on synthetic vulnerability data for static code analysis and classification tasks, not on synthesized agent trajectories for CTF challenges or interactive cybersecurity environments. The candidate addresses dataset bias in vulnerability detection, while the original constructs multi-turn interaction sequences for training LLM agents on offensive security tasks.

5. DeepBalance: Deep-Learning and Fuzzy Oversampling for Vulnerability Detection

URL: View paper

Brief Assessment

DeepBalance[29] focuses on vulnerability detection using deep learning and fuzzy oversampling techniques, not on synthesizing cybersecurity trajectories or training LLM agents for CTF challenges. The datasets and methodologies are fundamentally different.

6. Leveraging gans for synthetic data generation to improve intrusion detection systems

URL: View paper

Brief Assessment

GANs Intrusion Detection[26] focuses on generating synthetic network traffic data for intrusion detection systems, not on synthesizing agent trajectories for training LLMs in vulnerability discovery tasks. The domains and data types are fundamentally different.

7. An Ensemble Transformer Approach with Cross-Attention for Automated Code Security Vulnerability Detection and Documentation

URL: View paper

Brief Assessment

Ensemble Transformer Security[25] focuses on Java vulnerability detection using static code analysis and multi-model transformers, not on synthesizing cybersecurity trajectories for training LLM agents to solve CTF challenges.

8. A novel deep synthesis-based insider intrusion detection (DS-IID) model for malicious insiders and AI-generated threats

URL: View paper

Brief Assessment

Deep Synthesis Insider[27] focuses on synthesizing user profiles for insider threat detection using deep feature synthesis and generative models, not on creating agent trajectories for vulnerability discovery tasks. The datasets serve fundamentally different purposes in distinct cybersecurity domains.

9. AI-enabled Cybersecurity using Synthetic Data

URL: View paper

Brief Assessment

Synthetic Cybersecurity Data[24] focuses on generating synthetic data for training AI models in cybersecurity risk assessment using digital twin technology and adversary simulation. This differs fundamentally from the original paper's contribution of synthesizing agent

trajectories from CTF writeups for training LLM agents to solve vulnerability discovery tasks. The candidate addresses infrastructure security modeling, not agent training data for CTF challenges.

---

10. Approach to Forming Vulnerability Datasets for Fine-Tuning AI Agents

URL: View paper

Brief Assessment

Vulnerability Datasets[23] focuses on constructing datasets for vulnerability detection in source code using static analysis methods, not on synthesizing agent trajectories for CTF challenges or interactive exploitation tasks.

Contribution 3: ENIGMA+ agent scaffold with improved efficiency

Description: The authors develop ENIGMA+, an enhanced version of the ENIGMA scaffold that executes evaluation tasks in parallel rather than sequentially. This improvement dramatically reduces evaluation time from 1-3 days to under 5 hours for 300+ CTF challenges while maintaining evaluation quality.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Resource-Aware Multi-Fidelity Multi-Objective Multidisciplinary Design Optimization

URL: View paper

Brief Assessment

Multi-Fidelity Design[16] addresses multi-objective design optimization with flexible control over simulation model evaluations across different fidelities and disciplines. This is fundamentally different from ENIGMA+'s focus on parallel execution of CTF challenge evaluations in cybersecurity agent assessment.

---

2. Repoforge: Training a sota fast-thinking swe agent with an end-to-end data curation pipeline synergizing sft and rl at scale

URL: View paper

Brief Assessment

Repoforge[15] focuses on distributed evaluation infrastructure using Ray for SWE-Bench tasks, not on CTF challenge evaluation. The parallel execution approach targets different domains and technical requirements.

---

3. A Parallel GEM5-Based Simulation Infrastructure for Multicluster SoC Performance Evaluation

URL: View paper

Brief Assessment

Parallel GEM5[20] focuses on parallelizing hardware simulation infrastructure for SoC performance evaluation, not on agent evaluation scaffolds or cybersecurity tasks. The technical domains are entirely distinct.

---

4. Physics-Aware Compilation for Parallel Quantum Circuit Execution on Neutral Atom Arrays

URL: View paper

Brief Assessment

Quantum Circuit Compilation[19] focuses on physics-aware compilation for neutral atom quantum computers, addressing hardware resource allocation and circuit partitioning for quantum systems. This is fundamentally different from ENIGMA+, which addresses parallel execution of cybersecurity agent evaluation tasks in CTF challenges.

---

5. Parallel WaveNet: Fast High-Fidelity Speech Synthesis

URL: View paper

Brief Assessment

Parallel WaveNet[12] addresses parallel execution in speech synthesis, not agent evaluation scaffolds for cybersecurity tasks. The technical domains are entirely distinct.

---

6. Adaptive Job Scheduling in Quantum Clouds Using Reinforcement Learning

URL: View paper

Brief Assessment

Quantum Job Scheduling[13] focuses on distributed quantum circuit execution across multiple QPUs using reinforcement learning for job scheduling. This addresses quantum hardware resource allocation, not cybersecurity agent evaluation scaffolds or CTF challenge execution parallelization.

---

7. Parallel and High-Fidelity Text-to-Lip Generation

URL: View paper

Brief Assessment

Text-to-Lip[17] focuses on parallel decoding for lip generation from text, not on agent evaluation scaffolds or CTF challenge execution. The domains are entirely different (computer vision vs. cybersecurity agents).

---

8. Parallel Smell Agent Optimization (SAO): Collaborative Subpopulations for Accelerated Convergence

URL: View paper

Brief Assessment

Parallel Smell Agent[18] focuses on bio-inspired optimization algorithms with parallel subpopulations for computational efficiency, not on agent scaffolds for cybersecurity evaluation tasks. The domains and technical approaches are fundamentally different.

---

9. A Survey on Benchmarks of LLM-based GUI Agents

URL: View paper

Brief Assessment

LLM GUI Agents[21] is a survey paper focused on benchmarking GUI agents across diverse environments and metrics. It does not present a novel agent scaffold or execution framework that would refute the novelty of ENIGMA+'s parallel execution architecture for CTF evaluation.

---

10. Flash-Searcher: Fast and Effective Web Agents via DAG-Based Parallel Execution

URL: View paper

Brief Assessment

Flash-Searcher[14] focuses on parallel execution via DAG-based decomposition for web search tasks, not on improving evaluation scaffolds for cybersecurity CTF challenges. The candidate addresses task decomposition and concurrent reasoning paths, while the original contribution specifically targets reducing CTF evaluation time from days to hours through parallelized Docker container execution.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Cyber-Zero: Training Cybersecurity Agents without Runtime View paper
- [1] Enhancing Cybersecurity in Healthcare IoT Systems Using Reinforcement Learning View paper
- [2] Pentraformer: Learning agents for automated penetration testing via sequence modeling View paper
- [3] ASTRA: Autonomous Spatial-Temporal Red-teaming for AI Software Assistants View paper
- [4] Dynamic Role-Based Access Control Policy for Smart Grid Applications: An Offline Deep Reinforcement Learning Approach View paper
- [5] Adversarial Inverse Learning of Defense Policies Conditioned on Human Factor Models View paper
- [6] In Pursuit of Aviation Cybersecurity: Experiences and Lessons From a Competitive Approach View paper
- [7] Pentest-R1: Towards Autonomous Penetration Testing Reasoning Optimized via Two-Stage Reinforcement Learning View paper
- [8] Learning Your Way Without Map or Compass: Panoramic Target Driven Visual Navigation View paper
- [9] Offline Reinforcement Learning for Autonomous Cyber Defense Agents View paper
- [10] TD3R: Stabilizing the Offline-to-Online Reinforcement Learning by Restricting Policy Updates View paper
- [11] Real-time Network Intrusion Detection via Importance Sampled Decision Transformers View paper
- [12] Parallel WaveNet: Fast High-Fidelity Speech Synthesis View paper
- [13] Adaptive Job Scheduling in Quantum Clouds Using Reinforcement Learning View paper
- [14] Flash-Searcher: Fast and Effective Web Agents via DAG-Based Parallel Execution View paper
- [15] Repoforge: Training a sota fast-thinking swe agent with an end-to-end data curation pipeline synergizing sft and rl at scale View paper
- [16] Resource-Aware Multi-Fidelity Multi-Objective Multidisciplinary Design Optimization View paper
- [17] Parallel and High-Fidelity Text-to-Lip Generation View paper
- [18] Parallel Smell Agent Optimization (SAO): Collaborative Subpopulations for Accelerated Convergence View paper
- [19] Physics-Aware Compilation for Parallel Quantum Circuit Execution on Neutral Atom Arrays View paper
- [20] A Parallel GEM5-Based Simulation Infrastructure for Multicluster SoC Performance Evaluation View paper
- [21] A Survey on Benchmarks of LLM-based GUI Agents View paper
- [22] Generative AI-Enhanced Cybersecurity Framework for Enterprise Data Privacy Management View paper
- [23] Approach to Forming Vulnerability Datasets for Fine-Tuning AI Agents View paper
- [24] AI-enabled Cybersecurity using Synthetic Data View paper
- [25] An Ensemble Transformer Approach with Cross-Attention for Automated Code Security Vulnerability Detection and Documentation View paper
- [26] Leveraging gans for synthetic data generation to improve intrusion detection systems View paper
- [27] A novel deep synthesis-based insider intrusion detection (DS-IID) model for malicious insiders and AI-generated threats View paper
- [28] Evaluating Biased Synthetic Data Effects on Large Language Model-Based Software Vulnerability Detection View paper
- [29] DeepBalance: Deep-Learning and Fuzzy Oversampling for Vulnerability Detection View paper
- [30] Enhancing Large Language Models for Secure Code Generation: A Dataset-driven Study on Vulnerability Mitigation View paper
- [31] A Multimodal Framework for Advanced Cybersecurity Threat Detection Using GAN-Driven Data Synthesis View paper
- [32] Generative AI for Simulating Real World Dynamics Applications and Challenges View paper