
SHTR
AI vs AI in Cybersecurity
The Battle of Intelligent Offense and Defense

As artificial intelligence becomes more pervasive in cybersecurity, we are entering an era where **AI is both the attacker and the defender**.

This dynamic “**AI vs AI”** is shaping the future of *cyber warfare, threat mitigation, and digital resilience*.

AI Being Used for Designing and Executing Sophisticated Attacks

Enhanced Phishing and Social Engineering Attacks

- **Hyper-Personalization:** AI analyzes vast amounts of data from social media and other sources to craft highly convincing phishing messages tailored to individuals or organizations.
- **Deepfake Technology:** AI is used to generate realistic audio and video deepfakes, making impersonation attacks more convincing, such as CEO fraud or fake video calls for financial scams.
- **Chatbot Impersonation:** AI-powered chatbots mimic real customer service representatives, tricking victims into revealing sensitive information.

Malware Development

- **Evasive Malware:** AI enables malware to adapt and evade traditional detection methods by altering its code or behavior dynamically.
- **Polymorphic Attacks:** AI can create malware that changes its signature with every instance, making it difficult for signature-based antivirus software to detect.
- **Ransomware Optimization:** AI is used to identify high-value targets and optimize ransomware payloads for maximum impact.

Sophisticated Attacks..continued

Automated Attack Execution

- **AI-Powered Reconnaissance:** AI tools scrape public and private data sources to gather detailed information about targets, automating the reconnaissance phase of an attack.
- **Brute Force Attacks:** AI accelerates password-guessing attempts by analyzing patterns in commonly used credentials and adapting its methods in real time.
- **AI Bots for Distributed Denial of Service (DDoS):** AI coordinates botnets to launch more sophisticated and targeted DDoS attacks, capable of bypassing traditional defenses.

AI-Driven Cyber Espionage

- **Target Identification:** AI tools identify high-value individuals or organizations for espionage based on their digital footprints and communication patterns.
- **Automated Surveillance:** AI analyzes intercepted communications or data to identify key insights or plan attacks.

Adversarial Attacks Against *AI Systems*

Poisoning Training Data: Attackers manipulate the training datasets of AI systems to introduce biases or vulnerabilities, compromising their effectiveness.

Model Inversion: AI is used to reverse-engineer machine learning models, revealing sensitive data used during training.

Evasion Techniques: Attackers use adversarial inputs to confuse AI systems, such as slight alterations in images or data that deceive facial recognition or intrusion detection systems.

Challenges in Combating AI-Driven Attacks

Speed and Scale - AI allows attackers to execute attacks faster and at a larger scale than ever before.

Low Barriers to Entry- AI tools and frameworks are increasingly accessible, enabling even less-skilled attackers to launch sophisticated campaigns.

Detection Difficulty: AI-driven attacks often mimic legitimate behaviors, making them harder to detect and mitigate.

The Defensive Use of AI

Detect Anomalies:

Behavioral analytics powered by machine learning flag deviations in user or system behavior in real-time.

Predict Attacks:

AI models can identify patterns that precede attacks, enabling proactive defenses.

Automate Incident Response:

AI-driven SOAR (Security Orchestration, Automation, and Response) platforms can respond to threats within seconds.

Threat Hunting:

AI analyzes logs and telemetry data to uncover hidden threats or lateral movement across networks.

Phishing Protection:

AI filters incoming communications using NLP to detect suspicious language and impersonation attempts.

AI Use-cases

AI-Powered Anomaly Detection: Monitor the network and detect deviations from normal behavior in real-time.

Cross-Sector Collaboration: Share threat intelligence, especially around AI-driven tactics, techniques, and procedures (TTPs).

Red Teaming AI Models: Test AI defenses with offensive AI techniques to find and fix weaknesses.

Zero Trust Becomes Essential: With AI able to imitate humans and behaviors, assuming nothing and verifying everything is more critical than ever.

AI Use-cases.. continued

Threat Hunting with AI: Proactively identifying and hunting for emerging threats by analyzing large volumes of data from threat intelligence feeds,

Attack Surface Monitoring: Automate the identification of all connected devices and systems, continuously assessing the attack surface and prioritize security efforts based on vulnerability databases, and asset profiles.

Automated Response: Real-time incident detection and automated responses to mitigate the impact of cyber-attacks.

Risk Assessment Models: Continuously evaluate the security posture environments by assessing all components for vulnerabilities, misconfigurations, and outdated software.

AI Use-cases.. continued

Automated Vulnerability Scanning: Scan environments for vulnerabilities more efficiently and prioritize vulnerabilities based on the likelihood of exploitation, the severity of potential damage, and the criticality of affected devices.

Smart Patching and Updates: Automate the patching process.

AI for Remote Access Security: Secure remote access by using behavioral biometrics, anomaly detection, and multi-factor authentication (MFA) to ensure that only authorized personnel can access critical systems.

AI-Driven Policy Enforcement: AI can help enforce security policies by automatically monitoring compliance and flagging violations.

The Next steps….

Prioritize AI Security as a Strategy: As AI becomes integral to operations, its security must be treated with the same rigor as traditional IT and OT systems, ensuring robust governance, regulatory compliance, and ethical AI practices across all functions.

Empower Teams Through Education / Awareness: Build a culture of AI literacy, Train security, engineering, and operations teams and at the same time, educate executives and non-technical stakeholders on the implications of AI security.

Stay Adaptive and Future-Ready Against Emerging Threats: Proactive engagement through red teaming, scenario planning, and external collaboration will ensure resilience and long-term advantage.

ST

The dual-use nature of AI makes it essential for defenders to stay ahead of attackers by constantly innovating and improving their defensive capabilities.

THANK YOU

Sapann Harish Talwar