FORTINET

WHITE PAPER

Harnessing Artificial Intelligence in Telco Cybersecurity Ecosystems

Today's organizations and their cybersecurity teams face many challenges. As data volumes explode, increasing infrastructure complexity is expanding the potential attack surface, attracting ever more sophisticated hackers, and resulting in heightened cyber-regulatory pressure. In Telco environments - due to their highly complex technology, services, and partner environment – these challenges are multiplied.

Although AI is not new, the emergence of Generative AI (GenAI) together with its growing maturity and widening implementations, are establishing AI as a disruptive technology with tremendous potential for good (and bad), bringing it to the forefront of technology and business discussions.

For cybersecurity practitioners, AI promises the fruition of all their dreams: self-remediating, self-healing networks and ecosystems, instantaneous detection, response, and investigation of attacks and breaches, and the automation of everything that can be automated.

> "The telecommunications industry stands at the crossroads of AI, connectivity, security and sustainability in a digitized world. It drives the AI revolution and shoulders the responsibility of securing critical infrastructure and safeguarding our world." *(World economic forum, 2024)*

But we are not there yet. Although AI has played an important, albeit sometimes hidden role in cybersecurity for many years, its use and effectiveness across different aspects of cybersecurity is segmented and limited.

This paper provides a brief and realistic view of the key aspects of AI technology and examines how and where it can be used to enhance and augment cybersecurity ecosystems within Telco environments.

A brief history of AI

Although the evolution of AI spans decades, we can mark the birth of modern AI in the 1990s with the rise of Machine Learning (ML), where systems began to learn from data rather than by following explicitly programmed rules.

With the rise of deep neural networks and deep learning - fueled by advances in hardware (GPUs), big data, and the explosion of data from the internet, social media, and sensors - AI algorithms became more powerful. This gave rise to Generative AI (GenAI), which encodes a simplified representation of the training data, creating new representations based on the original data.

ML and GenAI are at the core of the AI revolution, but to better understand how and where they may fit into a Telco cybersecurity ecosystem, one must first understand their different objectives, characteristics, and limitations.

Although different subbranches of AI, ML and GenAI are heavily intertwined with ML fulfilling a critical role in the training of GenAI.

ML and GenAI

Machine Learning (ML)

ML systems learn and improve based on the data they consume. ML can assist with the classification of data, find natural groupings of observations, and help predict patterns and behaviors. In cybersecurity, this helps to identify viruses and malware, as well as direct attacks, and other abnormal behavior.

There are four major ways in which ML algorithms can be trained:

- **Supervised learning** presents the algorithm with labeled data and the desired outputs to help it discover a general rule that relates them.
- **Unsupervised learning** presents the algorithm with unlabeled data, leaving it to discover relationships and patterns on its own.
- **Semi-supervised learning** presents the algorithm with both labeled and unlabeled data.
- **Reinforcement learning** takes a more iterative approach where the system learns through trial and error and feedback from data analysis.

In cybersecurity, the specific ML algorithm training methods are based on the task in hand (prediction, classification, clustering, etc.), the environment in which the cybersecurity solution operates (static/dynamic/homogenous/etc.), and the cybersecurity technology or solution where the model is used (firewall, Web and API FW/WAAF, Security Incident and Event Management/SIEM, etc.).

The result is a ML model that acts as a program and performs the task for which it was trained on real data.

Once training is done and the ML model is generated, its size will vary according to the dataset and algorithm used. A specialized model dealing with specific datasets and tasks, such as Anti-Virus categorization, can limit the model footprint to several MBs, and reduce its consumption of compute and storage resources. These two characteristics make specialized ML models highly suitable technology to integrate within a network security function.

Another important factor to consider is the deterministic nature of the environment in which the AI will operate:

- **A deterministic** environment is where no randomness is involved. Given the same initial conditions and actions, the output will always be the same. Such environments are predictable and yield high accuracy. For such environments, ML models which are based on supervised and semi-supervised learning will be most efficient and effective.

- **A non-deterministic (stochastic)** environment introduces an element of unpredictability where the future state of the environment cannot be predicted with certainty. This can arise from random events or incomplete observability of the environment. In such cases, the ML model used must account for this uncertainty when making decisions or giving recommendations. For such environments, a mix of ML models based on unsupervised and semi-supervised learning provide the best results.

Given the same conditions, data, and parameters, deterministic ML will always produce the same result, whereas non-deterministic ML will not. The choice between the two therefore depends on the environment in which the model is intended to operate. Deterministic ML models are suitable for relatively static or deterministic environments, while non-deterministic ML models are most suitable for stochastic environments. It is important to remember that different ML models can be used serially or in parallel to achieve the required outcome.

Generative AI (GenAI)

ML uses predictive models to classify data, recognize patterns, and predict outcomes within a specific context or domain, while GenAI uses neural networks and deep learning to identify patterns as a basis for generating new content and augmenting data sets.

Modern GenAI systems are general-purpose or task-specific Large Language Models (LLMs), designed to interact with humans and generate intelligent and creative content (text, images, videos, etc.) when queried. To be effective, LLMs need to be trained on enormous datasets sourced mostly from the public domain (books, magazines, archives, articles, websites, etc.), and require significant compute and storage resources.

To realize an LLM's potential in a Telco - it needs to be augmented with the specific Telco context and data encompassing its complex environment (services, assets, networks, partners, suppliers, customers, etc.). This presents a challenge as a significant proportion of Telco data is private, protected, and sits behind network firewalls. Augmenting a non-sovereign LLM (an LLM that is not owned by the Telco) with such data, while maintaining privacy, confidentiality, and regulatory compliance, can therefore hinder the use of 3rd party GenAI LLMs.

Developing and training a completely sovereign LLM consumes significant time, resources, and effort, which has given rise to a growing market for pretrained, task-specific LLMs that can be augmented with Telco data and context to create sovereign LLMs in different domains, such as operations, customer engagement, etc.

Solutions are also available to enable Telcos to securely use non-sovereign LLMs with augmented private data. Examples of such solutions include private/sensitive data obfuscation and the use of cloud data platforms as secure intermediary access to public LLMs.

GenAI, particularly models like large language models, and other generative systems, are usually non-deterministic by nature. Generative models are intentionally designed to allow for diversity and creativity in their outputs. GenAI models can be made deterministic by controlling randomness factors, such as fixing random seeds and using deterministic sampling methods, but this limits their ability to generate diverse or creative outputs which is not always desirable.

In general, the non-deterministic nature of GenAI, and the compute resources it requires, limits its suitability for direct implementation in cybersecurity network functions and specifically for in-line security functions.

AI in The Telco Cybersecurity Ecosystem

Any Telco cybersecurity ecosystem has the following three key elements:

Threat Intelligence is the process of collecting and analyzing data to identify threat actors' modes of operation, motives, victims, and the tools, malware and techniques used to attack. It is a non-real-time process that takes place in the research labs of cybersecurity and 3rd party organizations and is "fed" to Telcos' cybersecurity infrastructures and security teams. Threat intelligence is essential for effective and continuous protection against current and future threats.

Collection and analysis of large volumes of networking and cybersecurity datasets are the core of the threat intelligence process and AI tools play a key role in dealing with the huge quantities of data and can alleviate some tasks from the cybersecurity analysts:

ML plays an important role in analyzing, categorizing, and identifying trends, attack chains, methods and tools, anomalies, suspicious behaviors, bad traffic and attacks. GenAI can be used to generate outbreak alerts, advisories, and related documentation. When applicable, GenAI can even write cybersecurity signatures to protect against identified threats.

The critical role of Telcos as the backbone of digital transformation and innovation will continue to drive network traffic and services alongside attacks against Telcos and their customers. Harnessing AI in threat intelligence is therefore essential to effectively securing Telcos and their customers.

Fortinet example

With over 100 billion security events per day, Fortinet's **FortiGuard Labs** has been developing and using AI for over 10 years. Events and their related datasets received from Fortinet equipment installed around the world, from 3rd party organizations, and from the dark web are analyzed using AI with neural networks and deep learning.
This AI-driven threat intelligence is made available to Telcos and enterprises around the world – ensuring their Fortinet cybersecurity infrastructure and security teams are up to date and effective against the latest threats.

Cybersecurity Network Functions are deployed in Telco networks to inspect and analyze network traffic and take appropriate action in real-time. As these devices operate directly in the path of Telco network traffic, they must operate at Telco-grade performance and scalability levels, supporting millions of concurrent sessions, throughputs up to several terabits per second, and a very low latency measured in microseconds.

Cybersecurity network functions (physical or virtual) must be deterministic in their ability to analyze large datasets and take decisions and actions in real-time, narrowing the scope of AI utilization to task specific ML models, where footprint and speed can be well defined and controlled.

Fortinet example

Fortinet **FortiWeb** Web Application and API Protection (WAAP) provides inline security for applications and APIs, and these are mostly non-deterministic environments which can be prone to false positives. To deal with such environments while minimizing false positives, FortiWeb uses two serially connected internal ML models.
The first learns the application and its API interactions and behavior to determine whether an HTTP request or API call varies significantly from previously observed behavior.
If an anomaly is detected, a second ML engine determines if this is a threat or a benign variance using highly trained threat models. These ML engines also facilitate BOT detection and help to identify any legitimate change in application and API behaviors.

Security Operations (SecOps) and management combine a set of tools and capabilities to provide a centralized investigation and remediation that is orchestrated, automated, and augmented to reduce cyber risk, cost, and operational effort. Key SecOps functions include correlation, enrichment, analysis, triage, validation, and response, based on information

from a wide range of sensors, including network functions such as routers and switches, as well as cybersecurity functions such as firewalls, NDR and EDR systems.

The different sensors send alerts and information to the SecOps platform or tools where they are processed, correlated and analyzed to provide Telco security teams with the ability to detect, mitigate, and respond to cyber threats. Key SecOps tools include SIEM for real-time threat detection through log aggregation, correlation, and analysis, SOAR for incident response automation, and NDR for network traffic monitoring for threat hunting.

Due to the complexity and scale of Telco networks and services, the importance of SecOps is magnified. SecOps does not happen inline, it is not real time, and it has access to large amount of compute resources. The use of both ML and GenAI tools are th erefore widely used:

- ML models are integrated within Sec Ops tools and are used for threat detection by correlating and analyzing a wide range of behavioral indicators. For example, to detect anomalous or malicious actions that security teams may easily overlook, AI and advanced analytics are used to monitor and corollate activity across users, devices, networks, emails, applications, files, and logs.
- GenAI tools are used to assist the SecOps team with incident investigations, response recommendations, creation of playbooks to automate incident response, and to generate incident reports, etc.

Fortinet example

ML models are widely integrated into key SecOps tools such as FortiSIEM, FortiSOAR, and FortiNDER, to perform tasks such as:

- User and Entity Behavior Analysis (UEBA) - establish baseline behavior, detect anomalies, and analyze trends – based on log data collection, correlation, and analysis.
- Suggest actions based on historic data analysis.
- Prioritization and recommendation.
- Network behavior baseline and anomaly detection.

FortiAI is a GenAI assistance built into analyst workflows to inform and expedite incident management and threat hunting. It augments and refines LLM results with the latest Fortinet threat intel, product knowledge, and use-cases, delivering accurate and actionable results.

Conclusions

AI is not the silver bullet that will resolve all Telco cybersecurity challenges, but it is increasingly becoming an indispensable tool due to its ability to rapidly process vast amounts of data, identify complex attack patterns, and automate responses.

The possible uses of AI are determined by the various components and roles of Telco cybersecurity and by the differences in nature, capabilities, and requirements of ML and GenAI.

With its great potential and massive investments, AI is rapidly evolving new capabilities and its use in Telco cybersecurity ecosystems will increase, enabling Telcos to further improve their ability to detect and prevent threats, respond to incidents faster and more efficiently, and to stay one step ahead of cybercriminals.

February 4, 2025 11:21 AM / EMKTG-368-0-1-EN