ITUPublications

International Telecommunication Union
Standardization Sector

ITU-T Technical Report

(12/2025)

XSTR.saAIoT

Security threat analysis for artificial intelligence of things on devices

Technical Report ITU-T XSTR.saAIoT

Security threat analysis for artificial intelligence of things on devices

Summary

Artificial intelligence of things (AIoT) integrates artificial intelligence (AI) technology with the Internet of things (IoT)and significantly enhances the intelligence and automation levels of IoT applications. By constructing AIoT systems and utilizing AI's advanced data processing and learning capabilities, real-time analysis and automated decision-making can be performed on the vast amounts of data collected by IoT devices. This not only improves efficiency and reduces costs but also demonstrates broad application prospects in fields such as smart homes, smart cities and industrial Internet.

Although the deep integration of AI and IoT in AIoT can achieve more efficient IoT operations and enhance data management and analysis, it also introduces a series of unpredictable security threats and complex challenges. This document identifies and analyses a range of potential security threats and challenges of AIoT on devices.

Keywords

Artificial intelligence, artificial intelligence of things, Internet of things.

Note

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

Table of Contents

Technical Report ITU-T XSTR.saAIoT

Security threat analysis for artificial intelligence of things on devices

1 Scope

This Technical Report specifies security threats for artificial intelligence of things (AIoT) on devices; in particular, it provides:

– a security threat analysis framework for analysing the specific security threats and challenges of artificial intelligence of things on devices;

– a landscape analysis for artificial intelligence of things on devices.

This Technical Report does not cover security threat analysis of general AI systems or IoT.

2 References

[ISO/IEC 27400] ISO/IEC 27400:2022, *Cybersecurity – IoT security and privacy – Guidelines*.

3 Definitions

3.1 Terms defined elsewhere

This Technical Report uses the following terms defined elsewhere:

3.1.1 AI system [b-ISO/IEC 22989]: Engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives.

NOTE 1 – The engineered system can use various techniques and approaches related to artificial intelligence (3.1.3) to develop a model (3.1.23) to represent data, knowledge (3.1.21), processes, etc. which can be used to conduct tasks (3.1.35).

NOTE 2 – AI systems are designed to operate with varying levels of automation.

3.1.2 artificial intelligence [b-ISO/IEC 22989]: Set of methods or automated entities that together build, optimize and apply a model (see clause 3.1.15) so that the system can, for a given set of predefined tasks, compute predictions (see clause 3.1.18), recommendations, or decisions.

3.1.3 artificial intelligence of things (AIoT) [b-ITU-T YSTP.AIoT]: Internet of things powered by artificial intelligence to achieve intelligent IoT applications and things.

NOTE – AI technology can be applied within the end to end of IoT infrastructure, and can especially be implemented in the device and the edge, to enhance the intelligence of IoT applications and things.

3.1.4 inference [b-ISO/IEC 22989]: Reasoning by which conclusions are derived from known premises.

3.1.5 Internet of things [b-ITU-T Y.4000]: A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.

NOTE 1 – Through the exploitation of identification, data capture, processing and communication capabilities, the IoT makes full use of things to offer services to all kinds of applications, whilst ensuring that security and privacy requirements are fulfilled.

NOTE 2 – From a broader perspective, the IoT can be perceived as a vision with technological and societal implications.

3.1.6 IoT device [b-ISO/IEC 20924]: Endpoint (3.1.15) that interacts with the physical world through sensing or actuating.

NOTE – An IoT device can be a sensor or an actuator.

3.1.7 IoT system [b-ISO/IEC 20924]: System providing functionalities of IoT.

NOTE – An IoT system can include, but not be limited to, IoT devices, IoT gateways, sensors, and actuators.

3.1.8 machine learning [b-ISO/IEC 22989]: Process of optimizing model parameters through computational techniques, such that the model's behaviour reflects the data or experience.

3.1.9 stakeholder [b-ISO/IEC 22989]: Any individual, group, or organization that can affect, be affected by or perceive itself to be affected by a decision or activity.

3.1.10 system [b-ISO/IEC TR 24028]: Combination of interacting elements organized to achieve one or more stated purposes.

3.1.11 thing [b-ITU-T Y.4000]: With regard to the Internet of things, this is an object of the physical world (physical things) or the information world (virtual things), which is capable of being identified and integrated into the communication networks.

3.1.12 threat [b-ISO/IEC 27000]: Potential cause of an unwanted incident, which can result in harm to a system or organization.

3.2 Terms defined in this Technical Report

None.

4 Abbreviations and acronyms

This Technical Report uses the following abbreviations and acronyms:

AI Artificial Intelligence

AIoT Artificial Intelligence of Things

CPU Central Processing Unit

DDoS Distributed Denial of Service

GPU Graphics Processing Unit

IoT Internet of Things

MITM Man-In-The-Middle

ML Machine Learning

TPU Tensor Processing Unit

5 Conventions

None.

6 Overview

With the rapid advancement of Internet of things (IoT) and artificial intelligence (AI) technologies, artificial intelligence of things (AIoT) is fostering the deep integration of AI and IoT devices, revolutionizing how people work, live and interact with the world. Smart devices such as wearables, drones and smart speakers are extensively utilized in both everyday life and industrial settings. These devices, equipped with advanced sensing, computing, networking and communication capabilities, can gather, analyse and transmit a wide array of data types, including images, video, audio, text and wireless signals. Progress in AI technologies, especially in areas like machine learning, deep learning and generative AI, has propelled the convergence of AI and IoT. The synergy between IoT and AI enhances decision-making capabilities, enabling proactive interventions, automated decisions and personalized services. It also facilitates more efficient collaboration between IoT-connected devices. As a result, AIoT significantly improves operational efficiency.

To comprehensively understand the specific security threats associated with devices in AIoT scenarios, this technical report organizes existing studies. Figure 1 illustrates the fundamental concepts of AIoT, including device-edge-cloud collaborative architecture, as described in [b-ITU-T YSTP.AIoT]. Specifically, in the context of AIoT, the capabilities of devices can be logically categorized into four main areas: AI-related capabilities, data-related capabilities, IoT-related capabilities and the collaborative capabilities between cloud, edge and device.

AIoT faces a range of unpredictable threats and complex challenges. These threats can be categorized into three main types: general IoT threats, general AI system threats and specific threats on AIoT. Similar to general IoT threats, AIoT faces vulnerabilities on devices, network attacks and supply chain threats. For example, IoT devices often collect large amounts of personal and sensitive data, such as location, health status and lifestyle habits. If this data is not encrypted or protected during transmission, it can be stolen by malicious attackers. Like general AI systems, AIoT systems are also susceptible to security threats, including data poisoning, adversarial attacks, model theft and backdoor attacks, which can disrupt normal AI model inference, leading to incorrect predictions and reduced generalization performance.

In addition to the above-mentioned general security threats of IoT and AI systems, AIoT, due to the integration of AI technology and IoT technology, not only has a wide-ranging impact on general IoT and AI threats but also brings many new specific security threats for the AIoT scenario. This document proposes a security threats analysis framework, which includes five layers, from bottom to top: hardware layer, system layer, data layer, network layer and application layer.

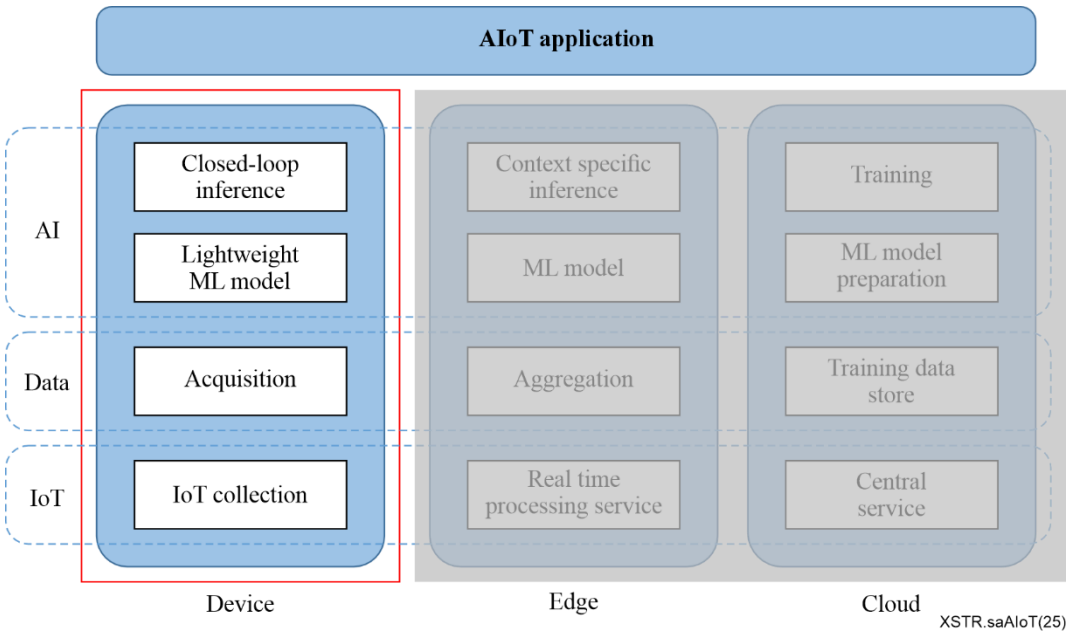

Figure 1 – The concept of AIoT on devices

6.1 The similarities and differences between AIoT and agentic AI

In terms of technical core, service objectives and functional focus, AIoT and agentic AI are significantly different.

AIoT is an integrated system of 'artificial intelligence + Internet of things'. Its core is to connect hardware devices such as sensors and controllers in the physical world based on the 'device-edge-cloud' architecture, realizing the collection, transmission and intelligent linkage of physical data. Essentially, it is an infrastructure serving the 'digital management and control of the physical world', functionally focusing on device collaboration and efficiency optimization of physical systems (e.g., fault early warning of industrial equipment, automatic adjustment of smart home devices) and relying on hardware devices and data transmission to form a closed loop.

Agentic AI is an agent system with autonomous decision-making capabilities. Its core is to simulate human logic to understand objectives, decompose complex tasks and dynamically adapt to environmental changes. Essentially, it is an intelligent executor 'replacing humans in decision-making and execution', functionally focusing on the autonomous advancement of complex tasks (e.g., handling multi-step customer service needs, responding to road conditions in autonomous driving). It does not rely on IoT hardware, achieves objectives only through decision-making algorithms and interaction logic, and does not take the connection of physical devices as a prerequisite.

Both are driven by artificial intelligence technologies (such as machine learning and deep learning) as their core. Their ultimate goals are to improve service efficiency, reduce the burden of human operations and to be applied in multiple fields such as smart homes, industry, medical care and transportation. Moreover, there is potential for synergy in practical scenarios, AIoT can provide real-time data support from the physical world for agentic AI (such as environmental and equipment data collected by sensors), while agentic AI can make autonomous decisions based on this data and drive AIoT devices to perform specific actions, jointly building an intelligent service closed loop of 'data collection-decision-making-action execution'.

6.2 Device, edge and cloud architecture and functions in AIoT systems

In an AIoT system, device, edge and cloud form the three core layers of the architecture. Due to differences in physical location, hardware resources, computing power and data processing methods, they assume distinct roles and functions. The device is the perception and execution peripheral that directly faces the physical world; the edge is the regional computing hub for nearby processing; and

the cloud is the core brain for overall coordination. Each has its own focus in terms of data processing scope, computing intensity and service objectives, jointly supporting the efficient operation of the AIoT system.

The device is the frontmost node connecting AIoT to the physical world, with its core positioning as a data collection source and instruction executor. Its hardware mostly consists of low-power, miniaturized devices, such as sensors (temperature and humidity, vibration, image sensors) and smart terminals (smart door locks, wearable bracelets). Restricted by size and energy consumption, their computing power and storage capacity are extremely limited. In terms of data processing, the device only performs minimal local responses. For example, a sensor directly triggers an alarm when detecting excessive temperatures or a smart door lock executes the unlocking action after locally verifying a fingerprint. It only transmits key results (such as abnormal signals) to the edge or cloud and does not process complex data. Its core value lies in realizing real-time capture and immediate execution of physical signals, ensuring low latency and high privacy.

The edge is the regional-level processing hub between the device and the cloud, with its core positioning as data offloading and proximity-based decision-making. Its hardware usually includes edge gateways, industrial servers or small local computing nodes, which have moderate computing power and storage capacity and can connect to multiple devices (such as dozens of sensors in a workshop or multiple cameras in a community). In data processing, the edge filters, aggregates and preprocesses the massive raw data from the device, for instance, screening out abnormal images captured by cameras, summarizing environmental data from multiple sensors to analyse regional trends and then uploading the streamlined valid data to the cloud. It can also make quick decisions based on local data (such as adjusting workshop equipment parameters). Its core value is balancing real-time performance and computing costs, reducing bandwidth pressure and latency on the cloud.

The cloud is the global intelligent brain of AIoT, with its core positioning as large-scale data storage and in-depth decision-making. It relies on cluster servers (CPU/GPU arrays) in remote data centres and possesses large-scale computing power and storage capacity. In data processing, the cloud receives aggregated data from multiple edges, conducts global analysis, complex model training and long-term trend prediction. For example, it optimizes the overall traffic scheduling plan based on traffic data from various urban regions and delivers the optimized models or rules to the edge and device. Its core value is breaking through local resource limitations, realizing intelligent coordination across regions and scenarios, and supporting the scaling and in-depth enhancement of intelligence of AIoT systems.

7 Stakeholders of AIoT systems

7.1 General

AIoT systems involve multiple stakeholders, each playing a specific role within this ecosystem. By clearly defining the security threat responsibilities of each stakeholder in the AIoT system and considering the inherent security threat levels associated with their roles, effective security threat guidelines can be established.

[ISO/IEC 27400] provides three types of roles in IoT systems: IoT service provider, IoT service developer and IoT user. Unlike in IoT systems, these roles need to be redefined and their responsibilities clarified in the context of AIoT systems.

The key stakeholders in AIoT systems are introduced as follows.

7.2 AIoT service provider

AIoT service providers manage and operate the services offered to AIoT users.

These services include general offerings from IoT service providers, such as connectivity services, data collection and management services and management services for IoT-related assets like IoT

devices. Additionally, AIoT service providers must offer advanced capabilities, including the deployment and maintenance of AI models/algorithms capable of performing complex reasoning tasks, as well as efficient data analytics services.

AIoT service providers need to tailor their offerings to meet the specific needs of AIoT users based on the current performance of their devices.

AIoT service providers not only need to meet the functional, non-functional and security requirements expected by general IoT users but also need to implement stringent controls to mitigate AI-specific security threats within the AIoT system, ensuring secure and stable provision of intelligent decision-making services.

7.3 AIoT service developer

General IoT service developers typically take on roles such as architects for IoT solutions or platforms, designers or implementers of IoT applications and designers or implementers of IoT devices. They should adhere to 'security and privacy by design' principles or use secure software development life cycles.

AIoT service developers, in addition to fulfilling the aforementioned roles in general IoT, must also conduct rigorous security threats analysis and control for the AI technologies used to empower IoT devices.

A sub-role of AIoT service developers is AIoT device developers, who are responsible for designing and producing hardware devices capable of loading AI models/algorithms. These devices can be operated and used by AIoT users or AIoT service providers.

For AIoT device developers, during the design phase, it is necessary to consider the performance and security requirements of AI technologies used in AIoT devices, including hardware, interfaces, memory and computing units. They must provide control measures to address AI-specific security threats in AIoT scenarios.

7.4 AIoT user

Similar to general IoT systems, users are the end users of AIoT services and can be categorized into human users and digital users. Human users are individuals who use AIoT services. Digital users are non-human entities that use AIoT services, such as automated services acting on behalf of human users.

AIoT users need to specify the functional and non-functional requirements for AIoT systems or services, as well as the functional requirements for AI technologies. These needs are to be met by AIoT service providers and AIoT developers, who are also responsible for ensuring the security of AIoT users.

8 Landscape for security threats of artificial intelligence of things on devices

This document depicts a layered view of specific threats of AIoT on devices in Figure 2, which includes five layers, from bottom to top: hardware layer, system layer, data layer, network layer and application layer. This five-layer model refines the security threats associated with the core capabilities in Figure 1 through layered detailing: 1) The security threats for AI-related capabilities in Figure 1 correspond to the application layer and system layer in Figure 2; 2) The security threats for data-related capabilities in Figure 1 correspond to the data layer in Figure 2; 3) The security threats for IoT-related capabilities in Figure 1 correspond to the hardware layer in Figure 2; 4) The collaborative capabilities between cloud, edge and devices in Figure 1 correspond to the network layer in Figure 2.

– The hardware layer includes sensors, actuators, embedded chips and other terminal components that interact directly with the physical environment. Unlike the data layers, it

cannot process data, it only captures and transmits raw physical signals. It is realistic to analyse security threats for the hardware layer as follows, but not limited to, hardware security vulnerabilities and physical manipulation. For example, in the smart grid scenario, attackers can interfere with the current sensor readings by using a strong magnetic field, inducing the edge model to misjudge the load status and trigger wrong power-off instructions.

– The system layer includes the operating system, middleware and artificial intelligence reasoning framework, and is responsible for resource management and task scheduling within the device. Unlike the network layer, it operates in complete isolation from external nodes. It is realistic to analyse security threats for the system layer as follows, but not limited to, resource overload, multi-tenant inference and cross-processor inference. In smart cities' urban traffic management, edge nodes are responsible for coordinating traffic light control and vehicle communication. If the edge device is subject to a man-in-the-middle attack, it may tamper with traffic scheduling instructions, causing regional congestion or even accidents.

– The data layer includes data acquisition, preprocessing, local storage and preliminary analysis, and other data operations performed independently on the device. It abstracts all data-related security threats. It is realistic to analyse security threats for the data layer as follows, but not limited to, data integrity, data noise, data missing, data duplication and data privacy breach. For example, data collection, data preprocessing and basic analysis are all carried out at the device and it is necessary to guard against threats such as data integrity, missing data and duplicate data.

– The network layer includes network communication, API calls, data transfer, model synchronization and collaborative inferencing between devices and between devices and cloud/edge nodes. Unlike the system layer, it operates across distributed endpoints. It is realistic to analyse security threats for the network layer as follows, but not limited to, model theft, model inconsistency and DDoS attacks. During the interaction between cloud-edge-device, cloud/edge nodes are often responsible for model training and global reasoning, and local devices need to download the latest lightweight models from the cloud or edge to achieve local real-time reasoning. Since the model files are transmitted over the network, attackers can intercept model data packets during transmission by attacking the cloud-edge-device communication network, posing a certain threat of model theft.

– Application layer refers to AIoT applications, including front-end user interaction and back-end model decision feedback, such as smart home, smart industry and smart transportation. Unlike all lower layers, it translates model outputs into real-world actions. It is realistic to analyse security threats for the application layer as follows, but not limited to, model performance degradation, model adversarial robustness degradation and model attacks. Due to the limited memory and computing power of AIoT devices, lightweight technology must be adopted to reduce the number of AI model parameters so that they can be deployed on these devices. This may lead to a decrease in model performance, a reduction in adversarial robustness, leakage of model assets and an increase in the threat of AI model attacks.

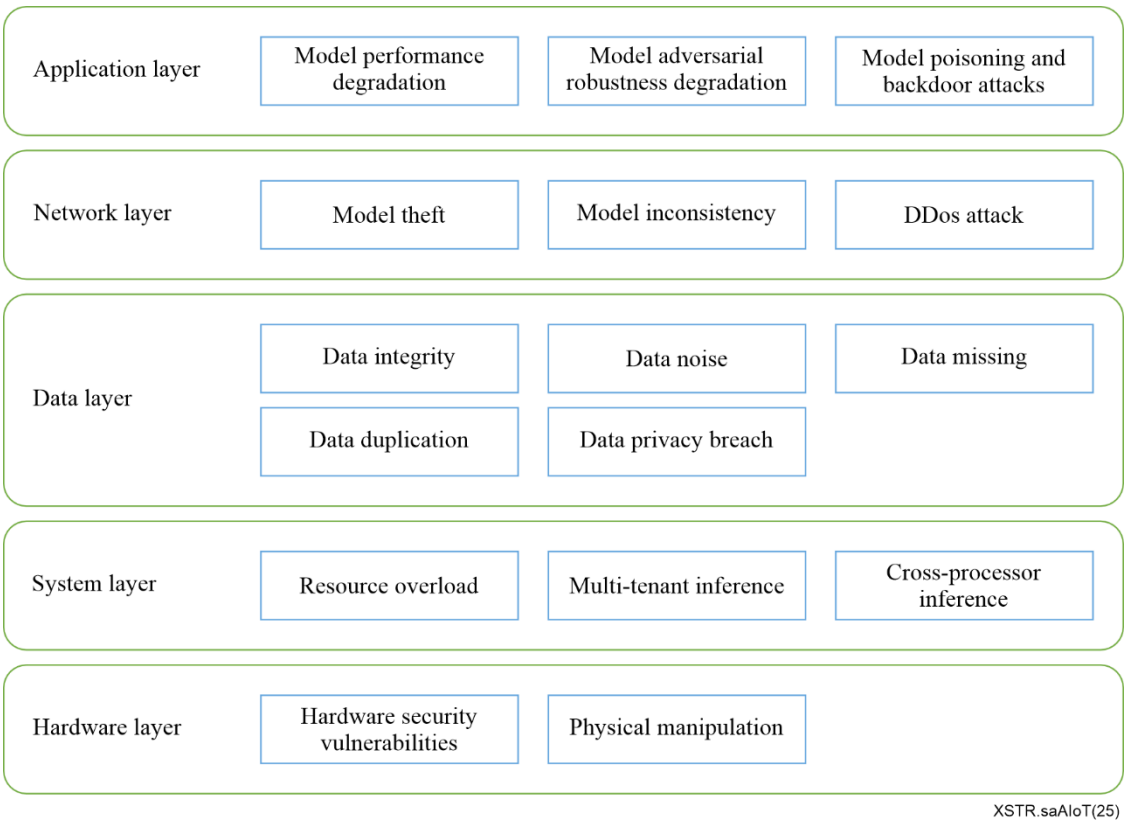

Figure 2 – The landscape for security threats and challenges of AIoT on devices

9 Security threats of artificial intelligence of things on devices

9.1 General security threats

In AIoT systems, general security threats primarily consist of two parts: general IoT security threats and general AI system security threats. These security threats are summarized as follows.

9.1.1 General IoT security threats

The IoT security threats are multifaceted, involving device security, network security, data security, user privacy and the overall ecosystem's security. Many IoT devices come with default passwords or lack necessary security hardening measures, making them easy targets for hackers. These devices typically collect large amounts of personal and sensitive data, such as location, health status and lifestyle habits; if this data is not encrypted or protected during transmission, it can be stolen by malicious attackers, leading to data breaches and privacy risks. Due to their widespread distribution and frequent role as interfaces for multiple systems, IoT devices are prime targets for network attacks like Distributed Denial of Service (DDoS), which can disrupt normal operations. The long production chain involving multiple hardware and software suppliers introduces supply chain risks, where malware can be introduced during manufacturing, installation or upgrades, potentially compromising devices with back doors or malicious programs that threaten cybersecurity. Furthermore, many organizations neglect long-term maintenance and real-time monitoring of IoT devices, making it difficult to promptly detect and respond to attacks or failures. Communication between IoT devices without proper encryption can result in data interception or tampering during transmission, and using outdated communication protocols can introduce additional security threats.

9.1.2 General AI system security threats

9.1.2.1 Data security threats

a) Data security threats in training phrase

During the training phase of AI models, there are multiple security threats that can disrupt the AI model training process through various attack methods such as data poisoning, data tampering, data forgery, label poisoning and feature poisoning. These attacks lead to erroneous predictions and reduced generalization performance. For example, data poisoning manipulates the training dataset to make the model classify malicious examples into desired categories, causing long-term damage until retraining or correction. Data tampering involves attackers modifying the training data, disrupting the normal behaviour of the AI system. Label poisoning deliberately mislabels a portion of the fine-tuning data to confuse the model, leading to biased or incorrect predictions. Feature poisoning modifies or injects malicious data into specific features or attributes of the training dataset, affecting the AI model's decision-making process and making it more susceptible to adversarial examples.

These threats collectively pose significant challenges to the integrity and reliability of AI systems.

b) Data security threats in inference phrase

During the inference phase of AI models, there are multiple privacy leakage security threats. For example, in a membership inference attack, an attacker can determine whether a particular training sample was included in the training dataset by querying the AI model's prediction results. Additionally, AI models have a memorization capability, which may cause them to retain information from their training data, potentially leading to the accidental exposure of sensitive information embedded in their learned weights during the inference phase.

c) Data management security threats

When conducting pre-training, fine-tuning and inference services for AI models, data needs to be transmitted among different entities or departments. This data usually includes various sensitive information and privacy, such as personal identity information, financial data, etc. During the data transmission process, if insufficient security management measures are not taken, attackers may intercept this data and thereby obtain sensitive information of users or organizations.

9.1.2.2 Model security threats

In AI systems, there are various types of attacks that threaten the integrity and privacy of models. For example, model extraction attacks allow adversaries to continuously collect input-output pairs through AI application services to train a similar substitute model, thereby stealing the functionality of the target model. This poses threats of confidentiality breaches or intellectual property theft. Model inversion attacks focus on reconstructing training data from the model's predictions, recovering private features of the model, and potentially inferring specific training samples, which could lead to the leakage of sensitive information. Adversarial attacks involve adding subtle and imperceptible perturbations to input data, causing the AI model to generate incorrect predictions. This affects model performance and can be used to execute unauthorized actions. Model poisoning involves manipulating the training dataset by introducing adversarial examples, causing the AI system to learn incorrect information and produce erroneous predictions. Backdoor attacks embed hidden triggers in the training data, causing the model to exhibit abnormal behaviour when these patterns are activated, giving attackers control over the model's output.

Each type of attack targets different vulnerabilities and can have serious consequences. Therefore, understanding and defending against these attacks is crucial for protecting the security of AI systems.

9.2 Specific security threats in artificial intelligence of things on devices

AIoT is the deep integration of AI technology and IoT, thus will face a series of new-type security threats. This section proposes a security threats analysis framework for the AIoT scenario, which includes five levels respectively: application layer, network layer, data layer, system layer and hardware layer.

9.2.1 Security threats in the AIoT application layer

Although model lightweight techniques have made significant progress in improving model efficiency and reducing resource consumption, they can also introduce a range of new security threats and privacy issues. Below are several security threats.

a) Model performance degradation

The process of lightweighting, whether through pruning, quantization or designing more compact network architectures, can lead to a decline in model performance. Therefore, when deploying AI models at the application layer, it is necessary to fully consider computing resources, model parameters and performance.

b) Model adversarial robustness degradation

Lightweight models, due to their fewer parameters or lower computational complexity, may not be as robust against adversarial attacks as larger models. This means they can be more easily deceived by carefully crafted inputs that appear normal to humans but cause misclassification by the model.

c) Model poisoning and backdoor attacks

Introducing malicious data or performing incorrect pruning during training can result in model poisoning, where the model is deliberately altered to produce erroneous outputs. Additionally, there is the potential for backdoor attacks, where the model gives predetermined incorrect answers when triggered by specific conditions.

In summary, while lightweight techniques offer convenience for deploying deep learning models, it is crucial to carefully consider and address related security and privacy protection issues before actual deployment.

9.2.2 Security threats in the AIoT network layer

Considering the limited memory and computing capabilities of AIoT devices, some devices may not be able to run the most efficient AI models relying solely on their own onboard resources. Therefore, it is still necessary for the device to interact with edge computing nodes or central clouds that have richer resources for model distribution, load balancing, data transmission and service communication. There are the following threats and challenges in the AIoT network layer:

a) Model theft

During AIoT cloud-edge-device interaction, cloud/edge nodes are often responsible for model training and global inference, while local devices need to download the latest lightweight models from the cloud or edge to achieve local real-time inference. As the model files are transmitted over the network, there is a certain threat of model theft as attackers can intercept model data packets during transmission by attacking on the cloud-edge-device communication network.

b) Model inconsistency

During the interaction between cloud, edge and devices, the AIoT devices need to request model synchronization from cloud/edge nodes to ensure the consistency between local inference and global inference. However, due to the instability of the network, various problems may arise during the model download process, increasing the threat of model inconsistency. Firstly, packet loss during network transmission may lead to the failure or unavailability of model data transmission. In such cases, even if a part of the model data reaches the device, it may not be usable due to incomplete data. Secondly, network caching issues may result in the model version not being updated in time, thereby

causing synchronization failure. The device end may still use the old version of the model for inference, failing to obtain the latest model parameters, which can lead to local malfunctions.

c) Distributed Denial of Service (DDoS) attacks

In AIoT, the dimensionality and complexity of network traffic data are higher. In addition to traditional traffic characteristics such as the number of packets per second and the number of bytes per second, there may also be specific data transmission patterns related to models. For example, there are a large number of operations such as model parameter synchronization and data sharing transmission, which generate unique traffic characteristics, making it more difficult to distinguish between DDoS attack traffic and normal traffic. Moreover, AIoT typically has higher real-time requirements for the network compared to IoT that does not require model-driven operations; any delay may affect the training and inference of models. Therefore, when detecting DDoS attacks, it is necessary to find a better balance between real-time performance and accuracy. On the one hand, it is essential to be able to quickly detect attacks and take measures to minimize the impact on model operation; on the other hand, false alarms must be avoided, as they may lead to normal model operations being incorrectly intercepted, affecting the normal operation of the models.

9.2.3 Security threats in the AIoT data layer

In AIoT systems, from data collection to data preprocessing, several key steps are involved. Each step requires careful planning and execution to ensure that the resulting data supports effective decision-making. Throughout this process, it must be paid not only to data quality but also to data security and privacy protection. There are the following threats in the AIoT data layer:

a) Data integrity

This refers to the accuracy and reliability of the collected data, ensuring that all stored data reflects an objective and true state. Achieving data integrity involves adhering to principles such as traceability, clarity, synchronization, originality or authenticity replication and precision.

b) Data noise

This involves interference in the dataset, such as electromagnetic interference or weather-related disturbances, which can affect the results of AI model training and inference.

c) Data missing

Data collected by AIoT devices often contains missing values or incomplete data, leading to the loss of important information.

d) Data duplication

Datasets may contain duplicate records, which can distort analysis results and waste storage resources. Identifying and removing these duplicates is a crucial step in ensuring data quality.

e) Data privacy breach

In AIoT, there are threats of privacy data leakage during cloud-edge-device interactions. On one hand, during data transmission, if adequate security measures are not in place, attackers may intercept data packets to obtain sensitive information. For example, a man-in-the-middle (MITM) attack can insert itself as an intermediary in the data transmission process, intercepting and tampering with data packets. On the other hand, if data is not encrypted during transmission, attackers can easily access and modify it. Even if some data is encrypted, it may still be vulnerable to decryption due to weaknesses in the encryption algorithm or improper key management.

9.2.4 Security threats in the AIoT system layer

In AIoT systems, performing AI inference on devices to make decisions is one of the most fundamental tasks, especially for latency-sensitive applications or scenarios where edge or cloud computing capabilities are unavailable. However, executing inference on devices comes with several threats. There are the following threats in the AIoT system layer:

a) Multi-tenant inference

Multi-tenant inference involves running multiple different AI models simultaneously on an AIoT device, usually from various concurrently running applications. The challenge for AIoT systems is how to effectively manage and handle inference requests from multiple tenants given the limited resources of the device. This includes ensuring fair resource allocation and avoiding performance degradation.

b) Cross-processor inference

AIoT devices can be equipped with multiple heterogeneous processors (e.g., CPU, GPU, TPU), allowing AI models to perform inference tasks across different types of processors within the device. This is because certain sub-computational tasks may need to be allocated to different processors for optimal performance. AIoT systems must efficiently schedule the computational resources of these heterogeneous processors to ensure high-performance and coordinated inference, thereby avoiding system downtime.

c) Resource overload

In AIoT systems, frequent data transmission and task offloading are required between multiple devices, edge computing nodes and cloud servers. If multiple devices request resources simultaneously, it can lead to resource contention, thereby causing resource overload. For example, if multiple devices simultaneously request model synchronization or data processing from the same edge computing node/cloud server, it may result in overload of edge and cloud resources. The threats of resource overload can be specifically categorized as follows: 1) computational resource overload: This occurs when the usage rate of CPUs and GPUs is excessively high, leading to increased task processing delays; 2) storage resource overload: This is characterized by insufficient storage space, preventing data from being written in a timely manner; 3) communication resource overload: This manifests as network congestion, resulting in increased data transmission delays. Regardless of the type of resource overload, it can lead to large-scale service unavailability.

9.2.5 Security threats in the AIoT hardware layer

a) Hardware security vulnerabilities

Some AIoT devices use open-architecture processors (e.g., RISC-V and ARM) to cut costs. However, if a secure boot mechanism is not established, malicious firmware may be implanted, which can alter local inference results. Most general IoT devices adopt closed architectures, so the risk of this kind is relatively low.

b) Physical manipulation

In some AIoT scenarios, such as in smart grids, attackers can interfere with current sensor readings by using a strong magnetic field, inducing misjudgement of AI models on the load status and triggering wrong power-off instructions. General IoT only needs to guard against data tampering, while AIoT also has to deal with model inference errors induced by the physical environment.

Bibliography

[b-ITU-T Y.4000] Recommendation ITU-T Y.4000/Y.2060 (2012), *Overview of the Internet of things.*

[b-ITU-T YSTP.AIoT] ITU-T Technical Paper YSTP.AIoT (2023), *Challenges of and guidelines to standardization on artificial intelligence of things*.

[b-ISO/IEC 20924] ISO/IEC 20924:2024, *Internet of things (IoT) and digital twin– Vocabulary*.

[b-ISO/IEC 22989] ISO/IEC 22989:2022, *Information technology – Artificial intelligence – Artificial intelligence concepts and terminology*.

[b-ISO/IEC 23894] ISO/IEC 23894:2023, *Information technology – Artificial intelligence – Guidance on risk management*.

[b-ISO/IEC 27000] ISO/IEC 27000:2018, *Information technology – Security techniques – Information security management systems – Overview and vocabulary*.

[b-ISO/IEC 27402] ISO/IEC 27402:2023, *Cybersecurity – IoT security and privacy – Device baseline requirements*.

[b-ISO/IEC TR 24028] ISO/IEC TR 24028:2020, *Information technology – Artificial intelligence – Overview of trustworthiness in artificial intelligence*.

[b-ISO/IEC TS 30149] ISO/IEC TS 30149:2024, *Internet of Things (IoT) – Trustworthiness principles*.