Financiado por la Unión Europea
NextGenerationEU

MINISTERIO
DE ASUNTOS ECONÓMICOS
Y TRANSFORMACIÓN DIGITAL

Plan de Recuperación,
Transformación
y Resiliencia

uc3m

6G-RIEMANN-SI Entregable E8

Definition of the use cases for privacy preserving network management

PROGRAMA DE UNIVERSALIZACIÓN DE INFRAESTRUCTURAS DIGITALES PARA LA COHESIÓN UNICO I+D 5G 2021

UNICO I+D

Fecha: 31/7/2022

Versión: 1.0

Propiedades del documento

Id del documento	E8			
Título	Definition of the use cases for privacy preserving network management			
Responsable	UC3M			
Editor	Albert Banchs			
Equipo editorial	**Partner**	**Name**	**Surname**	**Sections**
	UC3M	Albert	Banchs	All
Nivel de diseminación	Público			
Estado del documento	Final			
Versión	1.0			

Historial

Revisión	Fecha	Por	Descripción
1.0	31/07/22	Editor	Final version

Revisor

Equipo revisor	**Partner**	**Name**	**Surname**	**Sections**
	UC3M	Marco	Gramaglia	All

Descargo de responsabilidad

This document has been produced in the context of the 6G-RIEMANN Project. The research leading to these results has received funding from the Spanish Ministry of Economic Affairs and Digital Transformation and the European Union-NextGenerationEU through the UNICO 5G I+D programme. The information contained in this document is provided "as is" without any express or implied warranties, including but not limited to the implied warranties of merchantability and fitness for a particular purpose. The document writer shall not be liable for any damages, whether direct or indirect, arising out of or in connection with the use of this information. The user of this document assumes all risks and liabilities associated with its use and shall indemnify and hold harmless the document writer from any and all claims, losses, damages, or expenses, including attorney's fees, arising from the use of this information.

Table of Contents

Lista de acrónimos

3GPP - 3rd Generation Partnership Project

4G - Fourth Generation

5G - Fifth Generation

AI - Artificial Intelligence

API - Application Programming Interface

CLI - Command Line Interface

DNS - Domain Name System

EGMF - Exposure Governance Management Function

ETSI - European Telecommunications Standards Institute

IaaS - Infrastructure as a Service

MANO - Management and Orchestration

MDA - Management Data Analytics

MDAS - MDA service

NDAF - Network Data Analytics Function

NF - Network Function

NFV - Network Function Virtualization

NSaaS - Network Slice as a Service

O-RAN - Open Radio Access Network

ONAP - Open Network Automation Platform

OSM - Open Source MANO

SA5 - Service and System Aspects Technical Specification Group

SBMA - Service-Based Management Architecture

SD-WAN - Software-Defined Wide Area Network

SDN - Software-Defined Networking

SLS - Service-Level Specification

TS - Technical Specification

Resumen ejecutivo

El uso de la inteligencia artificial y el aprendizaje automático en la gestión y orquestación de redes de comunicaciones, incluyendo redes móviles y el Internet de las cosas (IoT), está convirtiéndose en un requisito para mejorar la eficiencia y la automatización en la operación de estas redes. En particular, la transición de las redes monolíticas como 4G a las redes multi-servicio y multi-tenant como 5G requiere la disponibilidad rápida y flexible de datos para admitir operaciones como el entrenamiento de modelos, la previsión y la clasificación de actividades en la red. La privacidad de los datos es un problema crítico a resolver para garantizar que los datos personales y confidenciales estén protegidos durante el intercambio de información entre diferentes partes interesadas en el ecosistema de la red. La detección de amenazas y el análisis de DNS para detectar ataques cibernéticos es otro tema importante que requiere el uso de técnicas de inteligencia artificial y aprendizaje automático. En general, la gestión y orquestación de redes de comunicaciones requiere el uso de soluciones inteligentes que puedan adaptarse a los requisitos de las redes de próxima generación y satisfacer las necesidades de privacidad de los usuarios y las partes interesadas.

Abstract

The field of network management and orchestration has seen significant advancements with the increasing adoption of automation and artificial intelligence. In particular, the operation of mobile networks has become a requirement for automation, which has been made possible by the availability of countless AI-based solutions. However, rapid and flexible data availability is essential to ensure the success of this automation. The use of AI and automation also extends to the domain of IoT, where predictive maintenance and anomaly detection are increasingly being utilized to improve network management. Furthermore, the analysis of DNS data can be used to detect and prevent phishing and other types of cyber attacks. However, with the increasing use of sensitive data in network management, privacy concerns must be considered and addressed. Overall, the successful implementation of network management and orchestration requires a comprehensive approach that considers the challenges and opportunities in automation, AI, IoT, and privacy.

1. Introduction

In order to ensure the success of future 6G Networks, network management solutions will play a critical role in providing new services and extreme personalization while minimizing resource usage. Achieving this level of efficiency and performance in an increasingly complex environment of infrastructure providers, network operators, and service providers would be impossible without automation algorithms, particularly those based on artificial intelligence.

One of the fundamental requirements for the effective implementation of AI-based network automation algorithms is the availability of data. Standardization bodies such as 3GPP have recognized the importance of data in developing analytics frameworks for network management. For example, the Network Data Analytics Function is a central element for data flows between network management and the 5G Core, collecting and analyzing data to optimize network operations.

However, as the scope of network automation broadens to include more stakeholders in the ecosystem, such as end-users and service providers, protecting the privacy and security of data becomes increasingly challenging. For example, data such as end-user trajectories can reveal personal information such as home and work locations, requiring proper anonymization before being shared with service providers. This is particularly important for providing tailored edge computing services, which require detailed end-user data.

Additionally, data exchanges between service providers and network operators must be conducted while maintaining privacy and security. For example, video streaming providers exchanging data with network operators to optimize network delivery must ensure that their data does not reveal any sensitive business-related activities.

In summary, network management solutions that leverage automation algorithms based on artificial intelligence will play a critical role in ensuring the success of future 6G Networks. However, these solutions must address the challenges of collecting and sharing data while maintaining privacy and security to fully realize their potential.

2. State of the art on network management solution

2.1. Network Management

To ensure access restrictions, 3GPP SA5 has introduced the Exposure Governance Management Function (EGMF), which is responsible for the abstraction, simplification, filtering, and aggregation of management services, including data services. Additional management service abstraction may be required due to a lack of trust relationship between the management service producer and consumer, such as when the service consumer is outside the operator's administrative domain. In the business model known as "Network Slice as a Service" (NSaaS), specific sets of management services must be made available to and consumed by an operator's customer, as specified in TS28.530.

In addition to the above, 3GPP TS28.530 lists lifecycle management services for a communication service instance, including activation, modification, Management Data Analytics (MDA)-assisted service-level specification (SLS) assurance, and termination. The MDA service (MDAS) is responsible for processing and analyzing network data (such as performance measurements, trace reports, QoE reports, alarms, configuration data, and other network analytical data) to detect specific events, make predictions about network performance, and generate analytics reports. The provided analytics reports may also include recommended actions. MDAS instances can be tailored to a specific use case and exposed to external consumers, who can subscribe to customized analytics reports via the EGMF.

To sum up, the Service-Based Management Architecture (SBMA) is a standardized approach to management services provision and consumption that has been introduced in 3GPP Release 15. The architecture defines a management service that offers management capabilities such as generic provisioning, fault supervision, and performance assurance. Management services are made available via a standardized service interface and can be accessed by any authorized consumer. Additional management service abstraction may be necessary to ensure access restrictions, and the Exposure Governance Management Function (EGMF) is responsible for this. The MDA service (MDAS) provides analytics reports based on processed network data and can be exposed to external consumers via the EGMF.

2.2. Network Orchestration

In the past, traditional network domains, including Network Functions (NFs) and the management plane, were already present in legacy networks up to 4G/LTE. However, the introduction of Software-Defined Networking (SDN), Network Function Virtualization (NFV), and containerization technologies have greatly enhanced the orchestration and configuration of these domains.

Resource virtualization is one such enhancement, allowing for on-demand provisioning of network services, following the successful Infrastructure as a Service (IaaS) paradigm. This feature provides the opportunity for creating, re-configuring, and terminating network services seamlessly.

Another critical advancement in network domains is the adoption of network programmability, which has boosted the adoption of API-based access to network configuration. This API-based approach enables flexible reconfiguration of NFs, in contrast to traditional network control through manual CLI-based approaches.

However, the scope of software-driven, application-agnostic management of general-purpose cloud resources such as orchestration and lifecycle management procedures have not been tackled by 3GPP. Vendors and open-source initiatives are currently offering standard-compliant solutions, such as Nokia CloudBand Suite or ETSI NFV MANO provided by OSM or other fora such as ONAP.

Radio resources, on the other hand, require similar orchestration and configuration based on network conditions like load and traffic patterns. For example, the configuration of inter and intra-data center networks is achieved through an SDN controller, and the transport network is managed by SD-WAN. SDN-based controller solutions are also extended to the network domain, as proposed by O-RAN for the access network.

While industry initiatives such as O-RAN have defined radio orchestration procedures, these resources' configuration remains a challenge. In conclusion, the orchestration and configuration of traditional network domains have significantly evolved with the introduction of SDN, NFV, and containerization technologies. Still, there are areas of improvement, such as the configuration of radio resources, that require further attention.

3. Privacy concerns of the network management solutions

The use of automation in mobile networks has become a necessity, just like in other technological fields such as suggestion systems and self-driving. With the availability of numerous AI-based solutions, autonomous network operations are expected to handle the transition from the 4G network to the multi-service and multi-tenant 5G network efficiently. However, for successful implementation, there is a need for the rapid and flexible availability of data to support model training, forecasting, and classification of network activities.

Efforts are already underway to standardize network automation aspects, such as those defined in 3GPP TS23.288, but implementing these in a multi-party setup targeted by beyond 5G communication is still in its infancy. Recent years have seen the development of new technologies for collecting and analyzing data in a centralized or federated manner, with impressive performance. However, all types of data must be treated in accordance with privacy-preserving principles, ensuring the privacy and integrity of personal data when exchanged among different stakeholders in the network ecosystem.

The stakeholders in a multi-tenant network ecosystem are diverse, including end-users, service providers, network operators, and infrastructure providers. These stakeholders exchange data to optimize the operation of the service, the network, or the infrastructure. Service providers may require data related to end-users such as their location, IP addresses, and browsing habits to optimize their services through the use of machine learning techniques. This data is personal and must be secured before being processed. On the other hand, service providers may need to exchange data about their service usage or operation internals with the network operator to optimize network operation. For example, security reports may be exchanged to enhance the network's security or expected load for their service. This data does not contain personal information but needs privacy properties, mainly for business operation.

Analogous situations apply at the infrastructure level, where the same provider can host several network operators. Data coming from operators about analytics related to the expected load, which is fundamental for the correct dimensioning of resources, should be appropriately anonymized before processing. Another challenge is the scalability of the system. With billions of IoT and mobile devices already in use and more to come, it is becoming inefficient to move, store, and process all data at the cloud. Instead, it will be analyzed at the source and in the network, and the necessary privacy guarantees will be enforced directly at the device using a federated approach.

4. Use cases

In the following, we discuss two potential use cases related to this topic.

4.1. Internet of Things

One potential use case for the privacy-preserving analysis of IoT data is in the realm of predictive maintenance. With the increasing prevalence of IoT devices in various industries, such as manufacturing, energy, and transportation, there is a growing need to monitor and analyze large volumes of sensor data in real-time to detect anomalies and predict when maintenance will be required.

However, this presents a challenge in terms of privacy and security, as this data often contains sensitive information about equipment, processes, and operations. For example, in a manufacturing plant, sensor data may reveal information about the products being manufactured, such as their size, shape, and materials used. If this information falls into the wrong hands, it could be used by competitors to gain an unfair advantage.

To address this challenge, privacy-preserving analysis techniques can be used to ensure that sensitive data is protected while still allowing for effective predictive maintenance. One approach is to use encryption techniques to secure the data both at rest and in transit, so that only authorized parties with the appropriate decryption keys can access it. Another approach is to use techniques such as differential privacy[1] to add noise to the data before it is analyzed, so that individual data points cannot be traced back to specific devices or individuals.

By using privacy-preserving analysis techniques, organizations can ensure that they are protecting sensitive data while still benefiting from the insights and predictive capabilities offered by IoT devices. This can lead to increased efficiency, reduced downtime, and ultimately, cost savings for the organization.

4.2. Cybersecurity

The use of DNS analysis to detect phishing and other types of cyber attacks is a critical aspect of cybersecurity that is applicable across a wide range of industries and sectors. With the increasing reliance on technology and digital communications, the risk of cyber attacks has become a major concern for organizations of all sizes and types.

DNS (Domain Name System) is a fundamental component of the internet that translates human-readable domain names into IP addresses. By analyzing DNS traffic, it is possible to detect malicious activity such as phishing attacks, malware distribution, and botnet communications.

Phishing attacks, in particular, are a major concern for organizations as they are a common tactic used by cybercriminals to gain access to sensitive data such as login credentials and financial information. By analyzing DNS queries and responses, it is possible to detect phishing attempts by identifying domains that are similar to legitimate ones but are actually controlled by attackers.

Similarly, DNS analysis can be used to detect other types of cyber attacks such as DNS cache poisoning, DNS tunneling, and DNS amplification attacks. By monitoring DNS traffic patterns, it is possible to detect anomalies that may indicate an ongoing attack and take appropriate measures to mitigate the threat.

The use of machine learning and other advanced techniques can further enhance the effectiveness of DNS analysis for cyber attack detection. By training machine learning models on large volumes of DNS traffic data, it is possible to detect patterns and anomalies that may not be immediately apparent to human analysts.

Overall, the use of DNS analysis for cyber attack detection is a critical aspect of modern cybersecurity. With the increasing sophistication of cyber attacks, organizations must remain vigilant in their efforts to protect their networks and systems from potential threats. By leveraging the power of DNS analysis and other advanced technologies, organizations can stay ahead of the curve and maintain a strong security posture in the face of evolving threats.

As with any analysis of network traffic, privacy concerns must be taken into account when analyzing DNS traffic for detecting phishing and other cyber attacks. DNS queries and responses can reveal sensitive information such as the names and addresses of websites and devices on a network, which can potentially be used to identify individuals and organizations.

[1] Dwork, Cynthia, et al. "Calibrating noise to sensitivity in private data analysis." Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3. Springer Berlin Heidelberg, 2006.

To address these privacy concerns, DNS analysis must be performed in a privacy-preserving manner. This includes measures such as the use of encryption to protect the confidentiality of the DNS traffic being analyzed, and the use of anonymization techniques to protect the privacy of the individuals and organizations whose DNS traffic is being analyzed.

Additionally, it is important to consider the legal and ethical implications of DNS analysis for detecting cyber attacks. Depending on the jurisdiction, there may be laws and regulations that restrict the collection and analysis of DNS traffic. It is important to comply with these laws and regulations to avoid legal and reputational risks.

Overall, while DNS analysis can be a powerful tool for detecting phishing and other cyber attacks, it must be performed in a responsible and privacy-preserving manner to ensure that individuals and organizations are not put at risk.

5. Conclusion

In this document, we discussed various aspects of network management and orchestration. We started by defining the concept of network management and how it has evolved over time to include new technologies such as software-defined networking (SDN) and network function virtualization (NFV). We then explored the challenges associated with managing and orchestrating networks in the context of 5G and IoT, including the need for automation and the importance of privacy.

We discussed the role of AI in enabling autonomous network operation, which requires the rapid and flexible availability of data for tasks such as model training, forecasting, and classification. However, to achieve this, privacy-preserving principles must be followed to ensure the integrity and privacy of personal data, particularly in multi-party setups like the ones targeted by beyond 5G communication.

We also discussed the use of DNS analysis to detect phishing and other cyber attacks. DNS analysis is a powerful tool that can be used to monitor network traffic and identify patterns that indicate malicious activity. However, it is important to balance the benefits of DNS analysis with the need to protect user privacy.

In conclusion, network management and orchestration are complex topics that require a deep understanding of the underlying technologies and their interactions. As networks become more complex and diverse, the need for automation and AI-based solutions will continue to increase. However, we must ensure that these solutions are designed and implemented in a way that respects user privacy and the principles of data protection.