ACCEPTED FROM OPEN CALL

A Tutorial on Next Generation Heterogeneous IoT Networks and Node Authentication

Keshav Sood, Shui Yu, Dinh Duc Nha Nguyen, Yong Xiang, Bohao Feng, and Xiaoning Zhang

Abstract

The extreme heterogeneity in Internet of Things (IoT) networks impacts the network performance, application service quality, and user experience, and has opened many vistas for adversaries to compromise and eventually knock down the networks. The concern for network security and network reliability will further increase in next generation networks (NGNs) such as 5G-IoT by many folds and requires a new level of security approaches. The first line of defense to secure network and its resources is node authentication. However, we note that without alleviating heterogeneity, NGNs cannot fully provide network security or meet the network performance demands. Further, a granular level of autonomy cannot be provided in NGNs, which impacts the IoT node authentication. This makes the authentication issue in heterogeneous networks a more challenging problem. In this article, we first highlight the heterogeneous NGN concept, its impact (particularly on IoT node authentication), and commonly used authentication protocols and industry standards. We also emphasize the need for enhancing NGNs' security by a novel multi-factor approach that also addresses heterogeneity. Unique research challenges and potential opportunities are highlighted.

Introduction

The next-generation networks (NGNs) are composed of millions of heterogeneous physical entities, nodes, multiple operational domains, complex protocols and technologies, different gateways, and so on [1], as shown in Fig. 1. Further, with 5G networking technologies, the networking market also continues to expand, enabling IoT networks in various sectors to provide flexible IoT applications such as airports connected to ePassports, digital facial data to check identity, many innovative smart city projects and critical infrastructure, Agriculture 5.0, the automotive industry, and so on. This shows that the IoT ecosystem has become an extremely important and integral part of our everyday lives. Henc,e the demand for IoT is growing everywhere and even more with the integration of IoT with 5G. Unfortunately, IoT devices have poor security settings, which opportunistically allow cyber-criminals to access IoT nodes and thus sensitive personal data to perform malicious activities [2].

Furthermore, 5G-IoT networks are composed of millions of heterogeneous physical entities, nodes, and more, which increases the threat landscape to a greater extent. Certainly, there are potential threats in such a large-scale heterogeneous communication environment. The key considerations in NGNs' infrastructure are the critical demands of high data flow rates, massive connectivity, low latency (e.g., in autonomous vehicles and remote surgery) as well as the coexistence of multiple technologies [2]. In the presence of complex heterogeneity, the existing models cannot fulfill all of these in every case [1, 2]. The impacts are high end-to-end latency, the required high power computational resources, the impact on the decision making ability of an autonomous and intelligent security and network performance related modules, and more. For example, if the authentication of a wireless node takes more time, it might be an opportunity for a hacker/attacker to knock down the network by simply using network bandwidth as a resource. The time to authenticate a node will increase proportionally with the increases in the degree of heterogeneity. The authentication is the first line of defense to protect network access. If it gets compromised, it could lead to devastating network failure consequences. The impact is more devastating in critical scenarios (where nodes are roaming) such as in battlefield or disaster situations.

The above emphasizes the need to alleviate heterogeneity eventually to enhance or provide strong authentication mechanisms. In the next secion, we comprehensively discuss this observation. Motivated by this, first, we emphasize that there is an urgent need to explore novel multi-factor authentication approaches that must take heterogeneity into account. Second, in the existing literature [3, 4, references therein], although the work on IoT node authentication is well synthesized, they do not fully incorporate the existing industrial protocols. To develop novel authentication solution/s without clearly analyzing (and comparing with) the existing industrial protocols and standards, the argument to emphasize and claiming the need of new authentication systems is not fully justified.

Our contributions in this article are:

- We provide a comprehensive discussion using critical scenarios to encourage researchers to consider complex network heterogeneity while designing novel node authentication strategies. We do not argue that the research in heterogeneous domains is not reported. However, we claim that we provide new insight to see these two factors jointly.
- We discuss key protocols used by cellular and IoT industries for node authentication. Also, key issues in these protocols are given. We provide a reasonable justification to emphasize that the heterogeneity in NGNs is a barrier, yet to be seen by researchers. In relation to this, we emphasize the need to investigate novel robust and efficient protocols or multi-factor

Keshav Sood, Dinh Duc Nha Nguyen, and Yong Xiang are with Deakin University, Australia.

Shui Yu and is with the University of Technology Sydney, Australia.

Bohao Feng is with Beijing Jiaotong University, China.

Xiaoning Zhang is with the University of Electronic Science and Technology of China, China.

Digital Object Identifier: 10.1109/IOTM.001.2100115

 2576-3180/21/$25.00 © 2021 IEEE

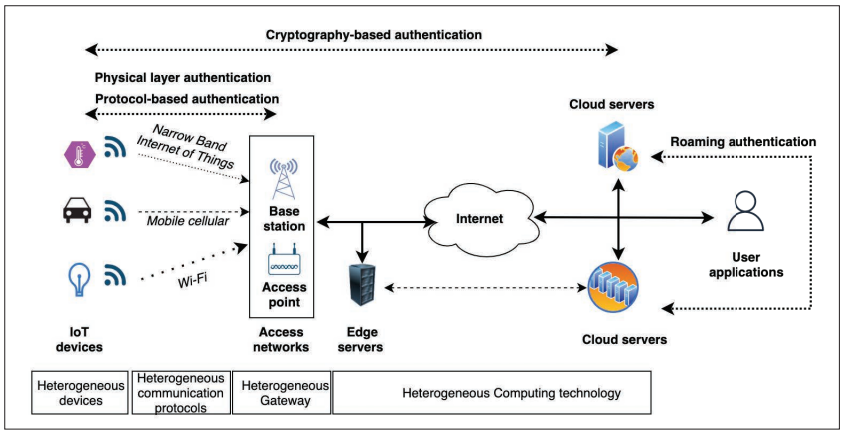

FIGURE 1. A high-level view of a typical heterogeneous network.

authentication approaches.

- A high-level view of the proposed multi-factor authentication scheme is given and discussed in detail. Significant future research directions to analyze this proposal are highlighted.

Key Observation and Benefits: The integral study of both heterogeneity and security has a wider scope of research opportunities, challenges, and applications including routing, authentication, billing, cooperation between nodes, intrusion detection, and more. The joint research in this domain is not mature, and although we have attained great achievements in some aspects of security, there are still many open challenges to be solved due to the extreme complexity of heterogeneous networks and impacts that have yet to be seen/studied. Therefore, it is vital to conduct systematic and comprehensive research on the common security problems (in HetNets), security issues related to management and convergence, multi-domain secure networking, security mechanisms and protocols, and so on.

Motivational Examples: How Heterogeneity Impacts Authentication?

Example 1: The existing systems are not fully autonomous, not highly scalable to accommodate high throughput, and lacks compatibility with NGN architectures.

In NGNs, the huge rate of traffic flows prevents the current intrusion detection systems (IDSs) from monitoring and analyzing the network traffic in real time. For example, Snort is a popular deep packet inspection (DPI) tool that can work effectively on transmission rates up to 1 Gb/s (it starts to discard packets from 1.5 Gb/s), but 5G delivers transmission rates up to 10 Gb/s. Moreover, due to the isolation between network slices (in 5G), solutions based on DPI may not be effective, especially when they are performed separately by individual tenants in an independent manner. In the future, a rapid response time of security applications will be extremely important to enable data analyses on the fly for anomaly detection in or near real time. Researchers have validated how heterogeneity impacts the response time of servers, and gives an opportunity for attackers (e.g., Wi-Fi deauthentication attack [5, 6]) to consume network resource even when the authentication application performs its task in static [2] and handover scenarios [7].

Example 2: Many existing approaches rely on centralized authentication, which is a great challenge. There are several solutions being used for authentication of computers running different operating systems. Different solutions[1] are available for Windows servers and MAC clients including Active Directory as an LDAP server, Quest Authentication Services, and PowerBroker Identity Services, among others. The challenge here is that constant update (multiple patch management solutions) for multiple operating systems is required. We cannot rely only on one vendor's patch release program; this eventually increases the administrative overhead. The HetNet is secured at multiple layers starting from the edge. Consider an example of a simple firewall (for Windows) that ensures the node's protection using Microsoft's Threat Management Gateway (TMG). However, the networks do not have borders, which motivates us to re-evaluate our security policies to protect the network edge.

In this case (heterogeneous multi-domain networking), encryption mechanisms such as IPsec and SSL are effective across multiple operating systems for end-to-end security.[2] Even with this, the end user is expected to have the vendor (company's) security policies configured at the end host device, which discourages network administrators simply because of the complexity of deploying solutions, which eventually impacts the security overhead and authentication.

Example 3: Specifically, the mobility of end nodes from one network to another (especially in vertical networks) imposes another challenge on HetNets to provide secure and seamless connectivity. Due to the heterogeneous nature of NGNs, the time to authenticate a node will increase proportionally with increases in the degree of heterogeneity. We have conducted an experiment using European Telecommunications Standards Institute (ETSI) OSM MANO (5G testbed). We use an ONOS controller, two workstations (one as the virtual infrastructure manager, VIM, interface and the other as OpenStack for VM deployment) of similar configurations (Intel Core i7−7700K @ 4.20 GHz CPU, 64 GB RAM). The experiments demonstrate, as in Fig. 2, how the heterogeneity affects the 5G network management, particularly the virtual network functions' (VNFs') deployment processing time (DT) and CPU utilization (in percent) at varying degrees of heterogeneity.

In Fig. 2, we note that the higher the degree of heterogeneity, the higher is the DT delay, which means the HetNets cannot quickly accommodate sudden change in the underlying network conditions; this results in the higher VNF deployment's process. This will eventually impact the node authentication time (in real-time scenarios), increases the average response time of threat detection deployed on a dedicated server, end-to-end delay, and the application-specific quality of service (QoS) requirements, and adversely affects the optimal use of network resources. We also note that the high degree of heterogeneity has a serious impact on the VIM CPU's performance. Consider a cloud network as an example (with a pay as you go model) in which the user must pay as long as he/she consumes resources. Unfortunately, a) the heterogeneity adds extra communication overhead, which affects the QoS, and b) the network stores an enormous amount of data (coming from heterogeneous devices and gateways), whichincreases the storage costs of the system. Therefore, it is reasonable to say that a high degree of heterogeneity has critical impacts on computational, communication, and data storage cost.

Example 4: The IEEE 802.11 protocol contains de-authentication frames. We understand that the de-authentication attack is a simple distributed denial of service (DDoS) attack with the aim of disconnecting communication between a user and a Wi-Fi wireless access point (WAP). If the adversary knows the medium access control (MAC) address of the victim (e.g., via network sniffing), it sends a de-authentication frame (with spoofed address of the attack target) to a WAP. In the context of heterogeneity, different wireless devices from different vendors brings complex heterogeneous MAC-level behavior as it expands the attack surface to a great extent. For example, these devices may not be conformant with the 802.11 standards, impacting the wireless network operations in many ways (e.g., unfair band-

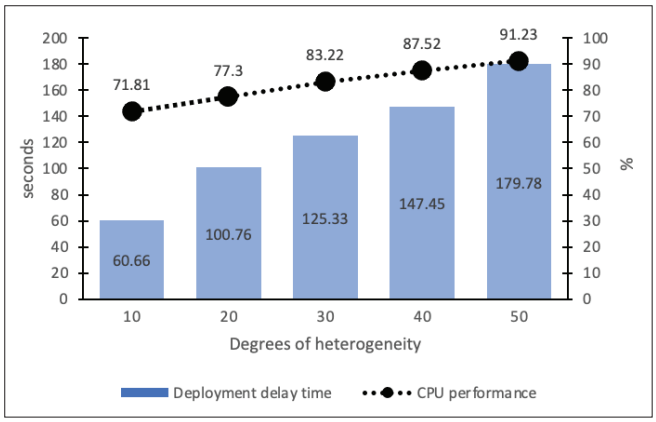

FIGURE 2. The impact of heterogeneity on VNF deployment delay and CPU utilization. Here, the degree of heterogeneity is defined as the complexity of heterogeneous elements in the experimental design. In our experiments, the degree of heterogeneity varies from 10 to 50 in which degree 10 means 10 VNFs. We consider each VNF as a cluster of small VMs that are heterogeneous in terms of RAM size, operational load, and switching hierarchy. We use the basic three-layer switching hierarchy. The number of VMs vary (2–6 in our case) in the clusters to introduce load and complexity.

width allocation, potential break-down of MAC functionality). This increases the authentication time of legitimate nodes, and with the excessive unfair use of a channel, a user can experience packet loss or simply loss of connectivity. This has serious impacts in many critical scenarios and application use cases.

IoT Authentication: Industry Initiatives, Authentication Use Cases, and Challenges

The industry initiatives in this problem domain are given in Table 1. In this section, we discuss critical authentication use cases and challenges.

Scenario 1 (Operational Technologies, OT): In this scenario, we understand the role of authentication and related challenges in manufacturing sector (Industrial IoT, IIoT, or cyber physical infrastructure). To improve the operational efficiencies in the diverse manufacturers' operational networks, which includes production monitoring, inventory management, predictive maintenance, and intelligent logistics, device-to-device connectivity is the norm. This sector is no different than others and is vulnerable to the growing cyber threats that can endanger manufacturing control systems drastically, especially those in the chemical, steel, and petroleum industries. The manufacturing processes in this domain are very sensitive, involving potentially unstable chemicals and high temperatures, and unfortunately, adversaries (in the absence of strong authentication mechanisms) could potentially mislead the endpoints controlling these processes and cause devastating situations, for example, critical plant/operation failures leading to injury or even loss of life at worst. This has emphasized the urgent need for designing and employing strong robust authentication/security.

It is noted that strong authentication is not as common as it should be in this sector for many reasons. For example;

1. Communications protocols do not require authentication; typically, Modbus is most preferred industrial automation protocol, which lacks any form of device authentication; this eventually has an impact on the communications integrity.
2. Further, manufacturing equipment often has long replacement cycles (roughly 10–20 years). Older devices near their end of life are often not fully integrated with security and do not have enough computing capabilities to perform resource-expensive cryptographic-based authentication.
3. One of the key and critical features of this sector is low end-to-end latency to ensure real-time operations of critical processes. We emphasize that the security solutions should not be time consuming or introduce latency, including authentication.

Despite these key limitations, the operational manufacturing sector/networks still have alternatives available for performing strong authentication. New endpoints are being designed with enough hardware to implement public key infrastructure (PKI)-based authentication. Also, authentication functionality gateways are being considered as a promising solution to offload computation-expensive tasks. In recent years, elliptic curve cryptography has grown among PKI-based methods to provide a security level equivalent to RSA with shorter key lengths, requiring less computing power and storage capacity to generate and protect keys. Many frameworks provide guidelines to manufacturers in device authentication, such as NIST SP 800-53, the Industrial Internet Consortium's (IIC) Industrial Internet Security Framework, and the IIC's Endpoint Security Best Practices document. We still observe that even though there are many frameworks, manufacturers typically choose cryptography-based device authentication approaches. Furthermore, these methods are popular, so the pace of adoption (of crypto solutions) could be slowed over the next few years by long replacement cycles of manufacturing endpoints and the need to evaluate alternate forms of authentication to optimize latency and cost. As we have mentioned, the attacks in this sector can be prevented by stronger authentication, and the crypto solutions are still time- and resource-constrained, so an alternate solution should be tested.

Scenario 2 (Authentication during Roaming): The flexibility in 5G due to software defined networking/network function virtualization (SDN/NFV) also adds attractive features to authenticate a user equipment (UE) using 3rd Generation Partnership Project (3GPP)-specified authentication protocols, 5G-AKA (in LTE it is EPS-AKA) and the Extensible Authentication Protocol (EAP) Authentication and Key Agreement (EAP-AKA) protocols to authenticate *things* in 3GPP and non-3GPP networks.[3] In previous networks (before 5G), 3GPP-specified authentication protocols were mainly used for node authentication. Authentication in multi-domain (multi-vendor-operated) networks is also possible as some organizations provide this flexibility. Thus, authentication and/or authorization (e.g., separate authentication/authorization server and credentials) can be performed independently by the external network operators even before the primary network allows a UE to connect to the external network using EAP request that the external network perform a secondary authentication/authorization and only permit connectivity after the external network approves [7].

Figure 3 provides a critical scenario to emphasize the need for node authentication while roaming in HetNets. To offer seamless network connectivity to an IoT node, traffic offloading is essential in both horizontal (within the same network cells) and vertical (from one network domain to another) networks. Horizontal roaming is understood as end-device mobility across different cells/zones administered by the same organization. Vertical network roaming is understood as end-device mobility across different administrative domains, regardless of whether it is in the same country or not. Seamless connectivity is the key metric of any network performance evaluation, especially in critical situations (battlefield, disaster, etc.), where continuous surveillance and monitoring of humans and assets is essential using IoT end nodes (unmanned aerial vehicles, surveillance sensors on smart vehicles, etc.) while roaming. In Fig. 3, a vehicle is moving from Zone A to Zone B. Service provider A provides the channel connectivity to this vehicle via a base station using cellular technology. Zone B's service provider provides connectivity to this vehicle via satellite links (note that the technologies as well as service providers are different).

Note that before successful offloading, the end node (vehicle) is only allowed to enter from home to foreign network after

Device management protocols
Protocols for IoT authentication leverages Public Key Infrastructure (PKI) capabilities and OAuth 2.0. Lightweight connectivity protocols, to manage end-end communication, an XML-based chat protocol (XMPP), Constrained Application Protocol (CoAP), or MQ Telemetry Transport (MQTT) are popular.
For device provisioning, authentication, and automation, the cellular or mobile organization are using Open Mobile Alliance Device Management (OAM-DM), LWM2M (Lightweight Machine to Machine), and TR-069 protocols. The other most common protocols are Kerberos and SSL/TLS.
The MQTT protocol and its variant MQTT-SN is a lightweight and robust protocol, and gives less overhead. OAM (its versions) is also a lightweight and structured protocol as well as the most preferable for low-capacity IoT devices in both static and mobile scenarios.
LWM2M is built on top of CoAP and is suitable for cellular networks especially in sensor network scenarios.
TR-069 specifications/protocol is heavy and very complex but still very popular for gateways and telecommunications devices.
Hardware-based Industrial IoT security solutions
ARM technology: ARM provides TrustZone for Cortex-A and Cortex processor to provide trusted software by hardware. For secure communication, the firmware includes trusted boot and secure runtime for switching between the secure and non-secure domains.
GSMA IoT SAFE: It provides secure IoT communications by leverages the physical SIM of mobile networks as a hardware *root of trust*.
Rapid7 project: It employs the Metasploit framework for exploiting the vulnerabilities from raw wireless and direct hardware.
Remarks: The IoT devices authentication is becoming more complex issue due to the resource and cost constraints of *things*. This leads the hardware-based security to rapidly becoming an industry-wide standard. Particularly, the SIM based *root of trust* based approaches is widely being acceptable by industries across the globe. Also, using secure boot mechanisms, the IoT organizations must verify the device software and firmware.
5G-IoT authentication projects
INSPIRE-5Gplus: This (the Horizon 2020 project) provides a fully automated end-to-end (E2E) network and service security management framework to manage 5G networks across multiple domains such as radio access network (RAN), core network (CN), and mobile edge computing (MEC).It contains several security management domains (SMDs) via security functional modules in each SMD which is responsible to support intelligent security automation of resources and services using software-defined security orchestration and management. The frameworks leverage on various emerging techniques and being utilizing as a Zero-Touch security management solution for 5G.
Verizon's 5G security: Their approach is a driving factor for any 5G enabled networks. It includes Enterprise Protections, Partnerships, and Global Backbone, to leverage on the existing security capabilities. It employs the new security features of 3GPP's 5G standards around the different parts of the network: User Equipment (UE), RAN, and CN. Finally, it utilizes the 5G'new capabilities such as Network Slicing, Orchestration, and Edge Computing to enable the new security services.
Tamarin prover: In [8] authors have also argued that 5G-AKA (Authentication and Key Agreement protocol) is vulnerable to the linkability attacks. An adversary can track the victims' mobile devices via this attack to compromise the privacy of the target users. To address this concern authors proposed a countermeasure with the inherent key encapsulation mechanism of ECIES (i.e., ECIES-KEM). They have used Tamarin prover (formal verification tool), to prove the effectiveness of the approach.
Remarks: The 5G Authentication and Key Agreement (5G-AKA) protocol using cryptographically approaches to guaranteeing two things. The *first* is to ensure that the subscriber's home network operator authenticates a) the UE (or IoT device in IoT context) as well as the roaming network the (that) UE is joining. It does not support that only the roaming network performs authentication. This will stop UEs from being tricked into getting connected any unauthorized networks. The *second* thing is it incorporates procedures to ensure that a UE is actually connected to the authorized roaming network. Eventually, it ensures that information about the UE and subscriber needed to establish a network connection [(e.g., Subscription Permanent Identifier (SUPI)] are only shared with authorized partner networks. Some organizations (such as Verizon's 5G network) are implementing these authentication procedures by default.
Standardization efforts
3GPP standards: They define mobile telecommunication technologies. One of the 3GPP research group's focus is to design global standards for interfaces of 5G networks. Within the 3GPP Technical Specification Group Service and System Aspects (TSG SA), the objectives of 3GPP TSG SA WG3 (SA3) is to define the requirements and specifying the architectures and protocols for security and privacy in 3GPP systems. Particularly, SA3 defines the architectures and protocols for security systems for IoT and vertical industries while SA5 is responsible for management, orchestration, network slicing, etc.
GSMA IoT SAFE: It ensures that the potential proposed approaches must meet the standards and guidelines set by GSMA IoT SAFE, (IoT SIM Applet For Secure End-2-End Communication) standards. The IoT SAFE provides the IoT end-to-end security solution by leveraging SIM as a root of trust. In the IoT SAFE, a SIM is embedded with a digital certificate to support TLS/(D)TLS protocol for establishing secure communication channels between devices and cloud servers.
Many industry and government organizations are working to enhance IoT security, for example, Cloud Security Alliance (CSA), Groupe Speciale Mobile Association's (GSMA), IEEE Standards Association, ioXt Alliance, Open Web Application Security Project (OWASP), U.S Department of Homeland Security, IoT Alliance Australia (IoT AA), etc.
Remarks
1. The authentication solutions must be integrated with other 5G Core Access components such as Access and Mobility Management Function (AMF) which handles the connection and mobility management tasks in 5G networks to ensure the authentication of devices under roaming. 2. Any proposed approach must assist the AMF abilities by authorizing mobility devices and informing AMF and SMF (Session Management Function). 3. For device authentication the cryptography-based authentication schemes are still very popular, and they prefer hardware-based authentication. In security context (by industries) hardware-based, i.e., "root of trust" is considered as a promising solution. The fifth generation (5G) mobile network is expected to offer massive IoT applications and services. We encourage potential readers to follow the given references and footnotes and see references therein [9]. 4. Although the alternate solutions are needed, but we do note that the solutions must meet the current industry standards (for protocols design to its implementation and deployment architectural design), example 3GPP (the 3rd Generation Partnership Project) and GSMA IoT SAFE. The IoT SAFE is advantageous to the IoT device security however there are still some concerns. Both SIM and IoT devices are resource constraints, and it becomes the fragile vulnerability of the cyber-attacks. Adversaries can attack the devices and steal digital certificates. 5. Moreover, TLS certificates technology also has weaknesses such as the certificate store poisoning or Certificate Authorities attacking. These issues potentially compromise the IoT SAFE in real-world applications. Therefore, novel/alternate multi-authentication solutions to the IoT applications to maintain the end-to-end security as well as to improve the "root of trust" of the IoT SAFE standards are required to be investigated. 6. Also, we need to ensure that the solutions not only secure the communication between the IoT devices and cloud servers but also authorize the IoT devices in the sessions.

TABLE 1. An overview of 5G-IoT industrial authentication protocols, security solutions, standardization efforts, and our remarks.

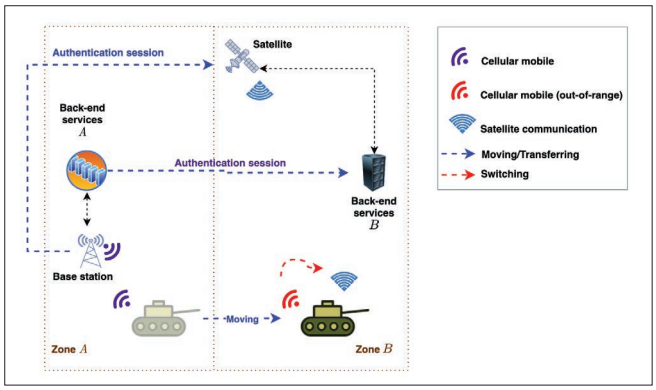

FIGURE 3. A critical scenario illustrating the need for authentication while roaming.

verifying the node's authenticity. In this case, the authentication sessions exchange seamless offloading/mobility management between a home network and a foreign network. We note that for business continuity or data availability the authentication process is more critical while roaming in vertical networks as discussed in the previous sections. Due to this if the end node is not seamlessly connected then the end node may simply lose the connectivity. Eventually, this could have devastating impacts in certain critical situations.

Some ongoing efforts, such as those being accomplished by the Internet Engineering Task Force (IETF) working group IPv6 over LP-WAN, are dealing with interoperability across networks to establish trust across different vertical network domains [3]. In view of this research problem, the LoRaWAN architecture (version 1.1) provides support for simple handover roaming across different LoRaWAN networks as an early approach to address interoperability between networks. We note that it requires explicit agreement negotiations between different administrative domains, in an ad hoc manner, case by case, exchanging sensitive security information. Therefore, roaming is very difficult in practice (let alone the authentication). Overall, we note that the existing solutions (to authenticate devices during vertical roaming) do not fully incorporate approaches that address heterogeneity, which means they cannot effectively enhance the network security of large-scale networks. Given this, we realize a need to investigate a lightweight (to enhance the performance) and secure solution to support IoT node authentication in vertical networks (we start from in-country scenarios). Overall, the authentication is necessary for the establishment of trust for inter-domain communication.

Scenario 3 (MFA Journey): Recently, the COVID-19 pandemic has transformed the way of working (onsite, remote, roaming). To discuss different unique scenarios and related authentication approaches, we show the MFA journey in Fig. 4 to discuss the need to investigate appropriate authentication solutions in different heterogeneous nodes' operational scenarios.

Case A (onsite and remote working). We assume that Bob works onsite, due to organizational regulatory requirements, Bob is only allowed to use a corporate laptop for MFA. In this case, the most preferable authentication approaches are OTP push or 3rd party authentication. Also, if Bob is using his mobile device for MFA, but his device contains highly sensitive (personal and corporate) information, PKI-based authentication approaches are effective. Now assume that Bob is allowed/required to use shared devices; then there are two cases: a) his personal phone is not allowed onsite assuming Bob is a call center help desk worker, and b) the onsite location has no connectivity (assuming Bob is a lab worker). In the former case, FIDO, biometric, hardware-based, and voice-recognition-based MFA are effective, and in the latter case, FIDO-based MFA is applicable. This is also valid for remote working scenarios where Bob uses shared devices for working.

Case B (roaming scenarios). Consider Bob is traveling and doing his corporate work using his personal laptop (BYOD concept). He is not willing to use any corporate app for MFA because of privacy reasons. In this use case, the pattern-based-Google authenticator, email, and SMS-based MFA approaches are effective. These scenarios emphasize that we must identify in which scenario each MFA approach would be better; therefore, use-case-based authentication approaches need to be investigated and deployed effectively.

Research Opportunities

We divide the research into two categories: traffic-based authentication and physical unclonable function (PUF)- based node authentication approaches.

First category (traffic-based authentication models): The approaches in this category collect and analyze IoT data streams to design intelligent IoT node authentication systems based on which they assess the node legitimacy status. But because of the change in underlying network topology, nodes' features, and so on, the model takes time to update and respond. In NGNs, the latency is expected to be less than microseconds. If the model cannot dynamically update itself in the strict time requirement, the model cannot maintain the network and user security. Previous researchers used the two best-known datasets, DARPA and KDD Cup [10]. These datasets were collected a decade ago and do not include the complex attributes of today's network, let alone next generation networks. Potentially, the University of New South Wales (UNSW) Canberra's systematic testbed constitutes heterogeneous data sources and the telemetry datasets collected from IoT devices, such as green gas IoT and IIoT actuators, and includes several normal and cyber-attack events from IoT networks. The testbed uses multiple virtual machines and hosts that have Windows, Linux, and Kali Linux operating systems to manage the interconnection between the three layers of IoT, cloud, and edge/fog systems. The testbed is also able to create new realistic datasets as desired by the user. We believe this is a good strategy to integrate/collect the dataset and test on 5G testbeds (ETSI OPEN MANO). The testing of any new hypothesis to enhance network security using this 5G dataset (or others available in the market) will be valid on NGNs. Mining heterogeneous data streams provides valuable insights into user patterns and behavior that can be used to design an intelligent security system. The ability to detect anomalies in much larger data and stream sizes in real time in the next generation networks is a major challenge.

We note that to remove complicated and time-consuming tasks such as data collection and model training and anomaly detection in real time, it might be promising to decouple high-level planes of learning and decision making from the low-level planes. By doing this, we can dramatically decrease the complexity for real-time data processing continually and boost performance of security models. To do this, the potential proposed framework can be placed as a separate module (to reduce latency and allow real-time data analysis) in 5G network orchestration (as NFV in MANO architecture). The assumption is that the orchestration of each network slice would be independent, or each distributed network would have its own network slice and each network slice would have its own independent orchestration module for node authentication.

Unfortunately, the above solution will not be highly efficient in all scenarios. There are some key reasons for this:

- Automation works on one aspect only.
- Cross-layer automation and intelligence is required.
- The traffic and node features are dynamic.
- Most often, we do not consider physical security features in security automation.

We consider an example of autonomous IoT access control models. In autonomous vehicles, a plethora of sensor devices sense diverse sensor applications to realize their function.

For example, advanced driver assistant systems (ADASs) use data from cameras, light detection sensors, radio detection, as well as vehicle status captured from the controller area network. This collected sensed data is used not only by the ADAS, but by multiple third-party applications. A malicious attempt at any one sensor might impact the overall intelligent decision making of the vehicle, and eventually the security and safety of the vehicle. In this case, the access control model cannot immediately identify the source/entry point of the malicious data stream as it only works in one aspect. The different access patterns on different vehicular networks are more complex to analyze as no one existing model can be applied on another vertical designed and connected by another vendor/technology. Thus, an open, fine-grained, and flexible autonomous access control model is required in this case. The autonomous security models should be dynamic and must be able to change with the change in context [11].

We reiterate that the conventional approaches for IoT security using autonomous models are not fully compatible, intelligent, or autonomous in real time. The primary reasons that conventional approaches will lack applicability in future networks is the high level of heterogeneity in future network nodes, technologies, and protocols. IoT networks are heterogeneous and will be even more so with the integration of 5G and 6G, which in themselves generate extremely heterogeneous sensory data. Unlike conventional IoT data streams, heterogeneous IoT data streams often involve varying modality of data in one set, which makes it very challenging to monitor objects and analyze data. Researchers have noted that the heterogeneous data streams affect the performance of intelligent autonomous IoT access control models and that even within one IoT system, a device typically has various options for communicating with other nodes.

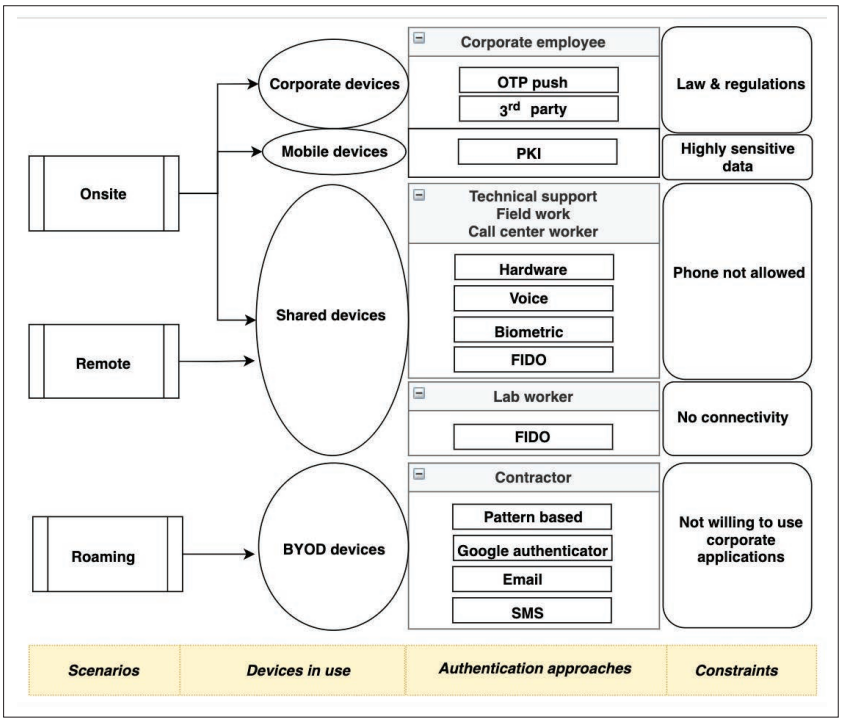

FIGURE 4. Multi-factor authentication (MFA) journey. For IoT devices MFA is an inevitable security solution. Almost all IoT devices are Internet connected and need to be protected as most of the time they become the point of access on open Wi-Fi networks. It is important to provide some degree of restriction and control to the IoT devices in any scenario where IoT devices are located (onsite, remote, and roaming). Use of static passwords to enable the connectivity is no longer the industry choice, and solutions beyond this (like MFA) are needed to ensure that an event that is occurring requires authentication or authorization. Further, the PUF-based MFA solution creates a unique profile of IoT devices as they use physical elements such as fingerprinting features of the devices.

Therefore, with the high degree of heterogeneity expected in future large-scale systems, a research approach driven by specific modalities or service needs is unlikely to lead to coherent architectural solutions for the overall problem of the next generation's heterogeneous networks. In fact, a robust and effective solution is practically impossible for NGNs.

Second category (PUF-based authentication approaches): We advocate that the use of physical unclonable functions (PUFs) of the wireless nodes is a promising alternative. In particular, the physical layer security must be considered. PUF-based systems are extremely robust and secure at low cost [12, 13]. These features are the variations in the manufacturing process to uniquely identify silicon chips. This is almost impossible to replicate these features as they are inherent variations; however, PUF-based solutions mostly employ deep learning models for classification (legitimate or non-legitimate nodes), which are not prone to machine learning attacks [14, 15]. PUF-based approaches can be used as standalone or multi-factor authentication in conjunction with application layer features. Our proposed framework (Fig. 5) is based on these observations.

The proposed 5G AUSF authentication module is integrated into the 5G network slicing and resides at the radio access network (RAN) segment. Each slice has its own orchestration module. The authentication module is a sub-module of each orchestration module that consists of two parts: a) the feature collection and b) anomaly detection. Using deep learning-based models, we collect RF features of underlying IoT nodes at base stations or gateways. The extracted features are then used to train a deep learning model (anomaly detection module) so that the model learns these unique RF features of the legitimate devices in the network. For anomaly detection, an auto-encoder can be used.

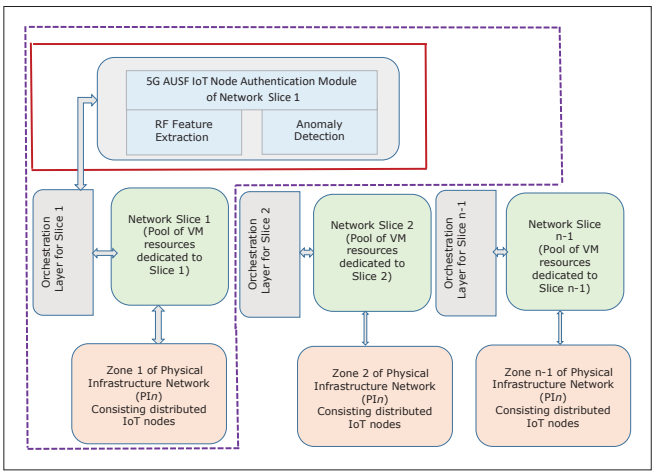

FIGURE 5. A high-level framework of the proposed slice-specific 5G authentication server function (AUSF) IoT node authentication strategy. The proposed strategy is part of each orchestration module. Each slice has dedicated virtual resources (VMs) as well as its own network management orchestration module. End nodes in each distributed underlying IoT network (divided in zones) are only connected to one dedicated network slice at one time.

The key benefits of this design are:

1. As the data computational load is shifted to network slice resources, it does not require any additional hardware at the Tx side, which means no (zero) computation/communication overhead on the resource–constrained IoT nodes.
2. This RF signature-based approach is independent of the upper layer protocols and thus addresses the heterogeneity issue of 5G–based IoT systems and does not degrade performance.
3. It is well suited for slice–isolated 5G–based IoT networks and can be effectively integrated into the standard architecture of 5G networks; thus, it easily enables network providers to offer authentication service to IoT applications and cloud–based services (e.g., in the form of a multi–factor authentication mechanism).
4. It is easier to manage the IoT node authentication requests (with different security requirements) using a slice–based approach.
5. It offers a high level of efficiency because the task of authenticating a huge number of IoT devices is shared between individual learning models run by different slices. As a result, the learning model used in the AD module does not need to learn the features of all the IoT devices in the network.

We note that proper integration and performance evaluation into the existing cloud (virtual) management technologies (e.g., ETSI OPEN MANO) are required. A proper threat model needs to be presented to establish the reality of the problem to highlight the attacker and its capabilities as well as the communication medium. This further requires evaluating the performance as we propose to integrate the solution in a network slice. The outcome will be a proof of concept to demonstrate how the proposal would fit in with a real-life MANO.

Summary and Future Works

We have emphasized that network heterogeneity has critical impacts on the IoT node authentication process. We have given a high-level view of our proposed multi-factor authentication scheme, which jointly addresses heterogeneity and helps enhance performance of authentication applications. In the future, we aim to investigate a secure and lightweight authentication process to alleviate network heterogeneity to provide secure and seamless handover to end nodes in vertical networks. Further, a theoretical model will be investigated to transform heterogeneous networks into homogeneous networks so that the existing theoretical models can be applied easily.

References

[1] K. Sood *et al.*, "Plug-In Over Plug-In Evaluation in Heterogeneous 5G Enabled Networks and Beyond," *IEEE Network*, vol. 35, no. 2, Mar./Apr., 2021 pp. 34–39.

[2] Q. Cui *et al.*, "Edge-Intelligenceempowered, Unified Authentication and Trust Evaluation for Heterogeneous Beyond 5G Systems," *IEEE Wireless Commun.*, vol. 28, no. 2, Apr. 2021, pp. 78–85.

[3] X. Wang *et al.*, "Attacks and Defenses in User Authentication Systems: A Survey," *J. Network and Computer Applications*, 2021, p. 103,080.

[4] H. J. Jo and W. Choi, "A Survey of Attacks on Controller Area Networks and Corresponding Countermeasures," *IEEE Trans. Intelligent Transportation Systems*, 2021, pp. 1–19.

[5] M. Vanhoef, "Fragment and Forge: Breaking Wi-Fi Through Frame Aggregation and Fragmentation," *30th fUSENIXg Security Symp.*, 2021.

[6] M. Conti *et al.*, "Fadewich: Fast Deauthentication Over the Wireless Channel," *2017 IEEE 37th Int'l. Conf. Distributed Computing Systems*, 2017, pp. 2294–2301.

[7] Y. Zhang *et al.*, "Robust and Universal Seamless Handover Authentication in 5G Hetnets," *IEEE Trans. Dependable and Secure Computing*, vol. 18, no. 2, 2021, pp. 858–74.

[8] Y. Wang, Z. Zhang, and Y. Xie, "Privacy-Preserving and Standard-Compatible Fakag Protocol for 5G," *30th fUSENIXg Security Symp.*, 2021.

[9] The Evolution of Security in 5g; https://www.5gamericas.org/wp-content/uploads/2019/07/5GAmericas_5G_Security_White_Paper_Final.pdf

[10] L. Yang and A. Shami, "A Lightweight Concept Drift Detection and Adaptation Framework for IoT Data Streams," *IEEE IoT Mag.*, vol. 4, no. 2, June 2021, pp. 96–101.

[11] H. Peng *et al.*, "Vehicular Communications: A Network Layer Perspective," *IEEE Trans. Vehic. Tech.*, vol. 68, no. 2, 2019, pp. 1064–78.

[12] D. Liu *et al.*, "Research on End-to-End Security Authentication Protocol of NB-IOT for Smart Grid Based on Physical Unclonable Function," *2019 IEEE 11th Int'l. Conf. Commun. Software and Networks*, 2019, pp. 239–44.

[13] B. Li *et al.*, "Physical-Layer Security in Space Information Networks: A Survey," *IEEE IoT J.*, vol. 7, no. 1, 2020, pp. 33–52.

[14] N. Shah *et al.*, "Introducing Recurrence in Strong Pufs for Enhanced Machine Learning Attack Resistance," *IEEE J. Emerging and Selected Topics in Circuits and Systems*, vol. 11, no. 2, 2021, pp. 319–32.

[15] P. Gope, B. Sikdar, and O. Millwood, "A Scalable Protocol Level Approach to Prevent Machine Learning Attacks on Puf-Based Authentication Mechanisms for Internet-of-Medical-Things," *IEEE Trans. Industrial Informatics*, 2021.

Biographies

Keshav Sood (keshav.sood@deakin.edu.au) is currently a lecturer at Deakin University, Melbourne, Australia. Previously, he worked as a research fellow in the Advanced Cyber Security Engineering Research Centre at the University of Newcastle, New South Wales, Australia. He worked on a project funded by the Defence Science and Technology (DST) Group, Australia. He was a trainee with the Terminal Ballistic Research Laboratory (TBRL, DRDO, Ministry of Defense) in Chandigarh, India. He is a Professional Engineer with Engineers Australia accreditation.

Shui Yu [SM] (shui.yu@uts.edu.au) is a professor in the School of Computer Science, University of Technology Sydney, Australia, and a guest professor at Zhengzhou University, China. He is currently serving on a number of prestigious Editorial Boards, including *IEEE Communications Surveys & Tutorials* (Area Editor) and *IEEE Communications Magazine*. He is a member of AAAS and ACM, and a Distinguished Lecturer of IEEE Communication Society.

Dinh Duc Nha Nguyen (nguyendinh@deakin.edu.au) received his Bachelor's degree from the Posts and Telecommunications Institute of Technology, Vietnam, and his Master's degree from Queensland University of Technology, Australia. He has been awarded a place on the Dean's list of excellent academic performance two consecutive times in 2020. He has more than six years of industry experience, mainly as a network analyst and software engineer. He is currently pursuing a Ph.D. degree at Deakin University, Melbourne, Australia.

Yong Xiang (y.xiang@deakin.edu.au) received his Ph.D. degree in electrical and electronics engineering from the University of Melbourne, Australia. He is a professor and the director of the Artificial Intelligence and Image Processing Research Cluster, School of Information Technology, Deakin University. He is an Associate Editor of *IEEE Signal Processing Letters* and *IEEE Access*.

Bohao Feng (bhfeng@bjtu.edu.cn) is currently an associate professor with the School of Electronic and Information Engineering, Beijing Jiaotong University. His research interests include software-defined networks, network function virtualization, service function chaining, and satellite communications.

Xiaoning Zhang (xnzhang@uestc.edu.cn) received his B·S., M.S., and Ph.D. degrees in communication and information engineering from the University of Electronic Science and Technology of China, Chengdu, in 2002, 2005, and 2007, respectively. He is currently a professor with the School of Information and Communication Engineering, University of Electronic Science and Technology of China. His research interests include network design, software defined networks, and network function virtualization.

Footnotes

[1] https://support.apple.com/en-au/guide/deployment-reference-macos/iorbeda89d1d/web}

[2] https://www.zte.com.cn/global/about/magazine/zte-communications/2008/3/en$_1$2/162489.html

[3] https://www.cisco.com/c/en/us/support/docs/wireless-mobility/wireless-lan-wlan/116493-technote-technology-00.html