Conference Proceeding Issue Published in International Journal of Trend in Research and Development (IJTRD), ISSN: 2394-9333, www.ijtrd.com

Challenges and Intelligent Authentication in 5G Networks Based on Machine Learning Techniques

P.Joseph Joshua Giftson,
B.E. Computer Science, II Year, St.Joseph's College of Engineering, Chennai, Tamil nadu.

Abstract: FIFTH Generation (5G) mobile communication technologies are on the way to be adopted all over the world as the next mobile network. Hence it is important to analyze its impact on the landscape of computing and data management. The increase in low-cost devices and access points with increased mobility and heterogeneity leads to the extremely complex and dynamic environment of 5G-and beyond wireless networks. Specifically, due to the open broadcast nature of radio signal propagation, widely adopted standardized transmission protocols, and intermittent communication characteristic, wireless communications are extremely vulnerable to interception and spoofing attacks. A large number of databases are shared in 5G technology and hence we face challenges regarding security and privacy. Here, we discuss some authentication approaches based on machine learning techniques for the 5G wireless network.

Keywords*: 5G wireless network, Artificial Intelligence, network slicing, Internet of Things, Machine Learning, data privacy, security management*

I. INTRODUCTION

Wireless communication and networking technologies are emerging as a promising vision for 5G wireless networks through realizing industrial and factory automation. Wireless networks are capable of interconnecting a large number of smart devices with the intelligent and reconfigure ability, and smartly make decisions by itself with the help of advanced technologies. Hence, it has wide applications in various domains, such as smart home/city/grid, e-health, intelligent transportation, automatic industry, meter auto reporting, remote sensing, and so on, which greatly improve the quality of our lives.

Although there are many advantages, facilitating and implementing intelligent wireless networks gives rise to several key challenges. First, considering the large amount of data generated by the huge number of smart devices, the applications of intelligent wireless networks face the challenges of collecting, accessing and processing the massive amount of data, to exploit and analyze the big amount of data toward the behaviour and characteristics discovery of wireless networks. Moreover, due to the extreme range of service requirements of wireless smart devices and the complex/dynamic environments, its applications are still not smart enough to tackle optimized physical layer designs, sophisticated learning, complicated decision making and efficient resource management tasks.

Even though cryptographic techniques have been widely studied for authentication, they may fall short of the desired performance in many emerging scenarios of 5G-and-beyond wireless networks. The weaknesses of conventional cryptography techniques are the increasing latencies, communication and computation overheads for achieving better security performance. Appropriate key management procedures are necessary for the conventional cryptographic techniques, and cooperation among multiple entities is required. This leads to excessive latencies and significant communication overhead and detecting compromised security keys cannot be readily achieved by the conventional digital credentials-based techniques.

Next-generation wireless networks must support ultra-reliable, low-latency communication and intelligently manage a massive number of Internet of Things (IoT) devices in real-time, within a highly dynamic environment. This need for communication quality-of-service (QoS) requirements as well as mobile edge and core intelligence can only be realized by integrating fundamental notions of artificial intelligence (AI) and machine learning across the wireless infrastructure and end-user devices.

II. MACHINE LEARNING

Machine learning was born from pattern recognition and it is essentially based on the premise that machines should be endowed with artificial intelligence that enables them to learn from previous computations and adapt to their environment through experience without being explicitly programmed. It uses statistical techniques to analyze observations/data/ experience by finding the patterns and underlying structures, in order to give devices the ability to "learn" automatically without human intervention and adjust actions accordingly. It is widely applied in computer vision, signal/language processing, social behaviour analysis, projection management, medical diagnosis, search engines and speech recognition.

III. ML TECHNIQUES

Four ML learning approaches are:

a)Supervised learning, b) Unsupervised learning, c) Semi-supervised learning, and d) Reinforcement learning.

a) Supervised learning algorithms are trained using labelled data, where both the input data and its desired output data are known to the system. It is commonly used in applications that have enough historical data.

b) Training of unsupervised learning tasks is done without labelled data. The goal of **unsupervised learning** is to explore the data and infer some structure directly from the unlabeled data.

c) Semi-supervised learning is used for the same applications as supervised learning but it uses both labelled and unlabeled data for training. This type of learning can be used with methods such as classification, regression and prediction. Semi-supervised learning is useful when the cost of a fully labelled training process is relatively high.

d) Reinforcement learning (RL), in contrast to the previously discussed learning methods that need to be trained with historical data, is trained by the data from implementation. The

goal of RL is to learn an environment and find the best strategies for a given agent, in different environments. RL algorithms are used for robotics, gaming, and navigation.

To perform these learning tasks, several frameworks have been developed. Among these, artificial neural networks (ANNs) constitute one of the most important pillars of machine learning, as they are able to mimic human intelligence, to model complex relationships between inputs and outputs, to find patterns in data, or to extract the statistical structure in an unknown joint probability distribution from the observed data.

IV. ADVANTAGES OF MACHINE LEARNING

- Machine learning has the ability to learn useful information from input data, which can help improve network performance.
- Machine learning based resource management, networking and mobility management algorithms can well adapt to the dynamic environment.
- Machine learning helps to realize the goal of network self-organization. For example, using multi-agent reinforcement learning, each node in the network can self optimize its transmission power, sub channel allocation and soon.
- By involving transfer learning, machine learning has the ability to quickly solve a new problem. It is possible to transfer the knowledge acquired in one task to another relevant task, which can speed up the learning process for the new task.

V. AN OVERVIEW OF 5G TECHNOLOGIES

5G is the fifth generation of cellular network technologies specified by the 3rd Generation Partnership Project (3GPP).It proceeds 2G, 3G, and 4G and their associated technologies, while introducing significant performance improvements.

a) Millimetre Wave Spectrum.

In addition to the classical spectrum below 6 GHz used by the majority of wireless communication technologies, 5G will operate in a high-frequency spectrum, from 28 GHz up to 95 GHz .This range is known as the millimeter wave (mmWave) spectrum. Compared to previous cellular network technologies, 5G will use a larger band of frequencies, thus, avoiding congestion. In comparison, 4G operates typically in the range 700-2600 MHz.

b) Massive MIMO and Beamforming.

5G uses the massive multiple-input and multiple-output (MIMO) technology. This technology consists of large antenna formations in both the base station and the device to create multiple paths for data transmission. With MIMO, 5G can achieve high spectral efficiency and better energy efficiency. Beamforming is a subset of massive MIMO. Beamforming controls the direction of a wave-front by manipulating the phase and magnitude of the signals sent by a single antenna placed in a formation of multiple antennas. Beamforming identifies the most efficient path to deliver the data to a receiver, while reducing the interference with nearby terminals. In addition, 5G uses a full-duplex technology which doubles the capacity of wireless links at the physical layer. With full-duplex, a device is able to transmit and receive data at the same time, using the same frequency.

c) Small Cells.

In addition to a larger spectrum and massive MIMO, 5G will comprise densely distributed networks of base stations in small cell infrastructure. This enables enhanced mobile broadband (eMBB) and low latency, providing an ideal infrastructure for edge computing. While small cells are typically used to cover hot spots, in mmWave 5G, they become a necessity due to the high-frequency (above 28 GHz) radio waves that cannot cover the same area as the classical low frequencies (below 6 GHz).

d) Device-to-Device Communication.

Similar to Bluetooth and WiFi (i.e. WiFi-Direct), 5G allows devices to communicate with each other directly, with minimal help from the infrastructure. This device-to-device (D2D) communication is a key feature of 5G that has the potential to accelerate the development of edge-centric applications. For example, in automotive applications, vehicles will be able to talk directly to each other, thus, reducing latency and avoiding the failure of the connection to the base station. Other use cases of D2D 5G communication are federated learning, where edge devices could share data among them, and blockchain where devices need to establish peer-to-peer (P2P) connections.

e) Virtualization.

5G networks are going to be highly virtualized. The software-defined networking (SDN), network function virtualization (NFV), and network slicing are given importance. SDN is an approach that separates networking data plane (i.e., data forwarding process) from the control plane (i.e., the routing process). This separation leads to easier configuration and management, and higher flexibility and elasticity. Complementary to SDN, NFV uses commodity hardware systems to run networking services that are traditionally implemented in hardware, such as routers and firewalls. With NFV, network flexibility is greatly improved, and the time-to-market is reduced, at the cost of lower efficiency.

f) Network Slicing.

Based on SDN and NFV, 5G networks will employ network slicing to multiplex virtualized end-to end networks on top of a single physical infrastructure. By separating infrastructure operators and service providers, 5G will better utilize hardware resources can provide services to both businesses and end-users.

g) Performance Improvements.

Compared to 3G and 4G, 5G has a lower latency of approximately 1 ms, increased energy efficiency, and a peak throughput of 10-20 Gbps. The increase in bandwidth gives better user experience, and allows more connected devices, such as drones, vehicles, and AR goggles, among others. While a 4G base station can only support around 100,000 devices, 5G can support up to a million devices per square km. A 5G network is designed to be flexible and suited for edge deployment, which further improves the end-to-end latency and overall user experience.

VI. SECURITY AND PRIVACY CHALLENGES

One distinguishing feature of 5G is network slicing, which enables applications with distinct requirements to share the same network. While virtualization has obvious advantages in terms of better exploiting the physical infrastructure and reducing the time to market, it poses security and privacy challenges.

The other features of 5G, such as improved bandwidth and latency, higher device density and D2D communication

may impact the security as well. Higher device density and increased bandwidth make it easier to conduct large scale DDoS attacks, especially using IoT devices. The fact that one 5G network slice comprises multiple virtualized resources managed by multiple providers makes it difficult to ensure isolation. It is possible to achieve virtual machine isolation with the secure design of hardware virtualization, but in slice isolation, the layers need to coordinate and agree on a cross-layer protocol. At each layer, the slice needs to be isolated to ensure security and privacy.

Applications based on device location, may see new devices moving in and out of range at high velocity. This type of ad-hoc communication with a high churn rate poses a new challenge for device authentication. Devices must establish identities of each other before communicating, by knowing the mapping of devices to their public keys. Existing public key infrastructures (PKIs) are too heavy weight because they are designed for enterprise identities. Large-scale consumer systems such as those used for end-to-end encryption, for example, iMessage and WhatsApp, meet the performance and scalability requirements, but they still rely on a centralized party. To decentralize the existing identity systems, we envision a blockchain-based solution which maintains a highly available and tamper-evident ledger storing identity information. However, existing blockchains are severely limited in their throughput and latency. Therefore, novel blockchain systems are needed to meet the performance requirement of future 5G applications. Current practices in enterprise security rely on collecting and analyzing data both at endpoints and within the network to detect and isolate attacks. 5G brings more endpoints and vastly faster networks. More endpoints mean a larger attack surface, raising the probability of the network being attacked to near certainty. Faster networks impact data collection, as it becomes unfeasible to store, and later analyze, highly granular data over long periods of time. Therefore, 5G demands a fundamentally new security analytics platform.

Apart from the security aspects, 5G also presents new challenges and opportunities in terms of privacy, as the improved bandwidth and reduced latency of 5G open up the possibility of transforming mobile devices into private databases that could be queried in real-time. For example, consider an online shopping service that provides recommendations to users based on their shopping histories. With current technologies, performing such recommendations requires the service provider to store users' shopping histories at the server-side, which has implications for privacy. In contrast, with the help of 5G, we may keep each user's shopping history in her local device, and let the service provider join hands with the users to perform recommendations in a privacy-preserving manner, e.g., by offloading to the users the part of the recommendation task that requires access to private data.

VII. AUTHENTICATION WITH INTELLIGENCE

We envision intelligent authentication approaches with the help of machine learning to address the above challenges for security enhancement and more efficient management in 5G-and-beyond networks. Intelligent authentication is required to meet multi objectives, namely high cost efficiency, high reliability, model independence, continuous protection, and situation awareness. The advantages of intelligent authentication based on machine learning is as follows:

High cost efficiency:

Due to the increasing number of resource-constraint devices in 5G, both communication and security management should be executed concurrently to achieve cost-effective authentication. The opportunistic selection of attributes and dimension reduction methods may help in decreasing the complexity of the authentication system relying on multi-dimensional attributes as well as in reducing communication and computation overheads. Furthermore, by performing training and testing with devices having enough energy and storage space, simplified and fast authentication can be achieved at resource-constraint devices.

High reliability:

Utilizing specific features and relationships in the multi-dimensional domain is extremely helpful for achieving security enhancement, since it is more difficult for an adversary to succeed in predicting or imitating all attributes based on the received signals and observations. The multi-dimensional information, such as time, frequency, network architecture, and communication process, provides broader protections for legitimate users. ML techniques could facilitate the authentication by analyzing and fusing the multi-dimensional information.

Model independence:

A data-based scheme through exploring machine learning techniques overcomes the difficulties in modeling uncertainties and unknown dynamics of the authentication process. Hence, the model-free authentication scheme removes the assumption of knowing structures of authentication systems, resulting in a more scalable authentication process design.

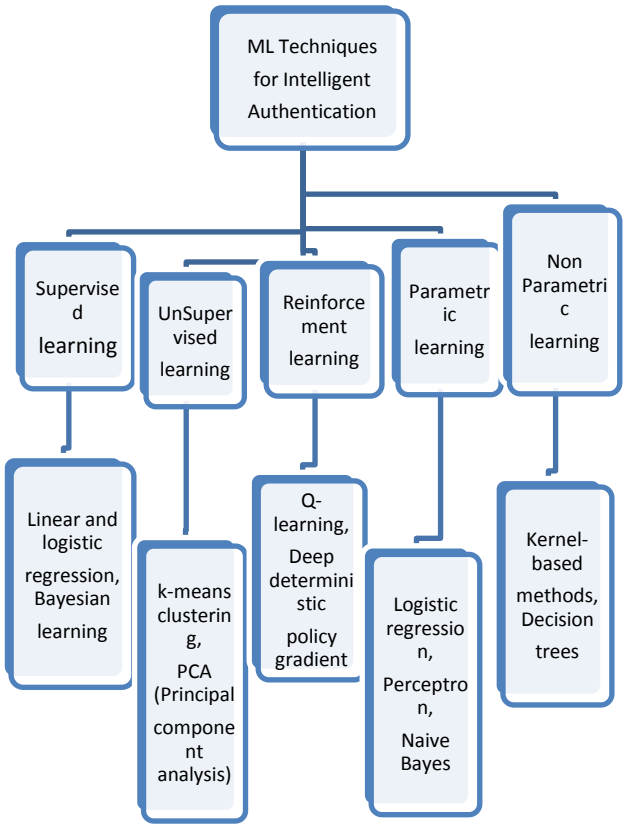
Continuous protection:

Utilizing the received information along with data transmission for authentication could provide identification after login and control the varying security risks. In achieving this, machine learning techniques may be explored for data analysis and processing, so that the seamless protection for legitimate devices can be achieved in 5G-and-beyond networks. Furthermore, continuous authentication may contribute to the quick authentication handover.

Situation awareness:

The situation-aware authentication observes the varying security risks and learns from the complex dynamic environment to enhance security. ML techniques provide powerful tools to learn the dynamic adversarial environment for self-optimization and self-organization, thereafter for achieving the automatic authentication and help in the detection of time-varying attributes and the adaptation of authentication process.

IX. ML TECHNIQUES FOR INTELLIGENT AUTHENTICATION



CONCLUSION

With 5G on the verge of being adopted as the next mobile network, it is necessary to analyze its impact on the landscape of computing and data management. The broad impact of 5G on both traditional and emerging technologies and the machine learning paradigms for intelligent authentication and its advantages were provided. The intelligent authentication design was developed to enhance security performance in 5G-and-beyond wireless networks. As a conclusion, machine learning techniques provide a new insight into authentication under unknown network conditions and unpredictable dynamics, and bring intelligence to the security management to achieve cost-effective, more reliable, model-free, continuous, and situation-aware authentication. Machine learning techniques in 5G will make the world even more densely and closely-connected, and will present us with vast amounts of possibilities and opportunities to overcome the challenges ahead of us.

References

[1] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. C. Chen, and L. Hanzo, "Machine Learning Paradigms for Next-Generation Wireless Networks," IEEE Wireless Commun. Mag., vol. 24, no. 2, 2018, pp. 98-105.

[2] He Fang, Xianbin Wang, Stefano Tomasin, " Machine Learning for Intelligent Authentication in 5G-and-Beyond Wireless Networks" 28 july 2019, IEEE Wireless Communications, 2019 - ieeexplore.ieee.org

[3] Mingzhe Chen, Ursula Challita, Walid Saad, Changchuan Yin, "Artificial Neural Networks-Based Machine Learning for Wireless Networks: A Tutorial" October 2017, IEEE Communications Surveys & Tutorials

[4] H Yang, X Xie, M KadochH Yang, X Xie, M Kadoch "Machine learning techniques and a case study for intelligent wireless networks" - IEEE Network, 2020 - ieeexplore.ieee.org

[5] Dumitrel Loghin, Shaofeng Cai , Gang Chen, Tien Tuan Anh Dinh, Feiyi Fan, Qian Lin , Janice Ng , Beng Chin Ooi , Xutao Sun, Quang-Trung Ta, Wei Wang, Xiaokui Xiao, Yang Yang, Meihui Zhang , Zhonghua Zhang "The Disruptions of 5G on Data-Driven Technologies and Applications" IEEE Transactions On Knowledge And Data Engineering, Vol. 32, No. 6, June 2020 Pgs 1179 -1198.

[6] H. Fang, X. Wang, and L. Hanzo, "Learning-aided Physical Layer Authentication as an Intelligent Process," IEEE Trans.Commun., vol. 67, no. 3, 2019, pp. 2260-2273.