

Machine Learning in IoT Security: Current Issues and Future Prospects

Gabriel Tosin Ayodele
Faculty of Engineering and
Informatics/ University of Bradford
Bradford West Yorkshire UK

Abstract:

The Internet of Things (IoT) connects billions of intelligent devices that can communicate with each other without human intervention. With an estimated 50 billion devices by the end of 2020, it is one of the fastest growing areas of computing history. On the other hand, IoT technology is essential to the advancement of various real-world intelligent applications that can improve people's quality of life. However, the interdisciplinary components involved in networking and deploying IoT systems raise new security concerns. Encryption, authentication, access control, network security, and application security solutions are worthless when it comes to IoT devices and their inherent shortcomings. Therefore, existing security measures need to be updated to properly protect the IoT environment. Machine learning and deep learning (ML / DL) have evolved dramatically in recent years, and machine intelligence has evolved from laboratory curiosity to viable machines. An important defense against new or zero-day attacks is the ability to intelligently monitor IoT devices. ML / DL is a powerful data exploration technique for revealing "normal" and "bad" behavior in the context of IoT components and devices. Therefore, machine learning and deep learning technologies are important not only to enable secure device communication, but also to transform IoT security into a security-based intelligence system. The purpose of this study is to provide a comprehensive overview of recent advances in machine learning and deep learning

techniques that can be used to improve IoT security solutions. It describes the many possible attack surfaces of IoT systems, the potential risks associated with each surface, and lists unique or newly established IoT security threats. It then details the ML / DL approach to IoT security, highlighting the potential, strengths, and weaknesses of each method. Learn about the possibilities and challenges of using machine learning and deep learning for IoT security. These opportunities and challenges may be used as future research paths.

Keywords—Deep Learning, Machine Learning, Internet of Things Security, Security based Intelligence, IoT Big Data, Attacks, Privacy

I. Introduction

The Internet of Things (IoT) is defined as a network of interconnected and dispersed embedded systems that interact via wired or wireless communication methods (Novo et al. , 2015). It is also described as an environment network of physical objects, having comparatively little computational capacity for data acquisition, processing, memory, and communication, incorporating electronics (including sensors and actuators), software programs, and networking. IoT items are smart appliances like smart light bulbs, smart adapters, smart meters, smart refrigerators, smart ovens, air conditioners, temperature sensors, smoke detectors, IP cameras, etc Clients that have updated items such as smartphones will also comes under this category, Examples of modern smart appliances include. This global scale bring new difficult situations to

control the devices, the volume of generated data, storage and communications, and, of course, processing and security and privacy. There is great literature available in the following domains of the IoT: architecture, communications, protocols, applications, security, and privacy, among others (Alfuqaha et al., 2015) (Granjal, Monteiro, Silva, 2015). software, and network connections. Smart appliances such as smart light bulbs, smart adapters, smart meters, smart refrigerators, smart ovens, air conditioners, temperature sensors, smoke detectors, and IP cameras are examples of IoT items, as are more modern devices such as smartphones. The large scale of IoT networks creates new challenging conditions in terms of device control, sheer volume of data, storage, communication, processing, and protection and privacy. Extensive research has been conducted on several areas of the IoT (architecture, communications, protocols, applications, security, privacy, etc.) (Alfuqaha et al. , 2015) (Granjal, Monteiro, Silva, 2015). This is because the amount of data produced by IoT devices is extraordinarily large, and conventional data acquisition, storage, and analysis methods may not operate at this level. In addition, the proliferation of data by IoT paves a new channel that traditional data processing systems cannot exhaust. In this respect, Machine Learning (ML) is considered one of the most suitable paradigms of computing for offering local intelligence at IoT devices (Saeid, Rezvan and Barekatin, 2018). It can also be defined as an ability of a smart device to modify or manage a state or perform an action as a result of an existing or gained awareness, which is also believed to be an essential part of an IoT system. From Figure 1 it is observed that the credence of monitoring IoT devices can provide an intelligent level answer to new or zero raw attacks. Secondly, due to the ability to learn instances of unknown new assaults and forecast them, ML/DL algorithms may help anticipate new attacks as they are mutations of previous

ones. For successful and safe systems, the IoT systems need to move from just enabling secure communication of the devices to security intelligent supported.

Figure 1

Illustration of the potential role of ML/DL in IoT security

The key research question of this survey is listed as follows:

- In earlier studies, what strategies were employed to solve IoT security issues using Machine Learning?
- What are the different IOT security treats?
- How might the arrangement of those featured treat be handled utilizing machine learning?

II. Iot Security Threats

IoT devices, on the whole, function in a variety of settings to achieve a variety of goals. As a result, balancing security requirements with the IoT framework's large attack surface is difficult. As a result, an unauthorized person might get access to these devices. Because of their limited computation and power capabilities, IoT devices are unable to maintain complicated security frameworks (Abomhara and Kien, 2015). Obtaining the IoT framework in this manner is a confusing and challenging effort.

A. Threats in IoT

There are two types of security threats: cyber and physical. Cyber risks are further divided into passive and active categories. The dangers are briefly discussed in the next sections.

1) Cyber Threats

Passive threats: Eavesdropping through communication channels or the network is a passive danger. An attacker can use eavesdropping to obtain data from sensors, follow sensor owners, or do both. On the black market, the collection of valuable personal information, particularly personal

health data, has become commonplace (Restuccia, Oro and Melodia, 2018).

Active threats: In active threats, the attacker is capable of not only listening in on channels of communication but actively manipulating IoT systems to change settings, regulate communication, refuse services, and so on. For illustration, active attacks on an IoT system might include impersonation (e.g., faking, Sybil, and man-in-the-middle), fraudulent inputs, data manipulation, and denial of service.

11) Physical Threats

Regardless of how distinct improvements are used at the real layer for IoT, the idea of actual follows commonly seem like and require social creating approaches. Furthermore, in order to launch genuine attacks, aggressors must be in close proximity to the gadgets/equipment, with different objectives in mind, such as completely destroying the equipment, limiting its lifespan, jeopardizing the communication component, changing the energy source, and so on. Furthermore, such changing devices may allow adversaries to make modifications to the directing tables and security keys, affecting communication with upper levels (Nawaz and Loscri, 2015). (Yang et al. , 2011).

B. Security Challenges in IoT Deployment

Security and protection have become two important factors affecting the commercial adoption of IoT administrations and applications. The current Internet is a security attack ground that spans from basic hacking, through hacks to all around business constructed hacks that have had a negative effect on several corporations such as the medical and commerce services. There are additional factors concerning IoT application and device security, stemming from the device limitations as well as from the context in which IoT runs. So far, security and protection issues in the IoT context have been examined comprehensively from numerous

viewpoints, including correspondent security, data security, protection, technical security, character the executives, malware research, and so forth (Ray et al., 2016).C. Holes in the Current Security Solution for IoT Networkse commercial acceptance of IoT administrations and applications. The current Internet is a magnet for security attacks ranging from simple hacks to corporate-level all-around constructed security breaches that have negatively impacted a variety of businesses, including medical services and commerce. The limitations of IoT devices, as well as the environment in which they operate, provide additional challenges for the security of both applications and devices. Until now, security and protection concerns in the IoT area have been widely investigated from many perspectives, such as correspondent security, data security, protection, technical security, character the executives, malware research, and so on (Ray et al., 2016). C. Gaps in the Existing Security Solution for IoT Networks Fundamental issues of security and privacy need to be addressed to facilitate the strategic deployment of IoT hence the need to look into its drivers. In particular, the word IoT has been thrown out of previous technologies, hence it becomes apparent if the security threats in IoT are new or inherited from previous technology. According to (Fernandes and Eykholt, 2017), authors studied and discussed the comparative analysis of threats that are relevant for IoT and traditional IT gadgets. They also focused mostly on the issues affecting privacy. Software, hardware, network, and applications are the key driving forces in arguing referential similarities and differences. These categorization show that security threat in the conventional IT and IoT are essentially similar. But the IoT's primary concern is resource constraints, and the problem of applying advanced security solutions on IoT networks becomes challenging.

III. To provide some context, they split out two particular areas of focus:

IOT Security and Machine Learning In this section, we briefly describe main machine learning algorithms and how they may be used in IoT systems. A. Basic of Machine Learning Algorithmstical components in the commercial acceptance of IoT administrations and applications. The current Internet is a magnet for security attacks ranging from simple hacks to corporate-level all-around constructed security breaches that have negatively impacted a variety of businesses, including medical services and commerce. The limitations of IoT devices, as well as the environment in which they operate, provide additional challenges for the security of both applications and devices. Until now, security and protection concerns in the IoT area have been widely investigated from many perspectives, such as correspondent security, data security, protection, technical security, character the executives, malware research, and so on (Ray et al., 2016).

C. Gaps in the Existing Security Solution for IoT Networks. It is critical to investigate the fundamental causes of security and privacy concerns in order to successfully use IoT. More specifically, the word IoT has been hurled out of previous technologies, therefore it's critical to identify whether the security concerns in IoT are new or a rehash of prior technologies' legacy. (Fernandes and Eykholt, 2017) compared and contrasted the security concerns faced by IoT and conventional IT devices. They also concentrated on concerns about privacy. Software, hardware, network, and applications are the key driving forces in arguing about similarities and differences. The security challenges in the conventional IT sector and the IoT are fundamentally comparable, according to these categorization. However, the IoT's main worry is resource limits, which make it difficult to adapt existing advanced security solutions to IoT networks.

iii. Iot Security and Machine Learning.

In this section, we discuss various machine learning algorithms and their applicability in IoT applications. A. Basic Machine Learning Algorithms The ML algorithms can be categorized as supervised, unsupervised, semi-supervised and Reinforcement learning algorithms.

Supervised Learning: Supervised learning is utilized when specified objectives are stated to be achieved from a given input. The data is labeled initially in this type of learning.

Unsupervised Learning: The environment just provides inputs for unsupervised learning, with no specific goals in mind. It may investigate similarities among unlabeled data and organize it into various categories without requiring labeled data.

Semi-supervised Learning: In the preceding two kinds, either all of the observations in the dataset have no labels or all of the observations have labels. Semi-supervised learning is somewhere in the middle.

Reinforcement Learning: No explicit goals are established in Reinforcement Learning (RL), and the operator learns via feedback after observing the environment. It does some acts and made a decision based on the reward it receives.

D. Machine Learning Techniques Used in IoT Security

As indicated in Table I, we will explore several machine learning techniques concentrating on the underpinning security and privacy issues in IoT networks. Authorization, threat detection and mitigation, Dispersed Denial of Service (DDoS) assaults, anomaly and detection techniques, and malware analysis are all things we examine.

TABLE I Machine Learning Techniques Used In Iot Security Machine Learning Algorithm Description Naivebayes

It's a classification technique that can be applied in both binary and multi-class

environments. Instead of computing the actual values, all of the qualities are considered to be conditionally independent (Zhou et al., 2017).

K-nearest neighbour It is a simple and effective supervised learning model that is used to associate new data points with existing comparable data points by exploring the given dataset. (Bekara, 2020).

K-means algorithm The K-means clustering method, which belongs to the unsupervised category of the ML family, is the most widely utilized well-known approach. K-Means clustering is a technique for classifying or grouping devices into K groups based on features or parameters (Kolias and Kambourakis, 2017).

Random forest and decision tree(dt)

It is a form of learning that is supervised. It creates a model by putting particular rules into action based on data attributes. This model is then used to forecast the value of a new targeted variable. In classification and regression issues, decision trees are utilized. These trees are essentially used to divide a dataset into numerous branches depending on particular principles (Roman, Zhou and Lopez, 2013).

Support Vector Machines (Svm) Svm is a low-complexity supervised machine learning technique for classification and regression (Bekara, 2020).

Deep learning It's a feed forward Neural Network (NN) in which each neuron is linked to another layer and there are no connections between the layers. Deep learning describes numerous layers that retain various levels of perception, with each layer receiving input from the previous layer and feeding the result to the next one (Azmoodeh, Dehghantanha and Member, 2019).

E. Limitations in Applying Machine Learning in IoT Networks

It covers the following typical features: huge and diverse amount of IoT traffic as well as its high and random speed. Unfortunately,

most common machine learning techniques on their own are not efficient or scalable enough to deal with IoT data which require drastic modifications (Qiu et al., 2016). In addition, IoT data is inherently imprecise by nature, and it is impossible to mitigate this stochasticity. The following sections discuss some of the most obvious evils of using machine learning algorithm in IoT networks.

Processing power and energy:

The Internet of Things devices are generally small, and their computing abilities are limited due to energy constraints. Smart IoT devices require real time data processing for real time applications, while the normal machine learning solutions are not designed to handle continuous streams of data in real time. Indeed, this frequency resulted in a number of issues, which arise when using traditional procedures designed for a vast amount of data. Lastly, with increasing data complexities, the amount of forecast an algorithm can do also decreases (Heureux & Member, 2017).

Data management and analytics:

Wireless data can originate from other networked in formation systems, sensing and communication devices and so on among them (Bogale, Wang and Le, 2018). Data generated in IoT networks are of syntax and semantic diverse form, format, and semantics leading to syntactic and semantic heterogeneity. Syntactic variation is evident in data types, file formats, encoding systems and data models. In the case of huge volumes of data and different types of data sets with different characteristics, such heterogeneity becomes a problem in terms of generalizable consistency.

iii. Machine learning: a solution to iot security challenges

From then, machine learning is essentially smart ways of optimizing execution rules with the help of model's insights or previous occurrences. In addition, it is used practical applications including Google having

devised machine learning to analyze threats to its android point and apps. However, in the case of DL, a type of ML, the model itself can come up with a way to decide about the precision of an expectation on its own. another perspective of DL models is presented as more suitable for arrangement and expectation undertakings in progressive IoT applications due to their Self-organizing nature with logical and individual assistance. While traditional methodology is found prevalent in different strands of IoTs for example applications, administrations, designs, conventions, data gathering, asset allocation, categorization, and analysis including security the valley associated with IoT demand very shrewd, reliable and consistent approaches. To this aim, ML and DL are interesting approaches for IoT networks for several reasons, first of all, because IoT networks produce a large volume of data. In addition, the ML and DL methodologies allow IoT frameworks to make intelligent and wise decisions increasing the value of data created by the IoT. DL methods may also be used in IoT devices to perform complex detect and acknowledge operations, making it possible to admit new applications and/or administrations that take into consideration constant working association with humans, smart devices and environment literals. The following are a few actual security-related uses of machine learning:

- Face recognition for criminology: appearance, lighting, obstacle (glasses, facial hair), make-up, hairstyle, and so on.
- Different penmanship styles for character recognition for security encryption.
- Proof of malevolent code: identifying noxious code in applications and programs
- Distributed Denial of Service (DDoS) detection: DDoS assaults on infrastructure can be detected via behavior analysis..

B. Adding ML and DL algorithms to IoT applications raises new concerns. These issues are complicated. For example, it is trying to cultivate an acceptable pattern of responding to data derived from different IoT applications. Furthermore, identifying information information is also a process which consumes much time. One other test is related to the least highlighted aspects of educational experience. Other concerns emerge when these models are organized on asset-required IoT devices where minimizing handling and accumulation is even more important higher (Bekara, 2020). Further the variations arising out of ML or DL computations are too excessive to be handled by the basic architectural schema and repetitively implemented. In this regard, it is crucial to elaborate the IoT security measures concerning the preceding, with respect to the impact on the ML and DL.

Conclusion:

The buying specifications of IoT devices add layer of complexity to acquire and integrate the different innovations that are needed from physical, transport distant transmission, flexibility design, to cloud. In recent years machine learning and deep learning have ushered in a number of strong scientific process that can be applied to enhance IoT security. The following IoT security threats and attack surfaces are described in this report. It offers a detail analysis of the envisioned uses of machine learning and deep learning in the protection of IoT.

References

- Abomhara, M. and Kjøien, G. M. (2015) 'Cyber Security and the Internet of Things : Vulnerabilities , Threats , Intruders', 4, pp. 65–88. doi: 10.13052/jcsm2245-1439.414.
- Ahmed, N., De, D. and Hussain, I. (2018) 'Internet of Things (IoT) for Smart Precision Agriculture and Farming in Rural Areas', IEEE Internet of Things Journal, 5(6), pp. 4890–4899. doi: 10.1109/JIOT.2018.2879579.

- Al-fuqaha, A. et al. (2015) 'Internet of Things : A Survey on Enabling', 17(4), pp. 2347–2376.
- Altawy, R., Youssef, A. M. R. M. and Member, S. (2016) 'Security Tradeoffs in Cyber Physical Systems : A Case Study Survey on Implantable Medical Devices', IEEE Access. IEEE, 4, pp. 959–979. doi: 10.1109/ACCESS.2016.2521727.
- Ammar, M., Russello, G. and Crispo, B. (2018) 'Journal of Information Security and Applications Internet of Things : A survey on the security of IoT frameworks', Journal of Information Security and Applications. Elsevier Ltd, 38, pp. 8–27. doi: 10.1016/j.jisa.2017.11.002.
- An, N. et al. (2017) 'Behavioral Anomaly Detection of Malware on Home Routers'.
- Arulkumaran, K. et al. (2017) 'A Brief Survey of Deep Reinforcement Learning', pp. 1–16.
- Azmoodeh, A., Dehghantanha, A. and Member, S. (2019) 'Robust Malware Detection for Internet Of (Battlefield) Things Devices Using Deep Eigenspace Learning'.
- Banerjee, B. A. et al. (2011) 'Ensuring Safety , Security , and Cyber – Physical Systems'.
- Bekara, C. (2020) 'Security Issues and Challenges for the IoT-based Smart Grid', Procedia - Procedia Computer Science. Elsevier Masson SAS, 34, pp. 532–537. doi: 10.1016/j.procs.2014.07.064.
- Bertino, E. (2020) 'Botnets and Internet'.
- Bogale, T. E., Wang, X. and Le, L. B. (2020) 'Machine Intelligence Techniques for Next-Generation Context-Aware Wireless Networks'.
- Bogaz, B. et al. (2017) 'Journal of Network and Computer Applications A survey of intrusion detection in Internet of Things', Journal of Network and Computer Applications. Elsevier, (January), pp. 0–1. doi: 10.1016/j.jnca.2017.02.009.
- Chauhan, J. et al. (2018) 'Institutional Knowledge at Singapore Management University Breathing-based authentication on resource-constrained IoT devices using recurrent neural networks', pp. 60–67.
- Du, H. et al. (2020) 'The Elements of End-to-end Deep Face Recognition: A Survey of Recent Advances'. Available at: <http://arxiv.org/abs/2009.13290>.
- Fernandes, E. and Eykholt, K. (2018) 'Internet of Things Security Research ', pp. 1–5.
- Fosso, S., Anand, A. and Carter, L. (2013) 'International Journal of Information Management A literature review of RFID-enabled healthcare applications and issues', International Journal of Information Management. Elsevier Ltd, 33(5), pp. 875–891. doi: 10.1016/j.ijinfomgt.2013.07.005.
- From, C. (2009) 'Securing Wireless Implantable Devices for Healthcare : Ideas and Challenges', (July), pp. 74–80.
- Granjal, J., Monteiro, E. and Silva, J. S. (2015) 'Pr E oo f Pr E oo f', pp. 1–20.
- Heureux, A. L. and Member, G. S. (2017) 'Machine Learning With Big Data : Challenges and Approaches', IEEE Access. IEEE, 5, pp. 7776–7797. doi: 10.1109/ACCESS.2017.2696365.
- Imran, M et al. (2017) 'The rise of ransomware and emerging security challenges in the Internet of Things', Computer Networks. Elsevier B.V. doi: 10.1016/j.comnet.2017.09.003.
- Items, R. et al. (2018) 'This is a repository copy of A Deep Recurrent Neural Network Based Approach for Internet of Things Malware Threat Hunting . White Rose Research Online URL for this paper : Version : Accepted Version Article : HaddadPajouh , H ., Dehghantanha , A orcid . org / 0000-0002-9294-7554 , Khayami , R . et al . (1 more author) (2018) A Deep Recurrent Neural Network Based Approach for Internet of Things Malware Threat Hunting . Future Generation Computer Systems , 85 . pp. 88-96. ISSN 0167-739X'.
- Kolias, C. and Kambourakis, G. (2017) 'DDoS in the IoT ', (January). doi: 10.1109/MC.2017.201.

- Lake, B. M. et al. (2012) 'Building Machines That Learn and Think Like People', (2012), pp. 1–58.
- Lane, N. D., Georgiev, P. and Qendro, L. (2015) 'DeepEar : Robust Smartphone Audio Sensing in Unconstrained Acoustic Environments using Deep Learning'.
- Mahmud, M. et al. (2018) 'Applications of Deep Learning and Reinforcement Learning to Biological Data arXiv : 1711 . 03985v2 [cs . LG] 7 Jan 2018', pp. 1–33. doi: 10.1109/TNNLS.2018.2790388.c.
- Mishra, P. et al. (2018) 'A Detailed Investigation and Analysis of using Machine Learning Techniques for Intrusion Detection', IEEE Communications Surveys & Tutorials. IEEE, PP(c), p. 1. doi: 10.1109/COMST.2018.2847722.
- Mohammadi, M. et al. (2017) 'Semi-supervised Deep Reinforcement Learning in Support of IoT and Smart City Services', X(X), pp. 1–11.
- Nawaz, A. H. and Loscri, V. (2014) 'The Internet of Things-A survey of topics and trends'. doi: 10.1007/s10796-014-9489-2.
- Nguyen, N. D. U. Y. et al. (2017) 'System Design Perspective for Human-Level Agents Using Deep Reinforcement Learning : A Survey', pp. 27091–27102.
- Novo, O. et al. (2015) 'Capillary Networks – Bridging the Cellular and IoT Worlds', (December). doi: 10.1109/WF-IoT.2015.7389117.
- Perera, C. et al. (2017) 'Context Aware Computing for The Internet of Things : A Survey', X(X), pp. 1–41.
- Qiu, J. et al. (2016) 'A survey of machine learning for big data processing', EURASIP Journal on Advances in Signal Processing. EURASIP Journal on Advances in Signal Processing. doi: 10.1186/s13634-016-0355-x.
- Ray, S. et al. (2016) 'The Changing Computing Paradigm with Internet of Things ', 2356(c), pp. 1–13. doi: 10.1109/MDAT.2016.2526612.
- Razzaque, M. A., Milojevic-jevric, M. and Palade, A. (2015) 'Middleware for Internet of Things : a Survey', 0(0), pp. 1–27.
- Restuccia, F., Oro, S. D. and Melodia, T. (2018) 'Securing the Internet of Things in the Age of Machine Learning and Software-defined Networking', 1(1), pp. 1–14.
- Roman, R., Zhou, J. and Lopez, J. (2013) 'On the features and challenges of security and privacy in distributed internet of things', COMPUTER NETWORKS. Elsevier B.V. doi: 10.1016/j.comnet.2012.12.018.
- Saeid, M., Rezvan, M. and Barekatian, M. (2018) 'Machine learning for internet of things data analysis : a survey', Digital Communications and Networks. Elsevier Ltd, 4(3), pp. 161–175. doi: 10.1016/j.dcan.2017.10.002.
- Sethi, P. and Sarangi, S. R. (2017) 'Internet of Things : Architectures , Protocols , and Applications', 2017.
- Wallace, B. Y. T. C. et al. (2016) 'ΒΙΟΕΚΧΥΛΙΣΗ ΟΞΕΙΔΩΜΕΝΩΝ ΜΕΤΑΛΛΕΥΜΑΤΩΝ ΝΙΚΕΛΙΟΥ ΜΕ ΤΗ ΧΡΗΣΗ ΕΤΕΡΟΤΡΟΦΩΝ ΜΙΚΡΟΟΡΓΑΝΙΣΜΩΝ No Title', Bulletin of the Seismological Society of America, 106(1), pp. 6465–6489. Available at: <http://www.bssaonline.org/content/95/6/2373%5Cnhttp://www.bssaonline.org/content/95/6/2373.short%0Ahttp://www.bssaonline.org/cgi/doi/10.1785/0120110286%0Ahttp://gji.oxfordjournals.org/cgi/doi/10.1093/gji/ggv142%0Ahttp://link.springer.com/10.1007/s00024-01>.
- Wang, T. et al. (2018) 'Deep Learning for Wireless Physical Layer : Opportunities and Challenges', pp. 1–14.
- Yang, Z. et al. (2011) 'Study and Application on the Architecture and Key Technologies for IOT', pp. 747–751.
- Zhang, D., Han, X. and Deng, C. (2018) 'Review on the Research and Practice of Deep Learning and Reinforcement Learning in Smart Grids', 4(3), pp. 362–370. doi: 10.17775/CSEEJPES.2018.00520.
- Zhou, J. et al. (2017) 'Security and Privacy for Cloud-Based IoT : Challenges , Countermeasures , and Future Directions', (January), pp. 26–33.