Henil Vedant – MS cybersecurity- 495670888

IDS for securing IoT Enviornment.

Abstract—**The aim of this paper is to integrate all the fundamentals and concepts of IOT learned during the class of 'CIS-600 IOT-APP Dev' and put them into practical use to understand and analyze various cyber-security integrated IOT applications and that is why various approaches to provide security using IOT and cybersecure applications come under the lens of Cyber-Physical system for the scope of this paper and subject. We break down these CPS and understand the underlying ML approach and its augmentation with IoT which is used to provide security. The explosive growth of the Internet of Things (IoT) applications has imposed a dramatic increase of network data and placed a high computation complexity across various connected devices. The IoT devices capture valuable information, which allows the industries or individual users to make critical live dependent decisions. Most of these IoT devices have resource constraints such as low CPU, limited memory, and low energy storage. Hence, these devices are vulnerable to cyber-attacks due to the lack of capacity to run existing general-purpose security software. It creates an inherent risk in IoT networks. The multi-access edge computing (MEC) platform has emerged to mitigate these constraints by relocating complex computing tasks from the IoT devices to the edge.**

Keywords— (Data-Poisoning attacks, Evasion attacks, Membership inference attacks, CPS-Cyber Physical Systems, GDPR- General Data Protection Regulations, NIST, Adversarial attacks, CAT Triangle, XSS/CSRF- attacks, IDS-IPS Systems.**).**

I. INTRODUCTION

The Internet of Things (IoT) describes the network of physical objects—"things"—that are embedded with sensors, software, and other technologies for the purpose of connecting and exchanging data with other devices and systems over the internet. These devices range from ordinary household objects to sophisticated industrial tools. With more than 7 billion connected IoT devices today, experts are expecting this number to grow to 10 billion by 2020 and 22 billion by 2025.[1] Over the past few years, IoT has become one of the most important technologies of the 21st century. Now that we can connect everyday objects—kitchen appliances, cars, thermostats, baby monitors—to the internet via embedded devices, seamless communication is possible between people, processes, and things. By means of low-cost computing, the cloud, big data, analytics, and mobile technologies, physical things can share and collect data with minimal human intervention. In this hyperconnected world, digital systems can record, monitor, and adjust each interaction between connected things. The physical world meets the digital world—and they cooperate. The Internet of Things (IoT) describes the network of physical objects—"things"—that are embedded with sensors, software, and other technologies for the purpose of connecting and exchanging data with other devices and systems over the internet. These devices range from ordinary household objects to sophisticated industrial tools.The internet of things, or IoT, is a system of interrelated computing devices, mechanical and digital machines, objects, animals or people that are provided with unique identifiers (UIDs) and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction.[1].By means of low-cost computing, the cloud, big data, analytics, and mobile technologies, physical things can share and collect data with minimal human intervention. In this hyperconnected world, digital systems can record, monitor, and adjust each interaction between connected things. The physical world meets the digital world—and they cooperate.[2].Some industries that integrate IoT solutions are: Healthcare, Retail, Businesses, Security, Automation etc.

The Internet can be described as the communication network that connects individuals to information while The Internet of Things (IoT) is an interconnected system of distinctively address able physical items with various degrees of processing, sensing, and actuation capabilities that share the capability to interoperate and communicate through the Internet as their joint platform.

II. PROBLEM STATEMENT

The IoT system has evolved with many use-cases. Each of the applications of the IoT system comes with its own adverse security issues. This section elaborates on the IoT system and the related security problems. We consider an Intrusion Detection System and how easily it has been spoofed based on new hybrid kind of attacks that cause the system to falsely spot malicious code and not prevent it from privilege escalation. A lot of the issues related to IoT security stem from the fact there is a lack of universal acceptance towards an architecture or manufacturing standards or physical handling of its sensors, actuators and data management. These open an array of possible options for the attacker to exploit. Combined with these IoT devices making use of state of art Machine Learning algorithms and deploying MLaaS platforms they have gone on to make significant improvements in the way data is handled and managed and processed. The inherent security flaws of ML is overlooked and that combined with IoT security flaws it becomes a dangerous attack vector for hackers to exploit. Physical handling of the security issues and users' ignorance due to their lack of awareness of security problems associated with the IoT systems are also critical problems. Moreover, there is no standard defined security architecture accepted by the IoT community. Different security architectures are adopted based on the use case, system requirement, the available technology, and the IoT network size. N-IDS are often domain specific and do not make use of more state of art defensive approaches in order to achieve their goal of safeguarding our systems. Unfortunately, traditional network security systems are not applicable in IoT systems due to the nature of the resource constraints. Furthermore, different network attacks have also emerged due to the rapid development and wide application of the IoT system applications. The amount of attacks will continue to rise as the IoT use cases expands. Being able to recognize and comprehend the intense rise in cyber-threats in the IoT system drastically decreases the risk of a network security and data breach. The goal is to identify through literature survey the current attack methodologies used to evade such Network Intrusion Detection Systems and after finding the

shortcomings propose a more new age, modern and more efficient IDS that makes use of state of art ML approaches like Deep Learning.

III. LITERATURE SURVEY

Recent papers and research articles reports the results of qualitative case study that correlates academic literature with five Industry 4.0 cyber trends, seven cyber risk frameworks and two cyber risk models[2]. While there is a strong interest in industry and academia to have a standard across all existing cyber risk frameworks, models and methodologies, an attempt to combine these approaches has not been done until present. Propose a new that is integrating old standards and governance into Industry 4.0 and offers a better understanding of a holistic economic impact assessment model for IoT cyber risk with new threats.There are various types of attacks that can be carried out and the sophistication of these types of attacks not only depend on the types of underlying technology IoT implements but also can be bundled with various other attacks to make the attacks more dangerous, as we shall see in the further sections of this paper.

Membership Inference attacks have been common and consistent throughout the cyber-security world and that is the case with IoT systems aswell. Membership inference attacks lets us learn if a particular sample was part of the training data or not, this in turn results in letting us know more about the model and how to spoof such a system by creating our payload. More research in relevant field has been done by Hossen Mustafa etc. all [14]. Data poisoning Attacks specifically are used to poison and create malicious payload that is deployed and is misclassified, in an email filtering system a spam mail that beats the system can be created using this method. Adversarial samples have long been known to spoof computer vision and computer graphic systems and these are the same that are deployed for facial recognition across all systems. DNS rebinding attacks have long been known to incur the system heavy losses and lead to data theft, information loss and system compromise. Besides this standalone malicious kits can be deployed which are connected to main server and cause various leaks and damage. Current laws, Regulations and data prevention laws that apply to IoT w.rt are all data regulation and prevention laws covered in "GDPR Compliance Verification in Internet of Things" by Masoud Barati etc.all. GDPR states that if an user effectively wants to delete his/her information from a given application or web service then he/she should be able to do so without just deletion on the front façade which means that the datapoints should be removed/deleted and re-trained model should again give the desired output and accuracy. This brings its own set of challenges for ML engineers to incorporate these changes in their existent and future architectures.

Detection Methods	Threats Identified	Placement Strategies
Network fingerprinting	Network Anomalies	Distributed locally in the IoT Devices
Artificial Neural Network (ANN)	Data Intrusion	Distributed locally in the IoT Devices
Artificial Immune-Base (ANN)	Network Intrusion	Distributed locally in the IoT Devices
ZeroR, KNN, SVM, NaiveBayes, Neural Network	Network Intrusion	Distributed with MEC
Negative selection algorithm	Network Intrusion (DoS/DDoS)	Distributed with cloud, Fog, and MEC
Stacked Autoencoder Deep Learning	Network Intrusion Attacks	Distributed with MEC
Multi-Layer Deep Network	Network Intrusion Attacks	Distributed Locally and Fog
LSTM Deep Learning	Network Intrusion Attacks	Distributed Locally and Fog
Markov Model and Virtual Honeypot Device	Network Intrusion Attacks	Distributed with Fog to MEC
Offensive Decoy technology	Cloud Data attacks	Distributed with cloud and Fog
Multilayer Perceptron (MLP)	Network Intrusion Attacks	Centralized on Fog
MapReduce approach	anomaly-based and misuse-based attacks	Distributed locally on IoT
Co-responsible distributed NIDS	Network Intrusion Attacks	Distributed between IoT and ISP's Server
Sample extreme learning Machine	Network Intrusion Attacks	Distributed between IoT, MEC, and Fog

Fig 1(a): These authors went on to publish papers with their detection methods and the threat identified along with their placement strategies, these papers had their shortcomings with augmentation of MLaaS services now being a part of the IoT chain and environment for setting up the infrastructure.

Attacks	Mode of Attack Initiation
Spoofing	Impersonation
Denial of Service (DoS)	Network Flooding
Distributed Denial-of-Service (DDoS)	Network Flooding
Jamming	Fake Signaling
Man-In-the-Middle	Eavesdropping Packets
Privacy Leakage	Attack Authentication Storage
Marai Botnet Attack	Malware Implant on Devices
Sybil Attack	Creates Anonymous Identities
AI-Based Attacks	Creates AI-powered Tools

IV. IOT AND N-IDS .

The augmentation of MLaaS and other modern services such as edge computing etc. make use of Machine Learning approaches in order to achieve their required goal as the technology is known to perform better at predicting and thwarting various kind of attacks that have been identified in the literature survey section. Below I present in detail each of the attack surface and how previous approaches have failed in implementing complete protection against these types of attacks.

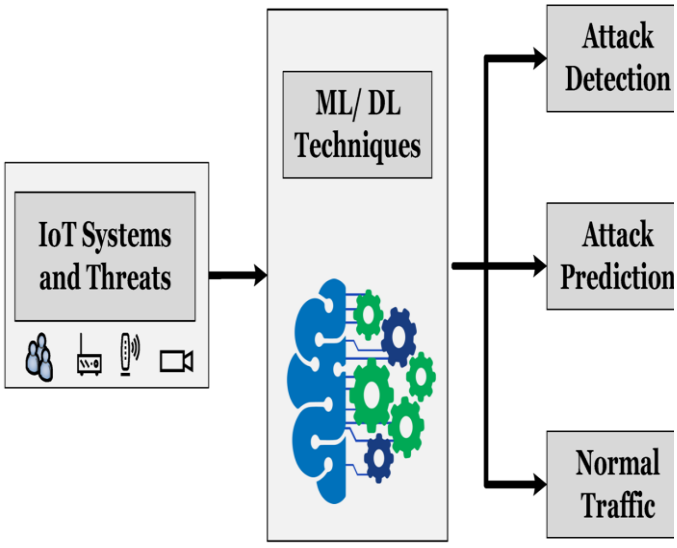
Following are some of the common type of attacks that are carried out over a network intrusion detection system.

A) Membership inference Attack:
Membership Inference Attack and its consequences in our problem setting would mean we have access to some trained samples and based on the features of this sample we craft our samples based on our inference and use these samples to launch our attack. The leakage happens because the model that was trained has inferred and lost its privacy due to the fact that there has been a member that has given up on its features. We get a face sample that was used to train the model and based on the sample we find out how the face classification takes place and craft our payload.

B) Data poisoning Attack:
The data that is being fed to the system for validation has been compromised and this data produces malicious output. False positive or false negative both errors can be generated.

C) Model Evasion Attack:
Model evasion attacks are carefully crafted adversarial examples that are created with the single goal that it can not be detected by human eye but computer gets spoofed to wrongly classified. In a dataset with millions of dataset this becomes a practically impossible task.

D) Adversarial Attack:
Adversarial attacks are attacks that are carried out in which we add noise (lambda, gaussian) to make misclassified predictions by the ML model. To human eye these seem to be normal.



A) Ddos Attack- Ddos attack is an attack that tries to shut down authentic and organic traffic from a network by flooding it with packets of malicious and arbitrary payload and make the system to slow and eventually faze out due to the amount of malicious activity drying up the computer's resource.

B) Probe- A recce attack that tries to gather information of the network, goal here is to eavesdrop and get valuable information as much as possible. This attack can also be seena s the first step towards reverse social engineering a more sophisticated and complicated attack.

C) U2R- This attack starts off with a normal user and based on the access level granted by the system on a network, he tries to gain access as a super-root user, the attacker attempts to exploit the vulnerabilities in a system to gain root privileges and access.

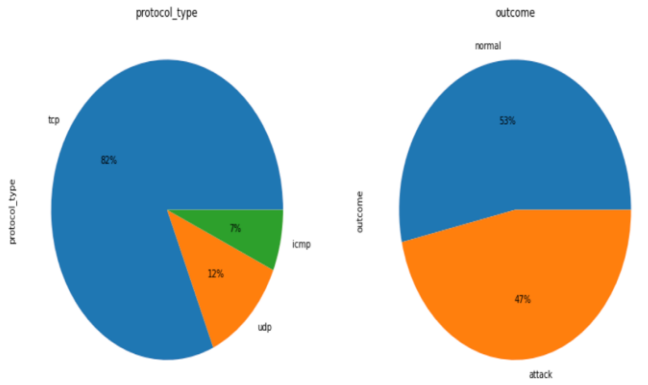
D) R2L – This attack tries to gain local access of a remote machine, it is understood the adversary has no access to local machine but through chaining of attacks he tries to get the privilege and gain access to the local system.
Machine learning approaches can be equipped with deep learning approaches and they can be used to provide solutions as per the above diagram. Here a generalized block diagram shows that ML and DL techniques can be used to divide and segregate the incoming traffic to the Network detection system in 3 parts:

a) **Attack Detection** – This is where the attack detection takes places.
b) **Attack Prediction** – This is where ML tries to predict which node of incoming traffic is going to cause an attack.
c) **Normal traffic**- Systematic and smooth handling of normal traffic so as to not be blocked by our ML approaches as False Negatives.
Typical N-IDS approaches can be classified as following, we would be creating our N-IDS based on state of art ML techniques along with DL approaches and show how they are better than classical statistical methods that do not protect against more modern attack approaches. N-IDS can have multiple approaches i.e it can be used to detect attack, prevent attack from being launched in the first place or identify intrusions in the system or thwart denial of services where the IoT environment essentially is starved of Computational resources.
Based on the extensive review and technologies discovered that are being used currently to prevent against network intrusion here we present the **Advantages of our proposed N-IDS for CIS-600:IoT equipped with state of art ML technologies that we have implemented**:

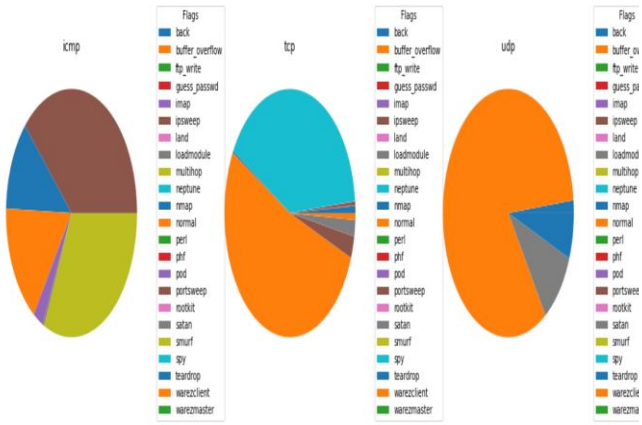
1) Protocol Display- Content of the packets being sent over the network can be observed to manually see if the type of data being sent is malicious, compared to firewall which only shows Ports and IP addresses.

Here we see a snippet of the above point, traffic of TCP,IDP & UDP, graph 2 shows its split between attack and normal.

```
pie_plot(data_train, ['protocol_type', 'outcome'], 1, 2)
```



2) Content of packets visible: When an NIDS performs protocol analysis, it looks at the TCP and UDP payloads. The sensors can detect suspicious activity because they know how the protocols should be functioning and what is within these packets.
We further show.info() for tcp,udp and icmp attacks as follows :

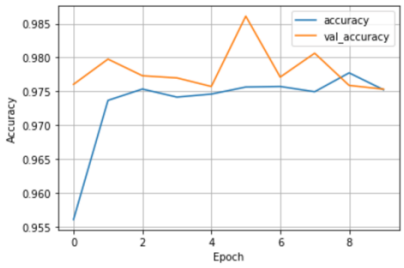


Our monitoring strategy uses a series continuous network protocols and capable of learning subtle distinction between threats and legal ones. Historical information fed the machine learning model classify Sybil and other attacks (ddos, probe, u2r). It helps to identify recurring patterns of Sybil, DDoS, U2R, Probe attacks and locate in long-term traffic chain.

3) In-par with NIST regulations- Because IDS gives greater visibility across your network for all protocols of transmission it also can implement IDS Logs and this can be used as part of documentation to meet a fair share of requirements in relation to keeping a footfall of the traffic over your network,

4) Boost Efficiency- Because IDS sensors can detect network devices and hosts, they can inspect the data within the network packets and identify the services or operating systems that are being utilized. This saves a lot of time when compared to doing it manually. An IDS can also automate hardware inventories, further reducing labor. These improved efficiencies can help to reduce an organization's staff costs and offset the cost of implementing the IDS.

```
plt.plot(history.history['accuracy'], label='accuracy')
plt.plot(history.history['val_accuracy'], label='val_acc
uracy')
plt.xlabel('Epoch')
plt.ylabel('Accuracy')
plt.legend()
plt.grid(True)
```

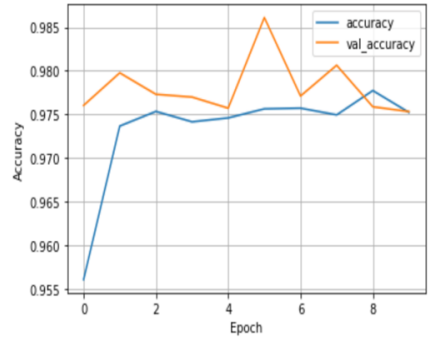


V. PRIVACY vs SECURITY vs UTILITY.

Here I come up with 2 scenarios of how user privacy can be violated and how it affects the security and how our N-IDS makes use of secure computing to make sure these privacy scenarios are not exploited by the adversary in order to get access to sensitive data.

1) As seen in above section a user with more privileges than it should have can cause situations of authentication breach in the system. A common way to achieve that identified through literature review is to make sure that the traffic which flows in has its origin traced and the content displayed to us, this way polymorphic worms that tend to hide in terms of evading our system at the time of injection can be traced and thwarted. To make the system perform up to mark it is very important for us to not incur loss in turn for utility that is our accuracy should be a true reflection of our classification of false positives and false negatives. It is always the case in previous IDS there have been certain amount of trade-off between security and utility, utility here means ease of use for the operator and security here means the hard bounds under which we make our IDS perform and generate outputs. It has been noted that the loss of training to testing samples is the grey area, meaning that the drop of accuracy in testing vs training is where the malicious samples tend to pass the IDS, key lies in creating an IDS such that our model does not incur losses while transitioning from a learned model to new model that processes and works with new unseen data.
The following results justify the fact there is no significant drop in accuracy between testing and training. Epoch here

is our optimization being performed on the model, higher epoch means more training time. There is a trade-off between the computational power and the accuracy our model predicts with. The approach is different from other prior privacy-preserving IoT systems that look at privacy only from the perspective of data management and worry about data breaches that reveal sensitive information. We warrant the fact that information leak can also be targeted through various ML approaches and that it is necessary to implement state of art technologies to thwart these attacks more actively after identifying them.



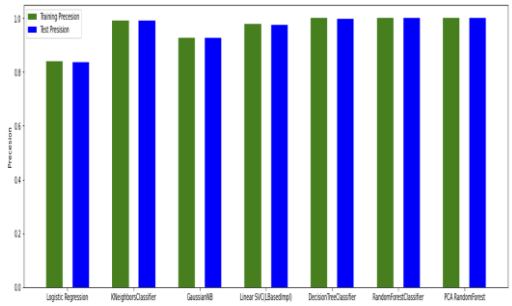


Figure on top is per epoch vs acc and val_acc, it shows as epochs increase the 2 converge.
Figure on bottom is training vs testing precision of each algorithm, we can see they match which means our testing results match training results.

VI. Proposed Draft, Policy and Regulation through Technology and Legislation.

As the Internet of Things (IoT) grows to connect an amazing diversity of devices to electronic networks, four new publications from the National Institute of Standards and Technology (NIST) offer recommendations to federal agencies and manufacturers alike concerning effective cybersecurity for these devices. Recently European countries have started looking into drastic measures to implement until the existing technology is either secured or made more compliant to their regulations and failure to do so would mean stopping the implementation of those technologies, not just putting the companies under regulation. One such country to be implementing such a drastic measure recently was Italy where they completely banned use of facial technology until further use. This was done to safeguard the interest of citizens of Italy with respect to the harms of using facial technology as surveillance. A uniform joint standard developed by industry experts and outlined by NIST based on current guidelines and it should fit in their (OMB) Office Management & Budget. A thorough act which makes reporting of all sort of IoT break-ins to the authorities. Secure, timely firmware updates for the system to be compliant with recent technology and up to date against all current OWASP guidelines. A government developed pen-test tool to check the system is secure against basic vulnerabilities which can be thwarted and secured at the first line of defense. Mandatory secure code development trainings and seminars assigned to make cyber-secure IoT space accessible not just for people in tech. but also layman.

Besides this it is important to safeguard critical data, stop endpoint data leaks, implement copy protection, regular security audits, containing ransomware attacks, periodic firmware updates etc.

The fact that most IoT solutions are provided by third parties, along with the pervasiveness of the collected data, raises privacy and security concerns. There is a need to verify which data is being sent to the third party, as well as preventing those channels from becoming an exploitation avenue. A new architecture with existing API definition languages to create contracts which define the data that can be transmitted, their format and constraint is required. To verify the compliance with these contracts, a Network Service architecture which validates REST-like API requests/responses against a Swagger schema should be implemented. Deal with encrypted traffic using a Service Function Chaining (SFC)-enabled Man-in-the-Middle (MITM), allowing verifications in “real-time.”.

All the data that is being used by machine learning models should even in the most primitive form have some sort of data prevention techniques on them as regulation through technology is a developing and upcoming field where we push the developers to make use of modern and existing mathematical and computational approaches to generate and assure mathematical guarantees about our data, its bound and privacy budgets. This can be achieved using either differential privacy or making use of federated learning to divide the task of data handling at various stages rather than one point of source contact. The above said have their own set of challenges and difficulties.

Another aspect could be to avoid using data which is sensitive and which we know can cause significant damage to the individual when leaked, various studies have gone on in the field of allurism in cyber-security which deals with the catastrophic event of a cps failing and there being maximum information, privacy and security loss to the user the system was supposed to safeguard. Such a user is generally unaware about the consequences of such a breach until fully made aware.

Zero Trust is a security framework requiring all users, whether in or outside the organization’s network, to be authenticated, authorized, and continuously validated for security configuration and posture before being granted or keeping access to applications and data.

Zero Trust assumes that there is no traditional network edge; networks can be local, in the cloud, or a combination or hybrid with resources anywhere as well as workers in any location.

Zero Trust is a framework for securing infrastructure and data for today's modern digital transformation. It uniquely addresses the modern challenges of today's business, including securing remote workers, hybrid cloud environments, and ransomware threats. While many vendors have tried to create their own definitions of Zero Trust, there are a number of standards from recognized organizations that can help you align Zero Trust with your organization.
Align to the NIST 800-207 standard for Zero Trust. This is the most vendor neutral, comprehensive standards, not just for government entities, but for any organization.

VII. CONCLUSION

With growing number of iot regulations we can find the right balance of trust and usability between the consumer and service provider, but along with that the paper also shows the hazard of blindly using technology without understanding its underlying threats and analysis. it also opens a wide range of research areas where further attacks, defense and aacademia research is required along with legislation, regulation and policies. besides this the people implementing and developing these technologies should have a moral duty and responsibility towards clean and secure computing for everyone. With growing number of IoT regulations we can find the right balance of trust and usablity between the consumer and service provider, but along with that the paper also shows the hazard of blindly using technology without understanding its underlying threats and analysis. It also opens a wide range of research areas where further attacks, defense and academia research is required along with legislation, regulation and policies besides this the people implementing and developing these technologies should have a moral duty and responsibility towards clean and secure computing for everyone. A detailed analysis and implementation which helps us understand a IoT enabled C.P.S in real world application and its challenges due to various reasons throughout the development cycle and its implementation. Future discourse for further studies to be driven in this direction for responsible IOT which gives mathematical guarantees for data and its management and security. A unique threat model with all attack surfaces and their prevalent defensive approach and a mitigation model to serve as reference. Create a theoretical pipeline for technology implemented and the defense that should be paired along with it.

REFERENCES

[1]- https://www.oracle.com/internet-of-things/what-is-iot/. – IoT Is here to stay.
[2]- P. J. Ryan and R. B. Watson, "Research Challenges for the Internet of Things: What Role Can OR Play?," Systems, vol. 5, no. 1, pp. 1–34, 2017.
[3]- The relation between IoT and Cyber physical Systems https://en.wikipedia.org/wiki/Cyber-physical_system.
[4]- Kumar, S., Tiwari, P. & Zymbler, M. Internet of Things is a revolutionary approach for future technology enhancement: a review. J Big Data 6, 111 (2019). https://doi.org/10.1186/s40537-019-0268-2.
[5]- IoT Considerations, Requirements, and Architectures for Smart Buildings—Energy Optimization and Next-Generation Building Management Systems,https://doi.org/10.1109/JIOT.2017.2647881.
[6]- A Practical Evaluation of Information Processing and Abstraction Techniques for the Internet of Things, https://doi.org/10.1109/JIOT.2015.2411227. Cognitive Internet of Things: A New Paradigm Beyond Connection
[7]- A Reference Model for Internet of Things Middleware, https://doi.org/10.1109/JIOT.2018.2796561.
[8]- A Survey of Clustering Techniques in WSNs and Consideration of the Challenges of Applying Such to 5G IoT Scenarios, https://doi.org/10.1109/JIOT.2017.2726014.
[9]- A Survey on Security and Privacy Issues in Internet-of-Things, https://doi.org/10.1109/JIOT.2017.2694844.
[10]- Fog and IoT: An Overview of Research Opportunities, https://doi.org/10.1109/JIOT.2016.2584538.
[11]- Operating Systems for Low-End Devices in the Internet of Things: A Survey, https://doi.org/10.1109/JIOT.2015.2505901Securing the Internet of Things: A Standardization Perspective, 10.1109/JIOT.2014.2323395.
[12]- Internet of Things for Smart Cities https://doi.org/10.1109/JIOT.2014.2306328.
[13] O. Garcia-Morchon et al., "Securing the IP-based Internet of Things with HIP and DTLS," in Proc. 6th ACM Conf. Security-Privacy), 2013, pp. 119–124.
[14] D.Hardt,"TheOAuth2.0authorizationframework,"RFC6749,Oct.2012.
[15] K. Hartke, "Practical issues with datagram transport layer security in constrained environments," draft-hartke-dice-practicalissues-00, IETF, 2012.
[16] K. Hartke and H. Tschofenig, "A DTLS 1.2 profile for the Internet of Things," draft-ietf-dice-profile-00, IETF, 2014.
[17] J.HuiandP.Thubert,"CompressionformatforIPv6datagramsoverIEEE 802.15.4-based networks," RFC 6282 (Proposed Standard), Sep. 2011.
[18] IETF, "Constrained RESTful environments (CORE) WG," 2013.
[19] IETF, "DTLS in constrained environment (DICE) WG," 2013.
[20] IETF, "Lightweight implementation guidance (LWIG) WG," 2013.
[21] IETF,"Authenticationandauthorizationforconstrainedenvironments(ace) mailing list," Apr. 2014.
[22] S. Kent and K. Seo, "Security architecture for the internet protocol," RFC 4301, Dec. 2005, updated by RFC 6040.
[23] S. L. Keoh, S. S. Kumar, and O. Garcia-Morchon, "Securing the IP-based Internet of Things with DTLS," draft-keoh-lwig-dtls-iot-02, IETF, 2013. [24] S. L. Keoh, S. S. Kumar, O. Garcia-Morchon, and E. Dijk, "DTLS-based multicast security for low-power and lossy networks," draft-keoh-dice-multicast-security-06, IETF, 2013.
[25] S.L.Keoh,S.S.Kumar,andZ.Shelby,"ProfilingofDTLSforCoAP-based IoT applications," draft-keoh-dtls-profile-iot-00, IETF, 2013.
[26] T.Kivinen,"MinimalIKEv2.Internet-draft,"draft-ietf-lwig-ikev2-minimal-01, IETF, 2013.
[27] S. S. Kumar, S. L. Keoh, and O. Garcia-Morchon, "DTLS relay for constrained environments," Internet-draft, IETF, 2013.