May 2022

Research-Based

DOCUMENTATION

Ahmad Naeem
Ahmadnaeem26831@gmail.com
Student at Superior University

Zumer Sarfraz
Bcsm-f19-448@superior.edu.pk
Student at Superior university

Hafsa Majid
Bcsm-f19-433@superior.edu.pk
Student at superior university

Federated Learning and its use in Cyber Security

1. Cyber Security

Cybersecurity is a process in which an individual or an enterprise tries to protect their data or information from various cyber threats as we know that the local and global devices are interconnected to each other via a network so, there are many risks of viruses, attacks or leakage of information. According to stats, in 2018 80,000 cyber-attacks happened per day and 30 million per year.

Cybersecurity is a process through which we can protect our data from these threats [4]. As we know that our data goes through the different layers in a network and if it's sensitive data or personal data of our user, it should be a priority to protect it [5]. The hackers can easily make access your data through some web servers, cloud computing, mobile networks, etc.

For protecting and securing your data from illegal or unauthorized access cybersecurity is spreading around the world [6]. Federated learning can be utilized to minimize cyber attacks and to achieve data privacy and security. It manages the set of techniques accustomed to saving the integrity of networks, programs, and data from unauthorized access. As fast as security grew, the hacking world grew faster.[35]

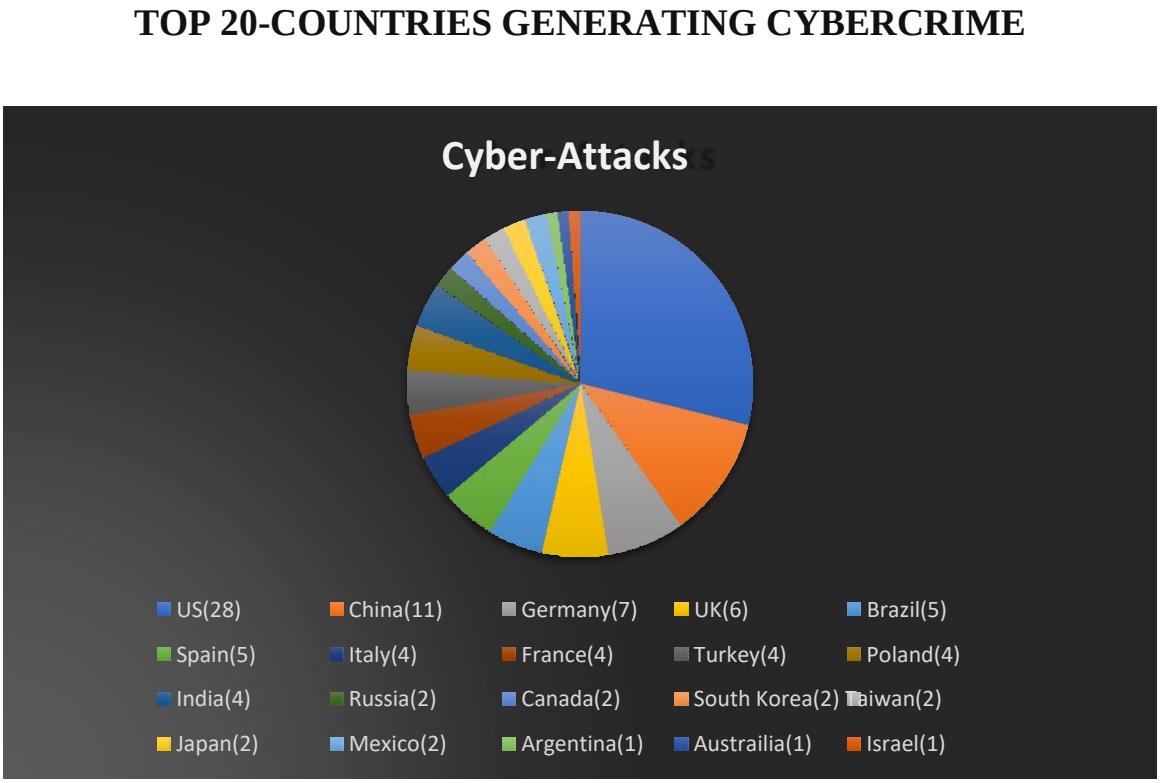


Figure 1.1 Cyberattacks

1.2 Trends in Cyber Security

- **Remote working from home:**

As we all know, following the Covid -19 epidemic, the majority of work has moved online. Offices and schools use several apps to communicate and access information, which increases the danger of cyberattacks because essential data can be separated from its sources. People working from home are using several local devices such as laptops, mobile phones, and computers, which has resulted in increased security challenges[20]. It may become difficult to distinguish between allowed and unauthorized access across several devices.

- **IOT:**

The Internet of Things (IoT) is a rapidly evolving trend in our culture. Smart microwaves, refrigerators, and watches, for example, are all controlled by home automation systems like Alexa[20]. As we may assume that there will be multiple devices connected in this network with few CPUs and storage, installing anti-virus and firewalls will be tough. This can put cybersecurity at risk.

- **Ransomware:**

Ransomware is malware that prevents the victim from accessing data and information unless the victim pays the attacker a ransom price. The ransom deadline may arrive at a particular time, and if the victim does not pay the ransom money, the data may be deleted forever[21]. Generally speaking, the ransom is a simple way for hackers to get money nowadays. If a device connects to the internet, it should have the most recent software security patches installed and anti-malware that detects and stops ransomware.

- **Increase in cloud security threats:**

It is one of the biggest cyber security trends nowadays. The necessity of cloud services is getting increased day by day due to remote workforces, especially after the pandemic. Cloud services offer good storage, efficiency, and vulnerability. The data stored on cloud services should be safe and secured so that there will be no data leakage.

2. Sub-domain of Cyber Security Prediction

Cyber security is a method, process, and technique that involved protecting data, information, computing system, software applications, and networks from cyber-attackers. Currently, cyberattacks became even more complicated and advanced. A report was found by computer crime and security analysis that malware was injected, computer theft, phishing, and bot attacks were administered on the premise of common cyber-attack concepts to get sensitive material, causing damage to the organization.
Cyber security could be a very imperative part of the cyber world. Because cyber security may be a method of defending information, software, network, and facts from unauthorized access or attacks. Cybersafety performs a necessary function within the realm of statistics technology. Securing the knowledge, network, records, and also application has come to be one of the most important challenges.

Cyber security may be an important part of the cyber world. Because cyber security could be a performance of protecting data, application systems, and data from illegal access or attacks. Cyber

security shows a necessary part in the arena of data technology.
Cybersecurity covers the physical and non-physical security of knowledge against internal and external attacks or threats. Non-physical threats like viruses and piracy come from illegal access to the system.

2.1 Sub-domain process:

Detection of Subdomain Vulnerability In Making a final list of the available or valid domains and subdomains, we've to seek out whether their DNS records are assigned to external services or not:

- **Checking DNS Record:**

We can check the DNS record of a subdomain by using Linux command dig or DNS lookup. If the subdomain appears to assign to an external service, then we'll move forward to test its HTTP response.

- **Checking HTTP Response:**

Either by visiting the subdomain we will see the HTTP response or using the Linux command curl we can also detect the HTTP response. If we get 404 responses and see a default 404 error page of the external service, we will initially assume the subdomain as vulnerable and move forward with the takeover process.

3. Introduction to Machine Learning:

Machine learning is a developing field that is getting used a lot nowadays. Mainly the goal of machine learning is to understand the sequence or structure of data and train that model according to that data so that it can be used by people. Machine learning algorithms trains the model by using some specific output results so that our model will be ready for all kind of known and unknown results. This raises the efficiency of our model.
There are mainly 2 kinds of machine learning:

- Supervised learning: In supervised learning, the input is given with the labeled output.
- Unsupervised learning: In unsupervised learning, the data is mostly unlabelled.[50]

Machine learning is a subfield of artificial intelligence (AI). The goal of machine learning generally is to understand the structure of data and fit that data into models that can be understood and utilized by people.

3.1 Importance of Machine Learning in Cyber Security!

Consider the three areas where most cyber-ML algorithms are finding application:

- **Intrusion detection:**

Intrusion detection aims to seek out illicit activities within a computer or a network through Intrusion Detection Systems (IDS). These systems traditionally supported patterns of known attacks, but modern deployments include other approaches for anomaly detection, threat detection [7], and classification-supported machine learning. A botnet may be a network of infected machines controlled by attackers and misused to conduct multiple illicit activities. Botnet detection aims to identify communications between infected machines within the monitored network and thus the external command-and-control servers. DGA automatically generates domain names and is often employed by an infected machine to talk with an external server(s) by periodically generating new hostnames. They represent a real threat to organizations because, through DGA which relies on language processing techniques,

it's possible to evade defenses supported by static blacklists of domain names. We consider DGA detection techniques supported ML.

- **Malware analysis**

Malware analysis is an especially relevant problem because modern malware can automatically generate novel variants with identical malicious effects but appear as completely different executable files. ML techniques are accustomed analyze malware variants and attributing them to the correct malware family.

- **Spam and phishing detection**

Spam and phishing detection include an outsized set of techniques geared toward reducing the waste of its slow and potential hazards caused by unwanted emails. Nowadays, unsolicited emails, namely phishing, represent the popular way through which an attacker establishes a primary foothold within an enterprise network. Spam and phishing detection are increasingly difficult thanks to the advanced evasion strategies employed by attackers to bypass traditional filters. ML approaches can improve the spam detection process.

3.2 Importance of Machine Learning in subdomain

While the domains seek advice from the domain names listed at the name System (DNS), they will even be malicious, disseminating malware, hosting phishing or spam or scam webpages, or facilitating command and control (C&C) communications.

- **Malicious domains:**

Malicious domains are one of the first resources leveraging attacks over the web. These malicious domains used for constructing malicious URLs are a really common and heavy threat to the protection of cyberspace. they will be accustomed lure users into becoming victims after they visit its phishing, drive-by downloads, spam, and other contents, which can lead to compromise of user's privacy, or may incur loss or may lead to a malware installation onto the user's machine. the name and therefore the subdomain name are the important parts of a URL. the mix of the name and also the subdomain name together form the hostname and represent the physical machine on which the resource is hosted.

- **Domain Name System:**

Domain Name System the Domain Name System (DNS) is one of the core protocol suites of the web, which takes the responsibility of directing requests for Internet resources to absolutely. Domaine domain System (DNS) may be a decentralized, well-distributed, hierarchical naming system for resources connected to the web, which translate human-readable and rememberable domain names into their respective IP (Internet Protocol) addresses, and sometimes the vice-versa too. In brief, DNS may be a simple service for lookup and translation of URL into IP address and contrariwise.

- **Malicious Domain Detection:**

Malicious Domain Detection The zoom of the Internet across varied thematic areas and their related threats keeps alive the challenge of malicious domain detection. Broadly malicious domain detection is often performed either on active DNS data or passive DNS data, where active data implies obtaining DNS data by purposefully sending DNS queries and recording the respective responses for analysis and detection, and passive data implies accessing DNS logs to induce real DNS queries and responses,

alternatively plant a sensor at the DNS servers to capture the queries and responses for analysis and detection.

4. BackGround:

As everyone knows, Machine learning requires lots of data which transfers to the central server for predictions and direction/advice. A trained ML model contains unintended features that may be utilized to extract personal information. Thus, local ML model parameters from a federated optimization algorithm may be exploited by an adversary to infer personal information, particularly when combined with related information like model organization and meta-data [1]. The downside of this process is there is no privacy for data, the cost is too high and the main problem is cyber attacks.

4.1 Rise of Federated Learning

Federated learning(FL) considers being the new dawn of Ai[2]. A federated machine is based on a decentralized model. It permits data exchange between users by sharing parameters and meta-information of machine earning models, while not violating the privacy necessities. every user stores their knowledge domestically and has no access to alternative users' knowledge. Federated learning is different from distributed learning.

In machine learning, data is uploaded to the server to train the model but in federated learning, models can be trained without sending their actual data to the server. It provides a method to protect user privacy.

For example, a user wants to autocomplete the model. Users don't want to share their private data anywhere. They want an efficient model without any data communication. So in this condition, Federated Learning can be utilized to construct a autocomplete model. They make a well-trained model and send it to n number of user devices [2].

There are Four steps in federated learning to train the model.

- **Transfer untrain model:**

  The n nodes (edge devices) receive the untrained model from the central server.

- **Local Training:**

  Each node trains the model locally by using its local data and making a new update parameter.

- **Parameters received:**

  The central server receives the updated parameters from the local environment.

- **Aggregation:**

  The central server aggregates all models makes a highly efficient model and sends it to nodes.

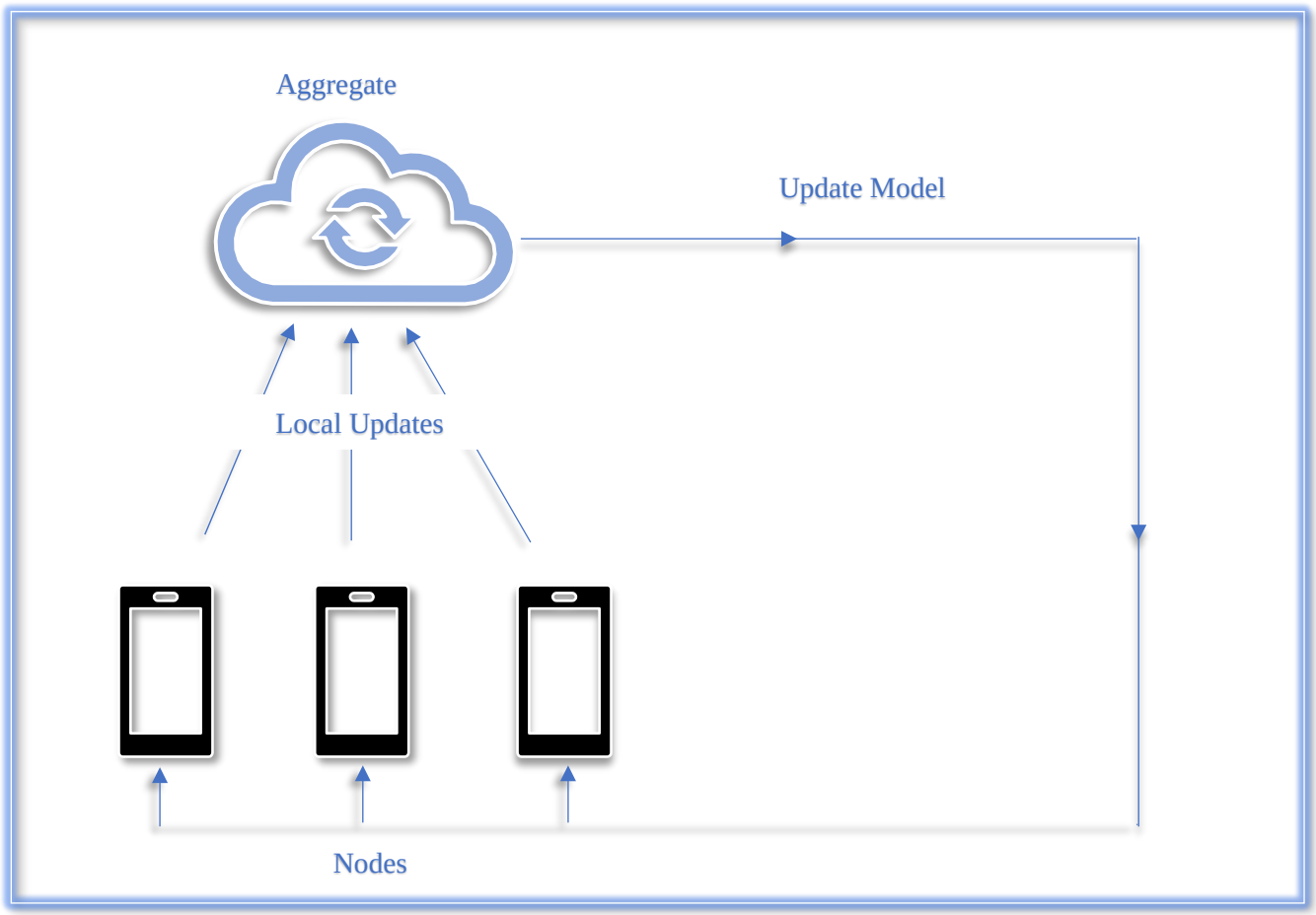


Figure 4.2 Implementation of federated learning

4.3 Federated Learning Functionalities

We can achieve the following important functions by utilizing federated learning.

- **Data Seclusion:**

The Central server sends the untrained model to the independent nodes the nodes train the model by using their local data without sending their data to the server. So that's how we can maintain our privacy.

- **Data Safety:**

There are following cyber attacks like SQL injection, man-in-the-middle, and DoS attack these attacks can be solved by using federated learning. There is no data communication between the node and server. Node just sending updates to the server.

- **Data Hiding:**

Data can cause cyber attacks if data is accessed by any unauthorized person. By using federated learning, data can be hidden from unauthorized persons. Only authorized people can access your confidential and private data.

- **Availability**:

Availability plays a vital role as it is most important that the data should be available to the authorized person. It includes the availability of data to the authorized user and the unavailability of data to unauthorized users [4][5].

- **Probity:**

The major part of cyber security is maintaining consistency, data can be accurate and complete. Attackers can alter the sender data before it reaches the receiver. In federated learning, your data can be secure because it is designed for privacy preservation and privileged data can not be leaked from the local environment.

5. Privacy Preservation

As we above stated, millions of applications are dependent on the internet. So, private data can be stolen by any third party easily. Cyber-attacks and data leakages are becoming more regular and expensive to handle[3]. Attackers steal the data held on the internet for training purposes, that data can be used to recognize people or other beneficial information that can be sold. Distributed learning can not solve these issues there is no privacy for data. Organizations need a model that solves all these issues.

There are the followings techniques to solve the issue of privacy preservation

- Differential Privacy[3]
- Homomorphic encryption[3]
- Multi-party computation[3]
- Federated Learning[3]
- Ensemble privacy-preserving techniques[3]

We just discuss federated learning. How federated learning solves the privacy issues.

Federated learning permits machine learning activity to be decentralized and minimizes the quantity of raw data displayed from independent nodes. The privacy of data is preserved and the danger of data-stealing is reduced. A Model can be trained in a local environment without sending raw data to the central server. Only the updated parameters are sent to the server by the independent nodes.

6. The problem in Federated Machine Learning

The implementation of FL for wireless networks has the following five main viewpoints and challenges:

- **Scalability**

FL should be scalable since more computers or processors can counteract the rising volume of data and provide a solution to the complexity and memory issues that arise in large-scale learning networks. It's critical to investigate challenges related to distributed training for a large-scale learning network.

- **Privacy and security**

Because only the locally obtained FL model is transferred to the middle in FL, each user's knowledge set is extra safeguarded. An eavesdropper can, nevertheless, do an approximate reconstruction of the

starting data, especially if the local and global model coefficients aren't safeguarded [7]. Furthermore, the local FL model may provide personal details.

Privacy in Federated Learning can be divided into two categories: global and native. Except for the BS in global privacy, the model creation at each iteration is undetectable to any or all or any unknown devices. Each iteration's model aggregation is private to any or all unknown third parties, as well as the BS in local privacy.

- **Asynchronous communication**

FL involves information exchange between wireless devices**.** Synchronous communication methods are simple, but they'll introduce stragglers among devices. a horny because of alleviating laggards during a heterogeneous environment is an asynchronous solution. Although asynchronous server parameters within the distributed data center are successful in addressing stragglers, assumptions of bounded delay could even be impractical in federated schemes.

- **Non-independent identical distribution (Non-IID) devices**

When training a joint model from differently distributed data across devices, challenges arise both in terms of knowledge modeling and analyzing the convergence trend of the relevant training process **[8]**. One key aspect of FL is managing heterogeneous settings and competing and distributed decision-making environments**.**

- **Joint communication and computation design:**

To deploy FL in an exceedingly very wireless communication network, each device possesses to transmit its multimedia data or local training results through an unreliable wireless link. it's vital to contemplate the multicell and multi-hop FL implementations for real scenarios **[9]**. additionally, the performance of FL learning schemes is degraded by limited radio resources. Thus, it's vital to contemplate the joint management of communications and computing resources to grasp efficient and effective FL.

7. BlockChain

There are some existing problems in federated learning as we discuss above. Organizations want a model in which they interact with minimum threats to privacy. So, blockchain can help us to sort out all the problems that organizations and companies face in FL.

Bitcoin was introduced in 2008 by Nakamoto based on a peer-to-peer payments system. Bitcoin is decentralized and now it is the largest cryptocurrency on Earth. Bitcoin gets support from blockchain which provide immutable data records and provides rewards to a node based on their contribution. Blockchain has several properties.

- Blockchain uses peer-to-peer networks so there is no need for any central server of a third party all nodes are equal. No need for any central node for network propagation.
- It is difficult to alter the data because blockChain is immutable.
- Through the encryption of private information user’s privacy can be prevented by attackers.
- Data can be traced by its source with the use of a special structure in the blockchain.

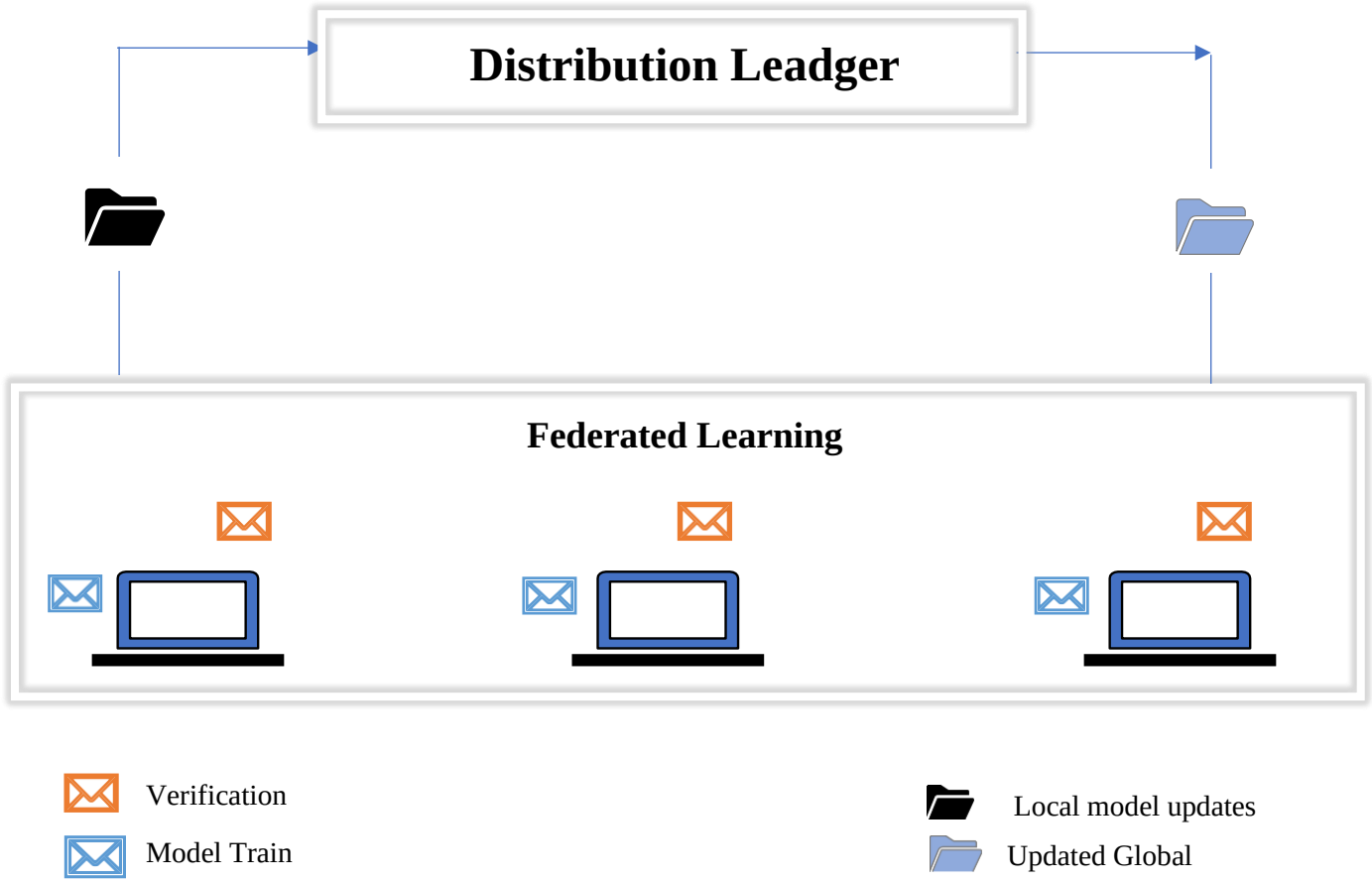
7.1 Blockchain Federated Learning:

Blockchain Federated Learning(BCFL) can solve the problems of federated learning. Fully coupled BCFL(FUBCFL) is an architecture of BCFL. In which clients not only trained the model but also verify the model and also generates the new blocks. In FL there is a central server that collects all the trained models from the nodes for aggregation. The role of the central server can take place by a blockchain. There are two methodologies of this framework. 1) Selected nodes collect the verified local model updates and then perform an aggregation process. 2) nodes can also take part in the aggregation process. Training data, local verified models, global models, and data that are produced during the learning process are stored in a distributed ledger.

- The client trains the model locally and verifies it.
- The local model is collected by nodes and then updates the global model.
- A new block is generated in distributed ledger and stores the verified models.
- And then updated models provide to the nodes according to their contribution.

In this mechanism, blockChain works as a distributed ledger that holds training data, a local updated model, and a global model. Private data is secured and malicious nodes are prevented.

Blockchain Environment



7.3 FuBCFl Implementation Diagram

8. Literature Work:

There are multiple papers in which different models and techniques are applied to predicting cyber attacks. Like, in the first paper (see table 6.1) they predict network intrusion by shallow learning with the technique of decision trees. And then also discuss how deep learning is more suitable in network intrusion than shallow learning(2020). In the second paper, they used a graph to predict the attack path and then use a recommender system to predict future cyberattacks(2018). Data mining techniques can also be used to predict cyber attacks. In the third paper, they are using a decision tree to build a model by extracting the data of historical attacks, and then this model is used to predict future attacks(2020).

The next paper tells us about the application of deep learning for predicting cyber attacks. The current detection implementation is deficient and increases the need to build more applicable prediction models and approaches. So, they apply different models like LSTM, and RNN to predict the attack that happens. The accuracy of LSTM is 93%(2020). The 26th paper, tells us how a zero-day attack can predict by using deep learning with the use of a bidirectional recurrent neural network(BRNN). The accuracy of this technique is 92% (2020).

The 27th paper applies the meta-learning method to predict the dos, R2L, and probe attacks by using Bi-LSTM, Gru, and LSTM techniques(2021). The 28th paper discusses how to enhance machine learning prediction in cyber security, they apply the deep learning method by using different techniques like Fuzzy Logic and generic Algorithm. The dataset they are using is DARPA k_DD99 and the accuracy get after applying this model is 94.60% (2021).

Ref.	ML Method	ML Type	ML Techniques	Cyber Attacks	Dataset	Accuracy	Year
[22]	Shallow Learning, deep learning	Supervised learning	Gradient Boosted Decision Trees, Convolutional LSTM	Network intrusion	Warden	63%	2020
[23]	Recommender systems	Supervised learning	Risk management	Graph attack	user submitted ratings		2018
[24]	Data mining technique	Unsupervised learning	decision tree algorithm	DDoS, PortScan	CICIDS	99%	2020
[25]	Deep learning		LSTM, RNN		CTF	93%	2020
[26]	Deep learning		Bidirectional recurrent neural network	Zero-day attack		92%	2020
[27]	Meta learner	Supervised learning	Bi-LSTM,Gru,NN-LSTM	Anomaly	(NSL-KDD)(UNSW-NB15)	99.64%	2021
[28]	Deep learning	Unsupervised learning	Fuzzy Logic and Genetic Algorithm	Dos, R2L, U2R, PROBE	Darpa K-DD99	94.60%	2022
[29]	IDS model	Unsupervised learning	Clustering	Dos, R2L, U2R, PROBE	K-DD 99 cup	94%	2020
[30]	Machine-learning	Supervised learning	Support Vector Machine Linear	Perpetrator	StandardScaler	95.02%	2020

[31]	Statistical Machine Learning	Supervised learning	NB,KNN,SVM,DT	Cyberattacks	CVE	96.70%	2019
[32]	Utilized machined learning	Supervised learning	ANN, DT, RT, XGBooost	MPS,MPG,EID	GIS	98.60%	2021

Table 8.1 Literature work

The 29th paper also predicts the same attacks which we are discussing in paper seven. But they apply the Intrusion Detection System(IDS) method by using the Clustering technique. The dataset they are using is the K-DD cup and its accuracy gets 94% which is close to the last model (2020). In the next paper, the author analyzes cyberattacks in several models. They used machine-learning methods and predict the effect of the defined features on the detection of the cyber-attack method and the perpetrator. The accuracy they get is 95.02 by using the Support Vector Machine Linear technique(2020). In the next paper, they predict attacks by using statistical machine learning. They use a dataset of CVE and apply NB, KNN, SVM, and DT techniques to 96.70% accuracy(2019). In the last papers, they utilized machine learning to predict following attacks like MPS, MPG, and EID with the use of ANN, DT, RT, and XGBooost techniques. The accuracy they get is 98.60%(2021).

9. Proposed Research Methodology:

The number of devices that are connected to the internet is exceed 30 million which is recorded in 2020. So cyber-attacks are frequently increasing because the worlds move towards online. Many industries are seriously affected by a cyberattack that's why they hesitate to adopt the technology of the internet of everything(IOE).

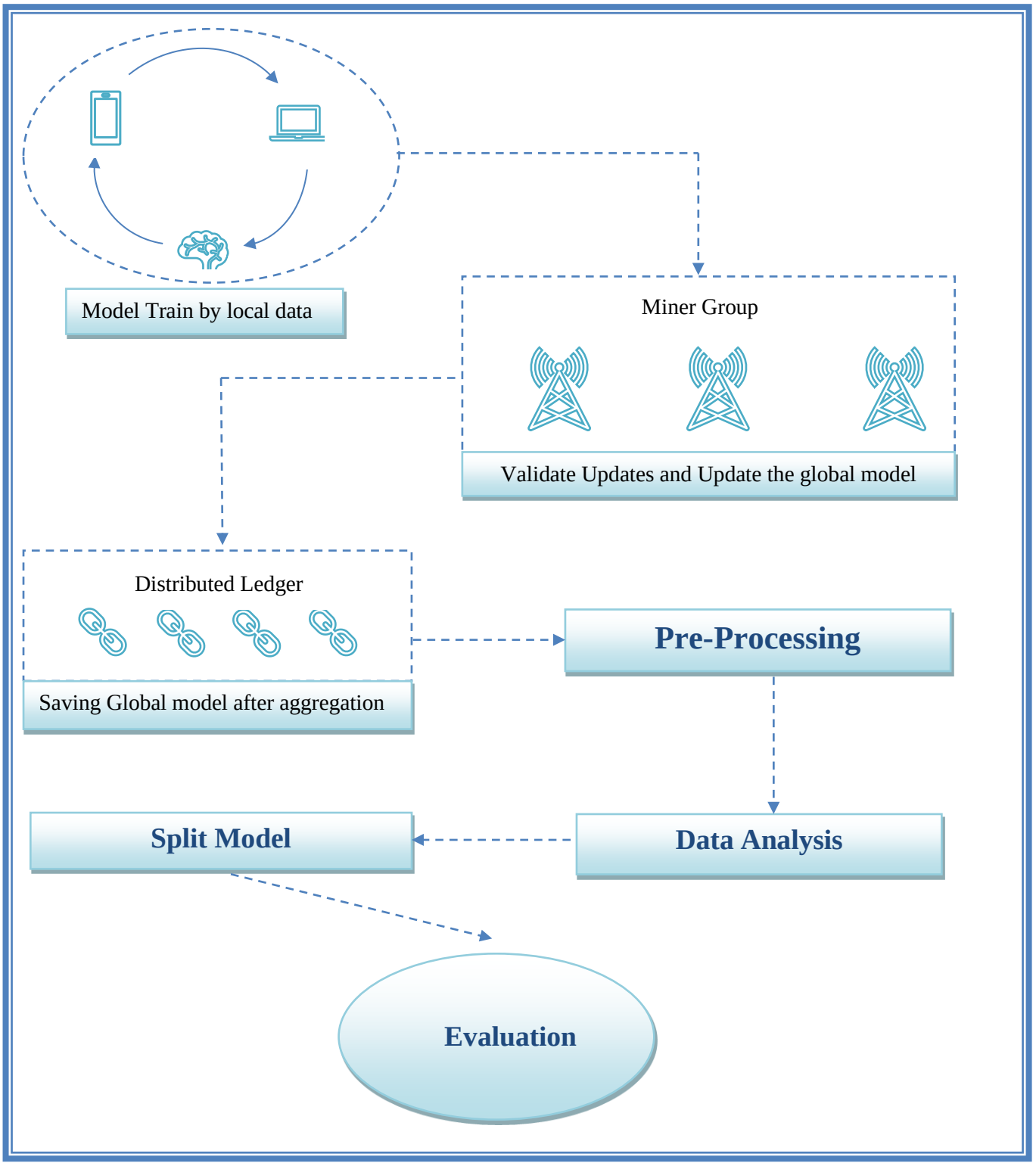
Computational federated learning solves the maximum number of threats of cyber attacks and makes an efficient model with the help of a central server. But, FML also has some problems with privacy preservation and security. There are the following issues with FML which we discuss above.

Cyber attacks on banks are increasing this day. The attacker gets all the money from the person's bank account after accessing the account data. The data can be accessed by a phishing attack. Phishing is a type of social engineering assault that is frequently used to obtain sensitive information from users, such as login credentials and credit card details. The victim receives a fake email from the attacker. The email looks real when the user clicks on it attacker access all the private data like a credit card number. So, that's how phishing attacks happen.

We proposed a model which helps organizations to work on IoT devices without any threat. A model is created with the involvement of blockChain-based federated learning. The model can predict future cyberattacks and make users' data save from the attacker.

Data is collected from different IoT devices and uses different parameters like replacing missing values, normalizing, and filtering examples for pre-processing. Pre-processing makes the dataset more efficient, reduces the duplicating values, and removes or replaces missing values.

Then apply a linear regression algorithm for data analysis. Data analysis collects raw data and converts it into useful information for decision-making by users. Then the next step is model building. We apply the FUBCFL method for model training. This method preserves the privacy of the user. This method has several functionalities, a model can be trained by the node itself and data are immutable it can not be changed and data can be traced by its source by using the special structure in the blockchain.



9.1 Proposed Research Methodology

Firstly, the node trains the model and verifies it, and then it makes the block in the distribution ledger which works as a blockchain. The distribution ledger contains the updated local models, global updated models, and data which produce during the learning process. After aggregation, the distribution ledger provides updated models to the nodes according to their contribution.

The next step is to split the data into training and testing. 70% of data is set for training and 30% of data is set for testing. The 30% data is tested on the training model. The algorithm we use for

evaluation is the decision tree. The Decision tree is a supervised learning algorithm. It is one of the most famous algorithms in today's era. In this algorithm, data is split into two or more homogenous sets based on independent variables. The accuracy of this model is more than the prevision model which was trained without blockchain.

10. Summary:

Federated learning is an advanced approach to training the models. There are a lot of functionalities that can help organizations work on the internet without any threat. The main issue federated learning solve is data privacy. Privacy is the main concern in today's era because lots of devices are attached to the internet and personal information can easily steal by a third party. Federated learning utilizes to train the model without exposing its actual data. But, federated learning also has some flaws in privacy preservation. Like sometimes nodes train models by applying unnecessary data which can not be useful for training models and parameters are transferred to the central server so there is a chance it may face a cyberattack during the data communication. All these issues can be solved by blockchain. Blockchain stores data in blocks form data that can not be changed its immutable and nodes can also participate in the aggregation process. Blockchain preserves the privacy of the user and makes an efficient model which can not be stolen by a third party.

References:

[1] **Nguyen Truong a,∗ , Kai Sun.:** Privacy preservation in federated learning: An insightful survey from the GDPR perspective. Computers & Security, 2021

[2]. **Aishwarya Srinivasan**, Difference between distributed learning versus federated learning algorithms, By **KDnuggets** on November 19, 2021.

[3]. **Dulari Bhatt**, Privacy-Preserving in Machine Learning (PPML), by analytics Vidhya on **February 3, 2022**

[4]. **G.NIKHITA REDDY1, G.J.UGANDER REDDY2**, A STUDY OF CYBER SECURITY CHALLENGES AND ITS EMERGING TRENDS ON LATEST TECHNOLOGIES, in 2014

[5]. **Mrs. Aswini Seth, Mr. Sachin Bhosale, Mr. Farish Kurupkar**, Research Paper on Cyber Security by ReserchGate in June 2021

[6]. **Mamoun Alazab, Sawarna Priya, Praveen Kumar Reddy**, Federated Learning for Cybersecurity: Concepts, Challenges, and Future Directions by ReserchGate in Oct 2021

[7]. **C. Ma, J. Li, M. Ding, H.H. Yang, F. Shu, T.Q. Quek, *et al,*** On safeguarding privacy and security in the framework of federated learning, IEEE Netw, 34 (4) (2020).

[8]. **Li X, Huang K, Yang W, Wang S, Zhang Z.** On the convergence of FedAvg on non-IID data. 2020.

[9]. **S. Luo, X. Chen, Q. Wu, Z. Zhou, S. Yu**, HFEL: joint edge association and resource allocation for cost-efficient hierarchical federated edge learning, IEEE Trans Wirel Commun, 19 (10) (2020).

[10]. **M. Chen, H.V. Poor, W. Saad, S. Cui,** Convergence time optimization for federated learning over wireless networks, IEEE Trans Wirel Commun, 20 (4) (2021).

[11]. **H.H. Yang, Z. Liu, T.Q.S. Quek, H.V. Poor**, Scheduling policies for federated learning in wireless networks, IEEE Trans Commun, 68 (1) (2020).

[12]. **Khaled A, Mishchenko K, Richtárik P.** Tighter theory for local SGD on identical and heterogeneous data. In: Proceedings of International Conference on Artificial Intelligence and Statistics; 2020 Aug 26–28; online; 2020.

[13]. **K. Wei, J. Li, M. Ding, C. Ma, H.H. Yang, F. Farokhi, *et al,*** Federated learning with differential privacy: algorithms and performance analysis, IEEE Trans Inf Forensics Secur, 15 (2020).

[14]. **Chen M, Shlezinger N, Poor HV, Eldar YC, Cui S.** Communication efficient federated learning. Proc Natl Acad Sci 2021.

[15]. **M. Chen, D. Gunduz, K. Huang, W. Saad, M. Bennis, A.V. Feljan, *et al***, Distributed learning in wireless networks: recent progress and future challenges, IEEE J Sel Areas Commun, 39 (12) (2021).

[16]**. Tian Y, Zhang Z, Yang Z, Yang Q. JMSNAS**, joint model split and neural architecture search for learning over mobile edge networks. 2021.

[17]. **Tong X, Zhang Z, Wang J, Huang C, Debbah M,** Joint multi-user communication and sensing exploiting both signal and environment sparsity. IEEE J Sel Topics Signal Process. In press.

[18]. **Yang Y, Zhang Z, Yang Q,** Communication-efficient federated learning with binary neural networks. IEEE J Sel Areas Commun.

[19]. **Q. Qi, X. Chen, C. Zhong, Z. Zhang,** Integrated sensing, computation and communication in B5G cellular Internet of Things, IEEE Trans Wirel Commun, 20 (1) (2020).

[20]. **Nikita Duggal**, Top 10 Cybersecurity Trends to Watch Out For in 2022, SimpliLearn, on Apr 11, 2022

[21]. Top Ten Cybersecurity Trends, Kaspersky, on 2022

[22]. **Mohammad Samar Ansari, Vaclav Bartosb, Brian Lee,** Shallow and Deep Learning Approaches for Network Intrusion Alert Prediction, ELSEVIER on 2020.

[23]. **Nikolaos Polatidis1 · Elias Pimenidis2 · Michalis Pavlidis1 · Spyridon Papastergiou3 · Haralambos Mouratidis1,** CrossMark on May 22, 2018.

[24]. **Md Anisur Rahman, Yeslam Al-Saggaf, Tanveer Zia,** A Data Mining Framework to Predict Cyber Attack for Cyber Security, on 2020

[25]. **Ouissem Ben Fredj, Alaeddine Mihoub, Moez Krichen, Omar Cheikhrouhou, Abdelouahid Derhab,** CyberSecurity Attack Prediction: A Deep Learning Approach, on 2020.

[26]. **Adeniji Oluwashola David, Olatunji Oluwadare Oluwasola,** Zero-Day Attack Prediction with Parameter Setting Using Bi Direction Recurrent Neural Network in Cyber Security, by Academia on 3rd Mard 2020.

[27]. **Mostofa Ahsan, Rahul Gomes, Md. Minhaz Chowdhury and Kendall E. Nygard,** Enhancing Machine Learning Prediction in Cybersecurity Using Dynamic Feature Selector, by MDPI on 21 March 2021.

[28]. **Hashim Albasheer, Maheyzah Md Siraj, Azath Mubarakali, Omer Elsier Tayfour, Sayeed Salih, Mosab Hamdan, Suleman Khan, Anazida Zainal, and Sameer Kamarudeen**, Cyber-Attack Prediction Based on Network Intrusion Detection Systems for Alert Correlation Techniques: A Survey, by MDPI on 15 February 2022.

[29]. **Hamed Alqahtani, Iqbal H. Sarker, Asra Kalim, Syed Md. Minhaz Hossain, Sheikh Ikhlaq, and Sohrab Hossain,** Cyber Intrusion Detection Using Machine Learning Classification Techniques, by check for updates on 2020.

[30]. **Abdulkadir Bilen and Ahmet Bedri Özer,** Cyber-attack method and perpetrator prediction using machine learning algorithms, by PeerJ on 9 April 2021.

[31]. **Athor Subroto, and Andri Apriyana,** Cyber risk prediction through social media big data analytics and statistical machine learning, by check for updates on 2019.

[32]. **MAHMOUD ELSISI, MINH-QUANG TRAN, KARAR MAHMOUD, DIAA-ELDIN A. MANSOUR, (Senior Member, IEEE), MATTI LEHTONEN, AND MOHAMED M. F. DARWISH,** Towards Secured Online Monitoring for Digitalized GIS Against Cyber-Attacks Based on IoT and Machine Learning, by IEEE on June 3, 2021.

[33]. **Zhilin Wang, Qin Hu,** Blockchain-based Federated Learning: A Comprehensive Survey, on 5 Oct 2021.

[34] **Multi-Vector Attacks Demand Multi-Vector Protection**". MSSP Alert. July 24, 2018.

[35] **P.S. Seemma1 , S. Nandhini2, M.Sowmiya3** Department of Computer Technology, Sri Krishna Arts & Science College, Coimbatore, Vol. 7, Issue 11, November 2018.

[36]. **P. Thomassen, J. Benninger, and M. Margraf**, "Hijacking DNS subdomains via subzone registration: A case for signed zones," Open Journal of Web Technologies (OJWT), vol. 5, no. 1, pp. 6–13, 2018.

[37]. **S. M. Zia Ur Rashid, MD. Imtiaz Kamrul, Asraful Islam, Department** of Electrical and Electronic Engineering, 7-9 February 2019.

[38] **Sree Narayana Gurukulam College of Engineering Kadayiruppu P.O**, Kolenchery Ernakulam dist., Kerala. Pin 682311.

[39] **Ing. Peter Lošonczi, Ph.D. MBA MSc. The University of Security Management in Košice**, SLOVAKIA, Issue on 26 July 2019.

[40] **Nils J. Nilsson,** Robotics Laboratory Department of Computer Science Stanford University Stanford, CA 94305.

[41] CyCon X: Maximising Effects **T. Minárik, R. Jakschis, L. Lindström** (Eds.) 2018.

[42] Gopinath Palaniappana , Sangeetha Sb , Balaji Rajendrana , Sanjaya , Shubham Goyal a , Bindhumadhava 2019.

[43] An Introduction to Federated Learning and Its Analysis Manjari Ganapathy, 5-1-2021.

[44] A blockchain future for the internet of things security: a position paper **Mandrita Banerjee a, Junghee Lee a, Kim-Kwang Raymond Choo**, 2018.

[45] A Survey of Blockchain Security Issues and Challenges, 500, **Lioufeng Rd., Wufeng, Taichung** 41354, Taiwan, 2017.

[46] Federated Learning: Challenges, Methods, and Future Directions, **Tian Li Anit Kumar Sahu Carnegie Mellon University**, 2019

[47]. Blockchain-Based Federated Learning for Intelligent Control in Heavy Haul Railway**, GAOFENG HUA1, LI ZHU 1, JINSONG WU 2,3, (Senior Member, IEEE), CHUNZI SHEN1, LINYAN ZHOU1**, (Student Member, IEEE), 2020.

[48]. **Dr. Divyakant Meva**, Issues and challenges in the blockchain: A survey by Research gate in 2018

[49]. **Osvaldo Simeone,** A Very Brief Introduction to Machine Learning With Applications to Communication Systems in 2018.

[50]. **Ouissem Ben Fredj, Alaeddine Mihoub** CyberSecurity Attack Prediction: A Deep Learning Approach, in 2020.

[51]. **Nikhita Reddy, Ugander G j Reddy**, A Study Of Cyber Security Challenges And Its Emerging Trends On Latest Technologies in 2014.

[52]. **Wajde Baiod, Janet Light,** Journal of International Technology and Information Management In 2021.

[53]. **Zibin Zheng and Shaoan Xie, Hong-Ning Dai, Xiangping Chen**, Blockchain challenges and opportunities: a survey in 2018

[54]. **Hemlata Kohad, Sunil Kumar, Asha Ambhaikar,** Scalability Issues of Blockchain Technology in 2020

[55]. **Rizwan Ahmad, Tanvir Fatima**, Importance of Cyber security and its sub-domains in 2020,

[56].**Jitendra Pande**, Introduction to cyber security In 2017

[57]. **Sachin S Bhosale,** Research Paper on Cyber Security in 2021