VISUALIZING ATTACKS IN TRANSMISSION MEDIUM USING KDD DATASET

**A PROJECT REPORT**

###### ***Submitted by***

**Aayush Kumar** (20BCY10045)

**Shrishti Srivastava** (20BCY10133)

**Avina Jain** (20BCY10135)

##### Kushagra Anand Mani (20BCY10145)

*in partial fulfillment for the award of the degree*

*of*

**BACHELOR OF TECHNOLOGY**

*in*

# COMPUTER SCIENCE AND ENGINEERING

****

**SCHOOL OF COMPUTING SCIENCE AND ENGINEERING**

**VIT BHOPAL UNIVERSITY**

**KOTHRIKALAN, SEHORE**

**MADHYA PRADESH - 466114**

##### DECEMBER 2021

**VIT BHOPAL UNIVERSITY, KOTHRIKALAN, SEHORE**

**MADHYA PRADESH – 466114**

**BONAFIDE CERTIFICATE**

Certified that this project report titled **“VISUALIZING ATTACKS IN TRANSMISSION MEDIUM USING KDD DATASET”** is the bonafide work of “**Aayush Kumar(20BCY10045), Shrishti Srivastava(20BCY10133), Avina Jain(20BCY10135), Kushagra Anand Mani(20BCY10145)**” who carried out the project work under my supervision. Certified further that to the best of my knowledge the work reported at this time does not form part of any other project/research work based on which a degree or award was conferred on an earlier occasion on this or any other candidate.

**PROGRAM CHAIR PROJECT GUIDE**

Dr. R. Rakesh Dr. Raja Soundaran

School of Computer Science and School of Computer Science and

Engineering Engineering

VIT BHOPAL UNIVERSITY VIT BHOPAL UNIVERSITY

The Project Exhibition I Examination is held on \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**ACKNOWLEDGEMENT**

First and foremost We would like to thank the Lord Almighty for His presence and immense blessings throughout the project work.

We wish to express our heartfelt gratitude to, Head of the Department, School of Computer Science and Engineering for much of his valuable support encouragement in carrying out this work.

We would like to thank our internal guide Mr. Raja Soundaran for continually guiding and actively participating in our project, giving valuable suggestions to complete the project work.

We would like to thank all the technical and teaching staff of the SCHOOL OF COMPUTER SCIENCE AND ENGINEERING who extended directly or indirectly all support.

Last, but not least, We are deeply indebted to our parents who have been the greatest support while we worked day and night for the project to make it a success.

**LIST OF ABBREVIATIONS**

|  |  |
| --- | --- |
| **Abbreviations** | **Meaning** |
| MITM Attack | Man in the middle attack |
| DDOS | Distribute Denial of Service |
| IT | Information technology |
| OT | Operational technology |
| OS | Operating system |
| TCP | Transmission Control Protocol |
| UDP | User Datagram Protocol |
| HTTP | Hyper Text Transfer Protocol |
| FTP | File Transfer Protocol |
| ICMP | Internet Control Message Protocol |
| OSI | Open Systems Interconnection |
| 4G | Generation |
| 3G | 3rd Generation |
| 5G | Generation |
| IDS | Intrusion Detection System |
| IPS | Intrusion Prevention System |
| LTE | Long Term Evolution |

**LIST OF FIGURES AND GRAPHS**

|  |  |  |
| --- | --- | --- |
| **TABLE NO.** | **TITLE** | **PAGE NO.** |
| 2.2.1 a | Security Model to prevent attacks on 4G | 14 |
| 2.2.1 b | Security Model to prevent attacks on 3G | 15 |
| 2.3 a | Privacy threats at the Application layer | 15 |
| 2.3 b | Privacy threats at the Presentation layer | 15 |
| 2.3 c | Privacy threats at the Session layer | 16 |
| 2.3 d | Privacy threats at the Transport layer | 16 |
| 2.3 e | Privacy threats at the Network layer | 16 |
| 2.3 f | Privacy threats at the Datalink layer | 16 |
| 2.3 g | Privacy threats at the Physical layer | 16 |
| 2.3 h | Raw Dataset | 17 |
| 2.3 i | Initial Graphs | 17 |
| 5.3 a | Modules imported | 23 |
| 5.5 a | Graphs Generated | 24 |
| 5.5 b | Up-link and Down-link of the connections | 25 |
| 5.6 a | Performance Graph | 26 |
| 5.6 b | Final Output (Softmax Algorithm – 98% efficiency) | 27 |
| 6.2 a | Flag on KDD Cups | 29 |
| 6.2 b. | Service on KDD Cups | 29 |
| 6.2 c | Target on KDD Cups | 29 |
| 6.2 d | Protocol type on KDD Cups | 29 |

**ABSTRACT**

The Fifth Generation (5G) of Networks aims at providing a broader connectivity all around the world. It provides more services as compared to other generations, which increased the use cases and applications. They have the ability to give speedy access with low latency. Simultaneously, they keep the option of providing slower speeds with fewer device resource requirements. Since the use of 5G networks is increasing at a high pace, hacker’s activity also increases. 5G network aims at increasing the transmission speed of data through various networking models. In our work, we mainly focus on 5G data transmission through OSI Layers, finding out security vulnerabilities and threats helps in minimizing these issues. After studying about various OSI Layers, Physical layer came out to be most vulnerable in compare with other layers. There are many attacks such as eavesdropping and data fabrication. These study makes us aware about various vulnerabilities and how we can mitigate them. For practical implementations, we had used KDD CUPS 1999 Dataset to analyze trends in various attacks in different protocols. We used LTE networks to analyze the trends and hence implementing those results for 5G networks, as 5G is not yet fully developed in India. After applying various Machine Learning Algorithms mainly KNN and Softmax, maximum attacks through TCP, UDP and ICMP mainly by means of HTTP protocol. Trojans, Malware, Rat Attack and around 10% Phishing attacks are the most common attacks in TCP and UDP. Softmax Algorithm provides more efficiency in analyzing the results hence giving 98% efficiency.

**TABLE OF CONTENTS1**

|  |  |  |
| --- | --- | --- |
| **CHAPTER NO.** | **TITLE** | **PAGE NO.** |
|  | List of Abbreviations  List of Figures and Graphs  Abstract | iv  v  vii |
| 1 | **CHAPTER-1:**  **PROJECT DESCRIPTION AND OUTLINE** Introduction 1.2 Motivation for the work  1.3 About Introduction to the project  1.4 Problem Statement  1.5 Objective of the work  1.6 Summary | 11  11  12  12  12  13 |
| 2 | **CHAPTER-2:**  **RELATED WORK INVESTIGATION**  2.1 Introduction  2.2 Existing Approaches/Methods  2.2.1 4G main threats and risks  2.2.2 3G main threats and risks  2.3 Observations from investigation  2.4 Summary | 13  14  14  15  18 |
| 3 | **CHAPTER-3:**  **REQUIREMENT ARTIFACTS**  3.1 Introduction  3.2 Hardware and Software requirements  3.3 Specific Project requirements  3.4 Summary | 18  18  19 |
| 4 | **CHAPTER-4:**  **DESIGN METHODOLOGY AND ITS NOVELTY**  4.1 Methodology  4.2 Modules Used | 20  21 |
| 5 | **CHAPTER-5:**  **TECHNICAL IMPLEMENTATION & ANALYSIS**  5.1Outline  5.2 Technical coding and code solutions  5.3 Working Layout of Forms  5.4 Prototype Submission  5.5 Test and validation  5.6 Performance Analysis (Graphs/Charts)  5.7 Summary | 22  22  23  23  24  26  27 |
| 6 | **CHAPTER-6:**  **PROJECT OUTCOME AND APPLICABILITY**  6.1Outline  6.2 Significant project outcomes  6.3 Project applicability on Real-world applications  6.4 Inference | 28  28  29  30 |
| 7 | **CHAPTER-7:**  **CONCLUSIONS AND RECOMMENDATION**  7.1Outline  7.2 Limitation/Constraints of the System  7.3 Future Enhancements  7.4 Inference | 30  31  31  31 |
|  | References | 31 |

**1. PROJECT DESCRIPTION AND OUTLINE**

**1.1 INTRODUCTION**

5G technology – The generation which is getting more into the grounds of networking world. The aim of 5G technology is to provide very high data rates and higher coverage through dense base station. To achieve this aim, various services, networking models, Machine Learning Algorithms will be required.

Wireless Communication whether 1G, 2G, 3G or 4G all the generations are prone to security issues. 1G wireless network was observed with illegal cloning in mobile phones. For 2G wireless networks, message spamming was the most common vulnerability attack. In case of 3G wireless networks, IP based communication opens the door for various security challenges. With growing generations, security issues and vulnerabilities are increasing at a much faster rate. With the upcoming demand of 5G networks, it is very important to study their security challenges and try to overcome them.

5G wireless network aims at increasing the transmission rate of data. As we already aware that OSI Layers or other networking models help in transmitting data. Hence understanding various vulnerabilities in OSI Layers help to overcome these issues.

Intrusion Detection System – a very known and popular way of detecting the vulnerabilities and various trends. Applying Machine Learning Algorithms on datasets helps us to find various trends in vulnerabilities.

**1.2 MOTIVATION FOR THE WORK**

As 5G technology is an updated version of 4G technology which only aims at increasing transmission speed of data. OSI Layers are responsible for transmitting data from sender to receiver hence understanding about various attacks and vulnerabilities in different OSI Layers helps to transmit data with more security. These will help in more secure transfer of data and also make people aware about various attacks.

IP and UDP are responsible for transfer of data. Applying Intrusion Detection System on such dataset help to recognize various figures and trends regarding vulnerabilities so that we can get an idea about various attacks. These will help in minimizing data loss and increases the security in data transmission.

**1.3 INTRODUCTION TO THE PROJECT**

Our project is basically a theoretical research work on finding out various vulnerabilities, attacks and security challenges associated with different OSI Layers in 5G transmission of data. In this we analyze all OSI layers, the protocols associated with them in data transmission. We also discussed about various protocols associated with prevention of various attacks.

The other part of our project is practical analysis of dataset. We use KDD CUPS 1999 dataset and apply various machine learning algorithms to detect various trends. Here we will be doing Intrusion Detection System so detect various figures in dataset.

**1.4 PROBLEM STATEMENT**

With the advent of 5G technology, the risks of data loss increased. To overcome data loss preventive measure need to be taken so that people can securely transfer their valuable information. Hence, we try to find out all possible attacks and vulnerabilities associated with each OSI layers.

Since 5G is not yet fully available in India, we are using 4G config files for Intrusion Detection System using KDD CUPS 1999 dataset so that we can get an idea about various vulnerability trends in 4G network which we can apply in 5G networks as 5G is just an updated version of 4G technology. Using various Machine Learning Algorithms, we can detect the vulnerabilities.

**1.5 OBJECTIVE**

Our objective is to find out as many vulnerabilities in OSI Layers so that we can overcome or mitigate them as soon as possible. Since 5G is not yet fully developed in India, analyzing these vulnerabilities help in mitigating them before 5G technology gets developed and adopted by people.

By analyzing 4G config files, we can get an idea about attacks that can be associated with 5G technology as well. By applying Intrusion Detection System, we can analyze trends and figures of various attacks.

Hence our main objective is to make the transmission secure so that various private and public organization can share their information without data theft and data loss.

**1.6 SUMMARY**

5G is a growing technology which supports various other technologies. The security of 5G technology is a major concern as data loss and data theft are very common these days. Transmission of data through secure medium is most important concern of private and public organization hence working in these field is very important. Data transmission follows various models out of which OSI Layers are most common, hence studying about their privacy and security is need. Also without practical implementations nothing can be finalized. Intrusion Detection System on 4G config files gives us an idea about vulnerability trends which we can apply in 5G environment to overcome security challenges.

**2. RELATED WORK INVESTIGATION**

**2.1 INTRODUCTION**

As technology evolves, new updates arrive every second, bringing with them potential hazards that might cause harm to the system employing the technology and lead to its collapse. Also the data transmission take places through various layers and at each layer there are some risks involved. So we are also analyzing different layers in OSI Model and and what are the risks each layers have and which layer has the maximum issues and what have been done previously to solve the issue and what can be done in future to make it more secure for the user.

We are attempting to forecast the weaknesses and security difficulties in 5G technology by doing research on OSI layers and performing Intrusion Detection on the KDD 1999 DATASET. Because 5G has yet to be launched in India, the config file is not yet accessible, although 5G is an improved version of 4G with a fast transmission rate.

So after searching for some dataset of 4G config file and 3G config file we are able to perform the intrusion detection on 4G config file and with the help of that we can analyze what are the existing vulnerabilities and what can be the risks in 5G and what are the steps we can take to reduce the threats. To safeguard IDI's internal servers, we will employ the tiered security approach. In the case of an intrusion, an intrusion detection system (IDS) will be put up to send out alerts and log all connections. An intrusion prevention system (IPS) will be installed to prevent the identified intrusions, and MAC address filtering will be used to prohibit or allow connections depending on the MAC address or physical address of each computer.

**2.2 EXISTING APPROACHES AND METHODS**

**2.2.1 4G**

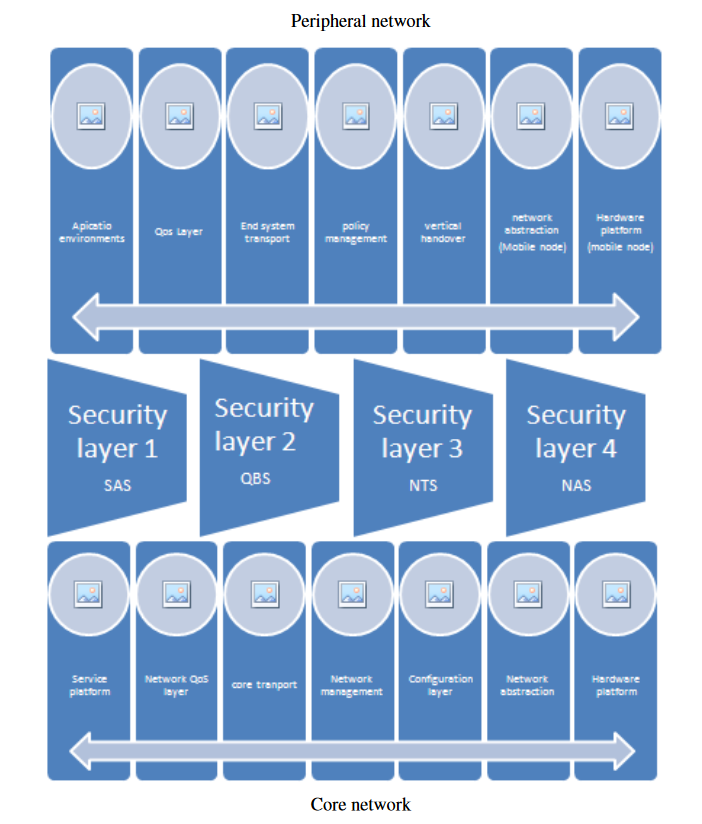
Main Threats and Risks **-** 1. Physical layer issues

2. WiMAX-MAC-Layer security issues

3. Denial of service security issues

4. Wi-Fi security issues

PROPOSED SOLUTION OF 4G Security –



*Fig 2.2.1 a Security Model to prevent attacks on 4G*

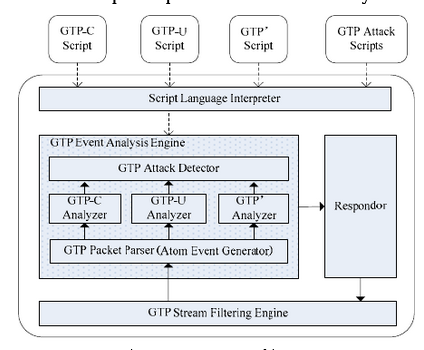
**2.2.2 3G**

Main Threats and Risks - 1. Protocol abnormal attack.

2. Infrastructure attack (GTP Deception).

3. Resource consumption attack**.**

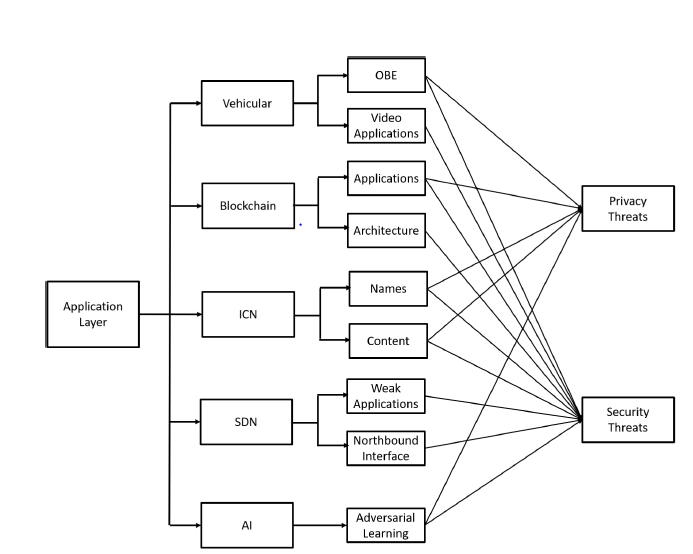
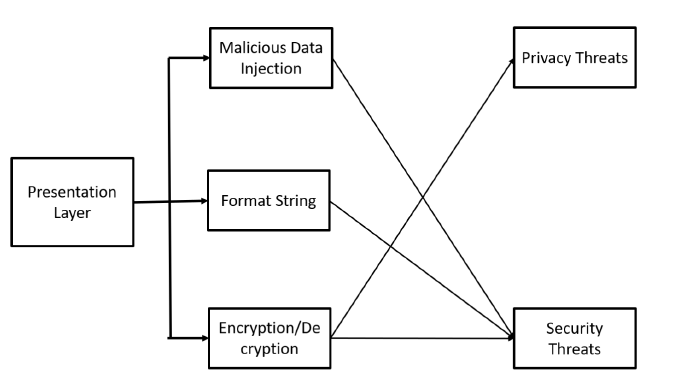
PROPOSED SOLUTION OF 3G Security –



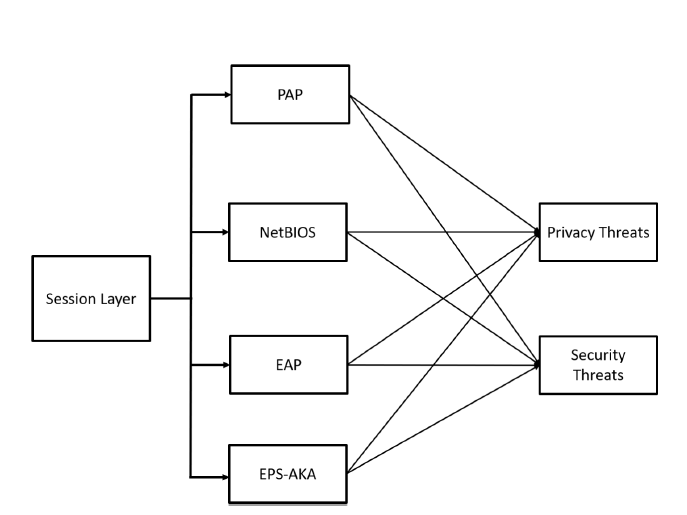
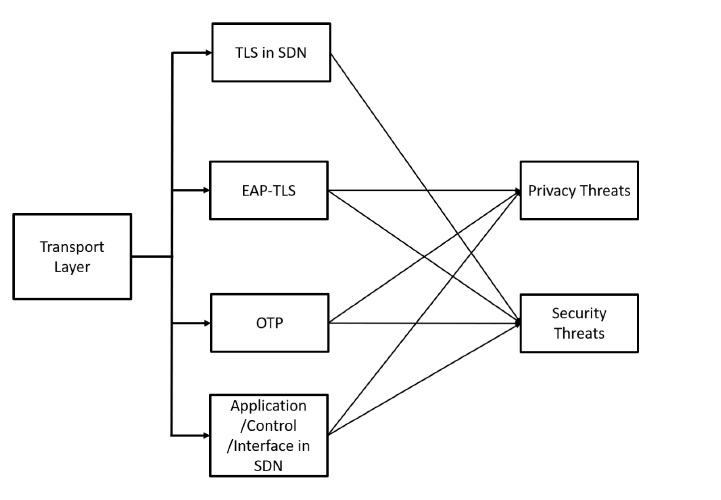
*fig 2.2.2 a Security Model to prevent attacks on 3G*

**2.3 OBSERVATION FROM INVESTIGATION**

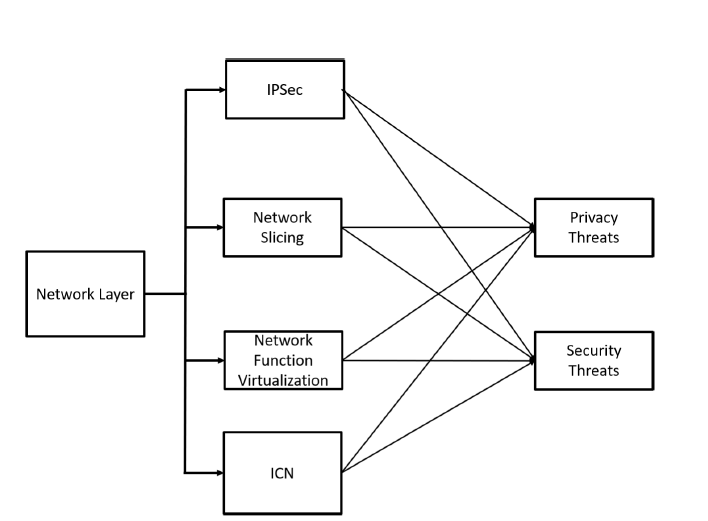
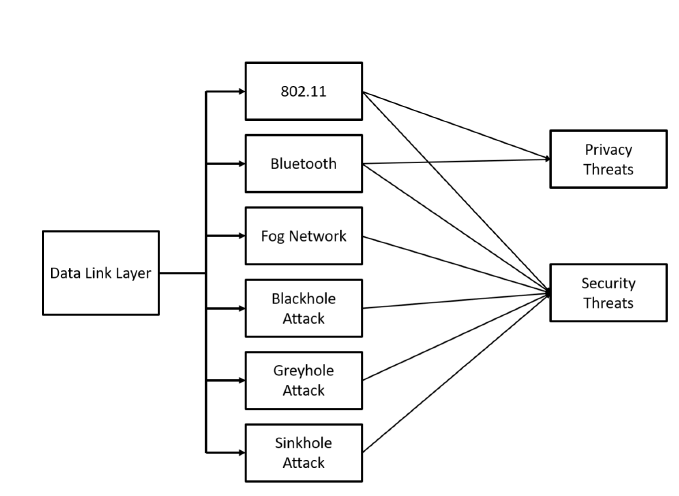
The observations we get from studying about security issues in OSI Layers in 5G transmission are-

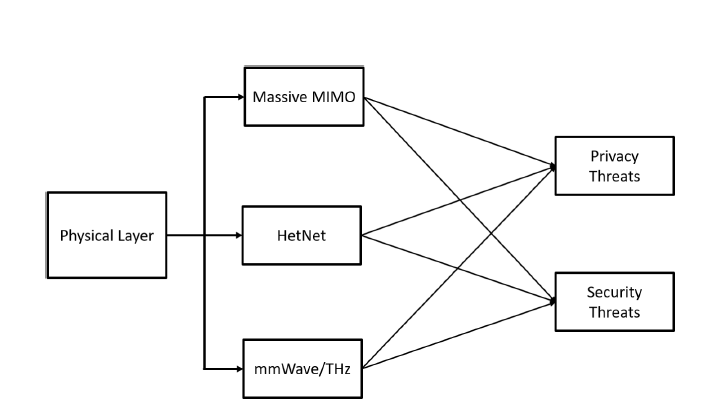
*fig 2.3 a Privacy threats at the Application**layer fig 2.3 b Privacy threats Presentation Layyer*

*fig 2.3 c Privacy threats at the Session layer fig 2.3 d Privacy threats at Transport layer*

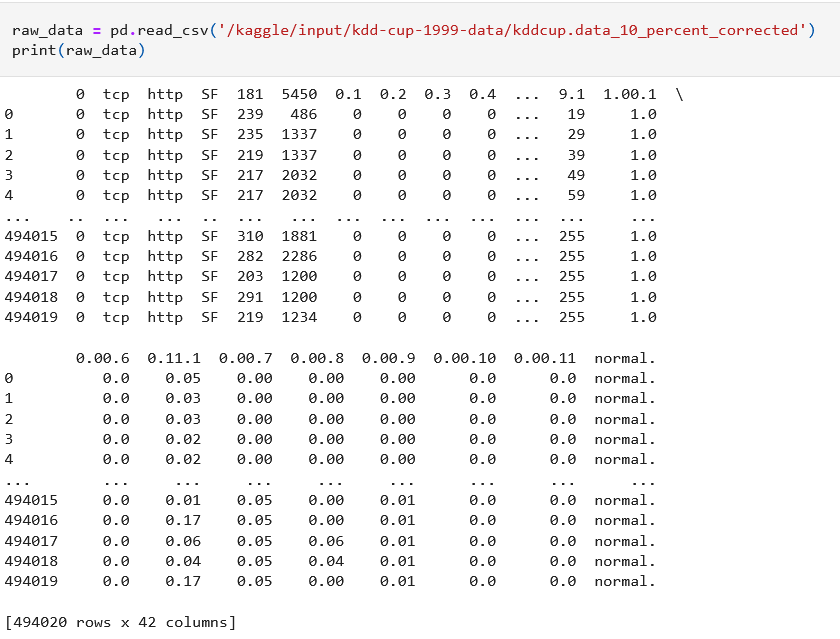
*fig 2.3 e Privacy threats at the Network layer fig 2.3 f Privacy threats at the Datalink layer*



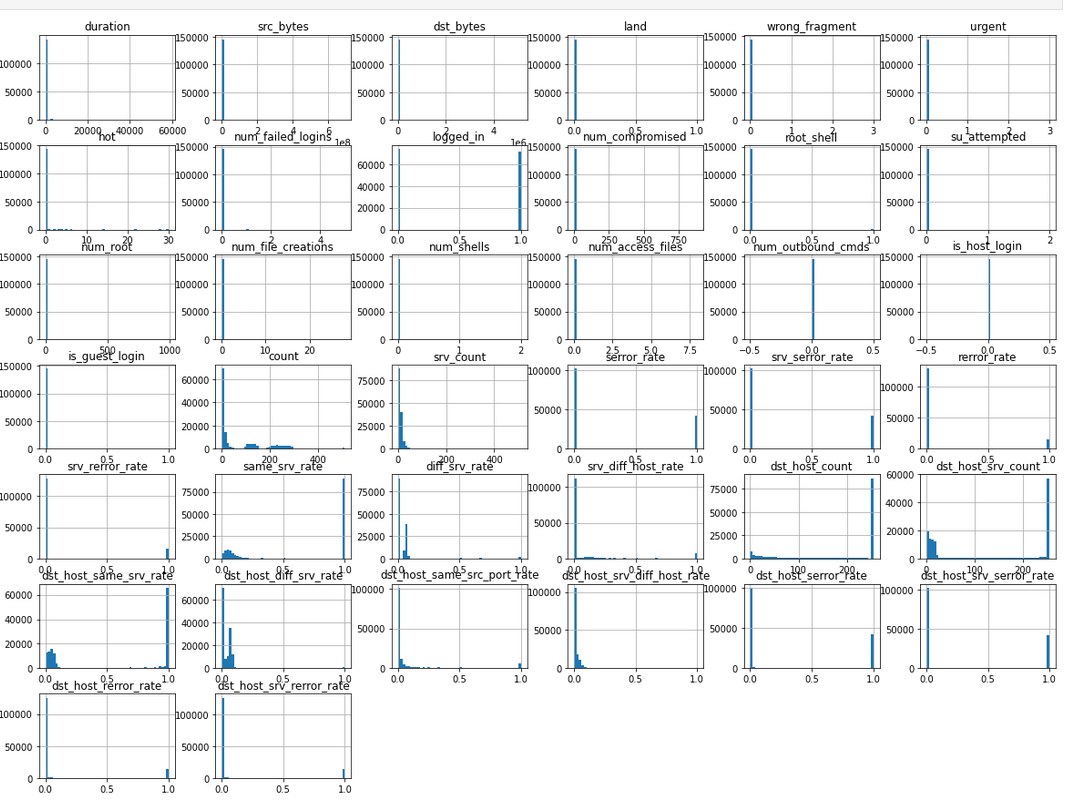
*fig 2.3 g Privacy threats at the Physical layer.*

**INTRUSION DETECTION DATA –**

Importing and analyzing the given “kddcup.data\_10\_percent\_corrected dataset -



*fig 2.3 h Raw Dataset*



*fig 2.3 i Initial Graphs*

**2.4 SUMMARY**

5G data networks are currently supporting linked real-world applications and have the potential to be the backbone of a future human civilization that is always connected. As a result, there are certain difficult design, implementation, and deployment difficulties affecting all aspects of 5G networks. Among these, timely identification of any unauthorized attempt is critical, particularly in the setting of a business information network. In addition, a real-time intrusion detection device principally based entirely on the usage of machine learning algorithm. Also, because data transmission is a layered process, each layer employs a variety of protocols, each with its own set of vulnerabilities. As a result of our investigation, we discovered that the Physical layer is the most important layer in data transmission, with numerous potential threats that can be mitigated using a variety of methods and algorithms.

**3. REQUIREMENT ARTIFACTS**

**3.1 INTRODUCTION**

For any project or research work, there is need of various tools, software, and other devices to complete the work. We also require various components to complete our research and practical work. For Intrusion Detection System, we need various components to get the work done.

**3.2 HARDWARE AND SOFTWARE REQUIREMENTS**

Hardware Requirement for our project work is –

* GPU

Software Requirement for our project work is –

* Virtual Environment (E.g. Kaggle)

**3.3 SPECIFIC PROJECT REQUIREMENTS**

**1. KDD data set**

NSL-KDD is a record set recommended to address a variety of flaws with the KDD 1999 records set. Despite the fact that this new edition of the KDD records set still has some of the issues raised by McHugh and will not be an ideal consultant of current actual networks, we believe it can be used as a powerful benchmark record set to assist researchers examine various intrusion detection methods due to the scarcity of public records units for network-primarily based totally IDSs. Furthermore, the enormous number of information contained inside the NSL-KDD tech and test modules is reasonable. This benefit allows the experiments to be done on the entire set without the need to select a tiny section at random. As a result, the evaluation conclusions of numerous researches can be consistent and comparable.

**2. Softmax Algorithm**

The Softmax regression is a type of logistic regression that converts an entry price into a vector of values that follows a probability distribution with a general sum of one.

The output values are in the range [0, 1], which is exceptional because we can avoid using binary classes and include as many lessons or dimensions as we like in our neural community model.

SoftMax is frequently referred to as a multinomial logistic regression because of this, when educating a data set, the feature is frequently utilized to compute losses that can be projected.

Softmax regression has been used in discriminative modes, as well as Cross-Entropy and Noise Contrastive Estimation.

These are the most useful among a variety of tactics that aim to improve performance. These are only a few of the many techniques that have been proposed to improve today's schooling system in order to increase the likelihood of correctly guessing the suitable phrase or sentence.

**3. KNN Classifier**

The algorithms calculate the distances between this and all other K numbers of datapoints inside the dataset near the preliminary factor for a given statistics factor within the set and vote for the class with the highest frequency. Euclidean distance is commonly used as a measure of distance. As a result, the final consequent version is merely a collection of categorize facts arranged in a space. This set of rules is widely recognized for a variety of applications, including genetics, forecasting, and so on. In this example, the set of rules is good, while the greater functions are free, and out suggests SVM. It is a fact that KNN reduces overfitting.

**4. DESIGN METHODOLOGY AND ITS NOVELTY**

**4.1 METHODOLOY AND GOAL**

The methodologies used in our project are-

* **Neural Network**

Neural networks help in the development of financial procedures such as time-collection forecasting, algorithmic trading, securities categorization, credit score risk modelling, and the construction of customized signals and fee derivatives. In addition to the neural community of the human brain, there is a neural community. A "neuron" is a mathematical feature in a neural community that gathers and categorizes records based on a specified architecture. The community resembles statistical procedures like curve fitting and regression analysis. Different methodologies of technical evaluation cannot discern dispersed nonlinear interdependencies and styles that networks can.

There are three key components: an enter later, a processing layer, and an output layer. The inputs can be weighted primarily depending on several factors. Within the processing layer, which is concealed from view, there are nodes and connections between those nodes that are meant to be equivalent to neurons and synapses in an animal brain.

* **SoftMax Algorithm**

SoftMax turns a vector of K real values into a vector of K actual values that add up to one. The SoftMax translates the entering values, which might be positive, negative, zero, or multiples of one, to values between zero and one, allowing them to be interpreted as probabilities. SoftMax translates a little or negative input to a small probability, and a high input to a huge chance, but it will generally remain between zero and one.

The SoftMax feature is also known as the soft argmax feature or multi-elegance logistic regression.

This is because SoftMax is an extension of logistic regression that may be used for multi-elegance classification, and its components are comparable to the sigmoid feature in logistic regression. When the lessons are mutually exclusive, the SoftMax feature can be quite useful in a classifier.

Many multi-layer neural networks feature a penultimate layer that generates real-valued evaluations that are challenging to scale and control. The SoftMax may be particularly beneficial in this circumstance since it converts the ratings to a normalized probability distribution that can be shown to a person or used as input to multiple systems. As a result, adding a SoftMax feature to the neural network's last layer is extremely common.

* **KNN ALGORITHM**

The k-nearest neighbors (KNN) set of criteria is a statistics category approach for evaluating the likelihood that a statistics factor will become a member of one organization or another largely depending entirely on which organization the statistics factors closest to it belong to. The k-nearest neighbor set of rules is a type of supervised system that learns a set of rules to solve category and regression problems. However, it is most commonly employed for category difficulties.

**4.2 MODULES USED**

* **NumPy:** For including assist for massive, multi-dimensional arrays and matrices, alongside a massive series of high-stage mathematical features to perform on those arrays.
* **Pandas:** Panda’s module runs on pinnacle of NumPy and it is far popularly used for recording technological knowledge and recording analytics. NumPy is a low-stage records shape that helps multi-dimensional arrays and a wide variety of mathematical array operations. Pandas have a higher-stage interface. It additionally offers streamlined alignment of tabular records and effective time collection functionality.
* **Matplot.lib.pyplot:** Matplotlib is a complete library for developing static, animated, and interactive visualizations in Python. Matplotlib makes clean matters clean and difficult matters possible. Create a booklet with first-class plots. Make interactive figures which can zoom, plan, update.
* **Seaborn:** Seaborn is an open-supply Python library constructed on pinnacle of matplotlib. It is used for records visualization and exploratory records analysis. Seaborn works without problems with data frames and the Pandas library. The graphs created also can be custom designed without problems.
* **Time:** Python's time module offers a feature for buying nearby time from the range of seconds elapsed because the epoch known as local time.

**5. TECHNICAL IMPLEMENTATION AND CODING**

**5.1 OUTLINE**

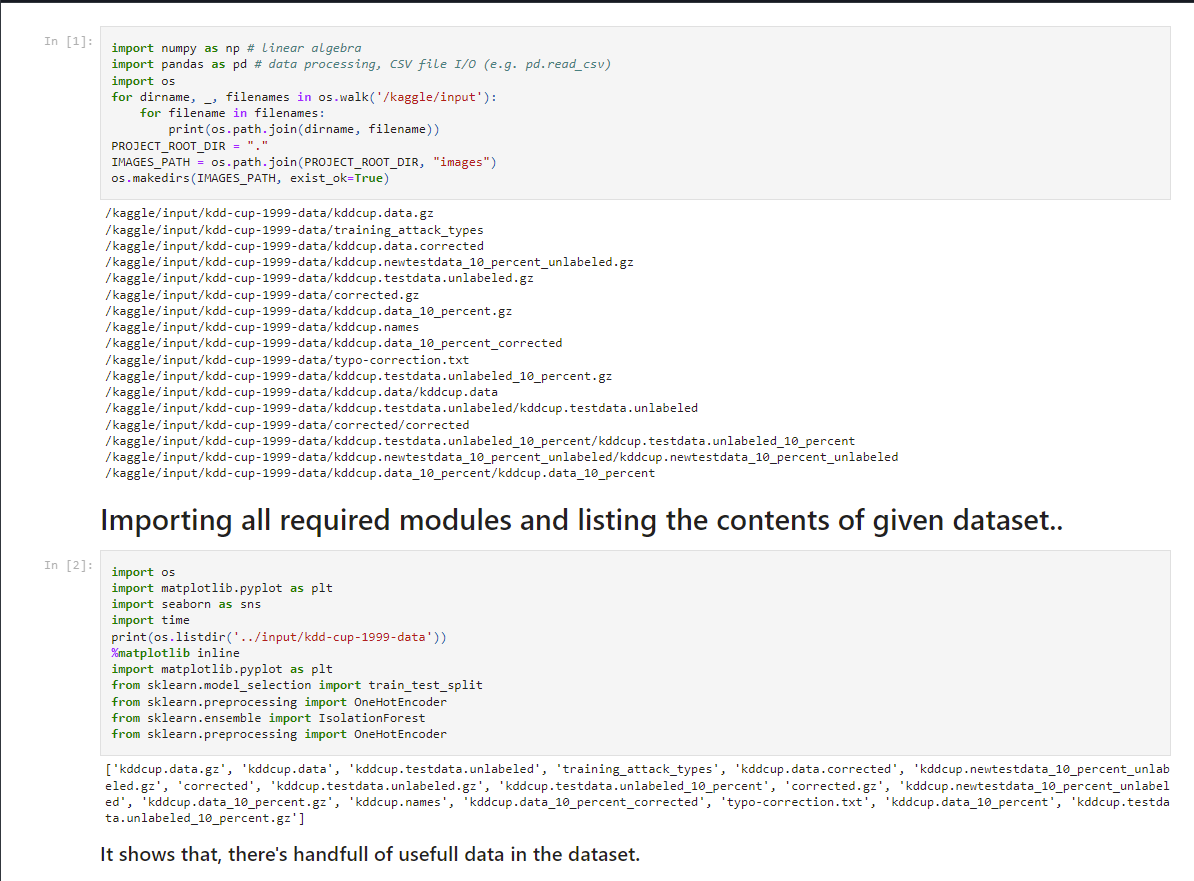
An Intrusion Detection System (IDS) is a framework that screens network traffic for dubious action and issues ready when such action is recognized. A product application filters an organization or framework for vindictive movement or strategy infringement. Any pernicious undertaking or infringement is frequently answered to a head or assembled halfway using security data and occasion the board (SIEM) framework. A SIEM framework incorporates yields from various sources and uses alert sifting calculations to recognize hazardous ways of behaving from misleading problems. Inspite of the fact that interruption identification frameworks screen networks for possibly destructive exercises, they are additionally inclined to phony problems. Therefore, undertakings should adjust their IDS gadgets when they first send them.

**5.2 TECHNICAL CODING AND IMPLEMENTATION**

For solving this dataset, we have used the "KDD 1999 intrusion detection dataset" and the "Kaggle ML environment for implementing ML algorithms". For this dataset, we have used the pipeline method and two algorithms, namely KNN and Softmax. While implementing, we came to know that Softmax is more efficient and less time-consuming, so we worked on the Softmax algorithm and got approximately 98% accurate results.

**5.3 WORKING LAYOUT OF FORMS**

For practical implementations and for applying various machine learning algorithms, modules and methods are required. These modules need to imported first so that we can apply methods and algorithms to solve the given dataset.



*fig 5.3 a Modules imported*

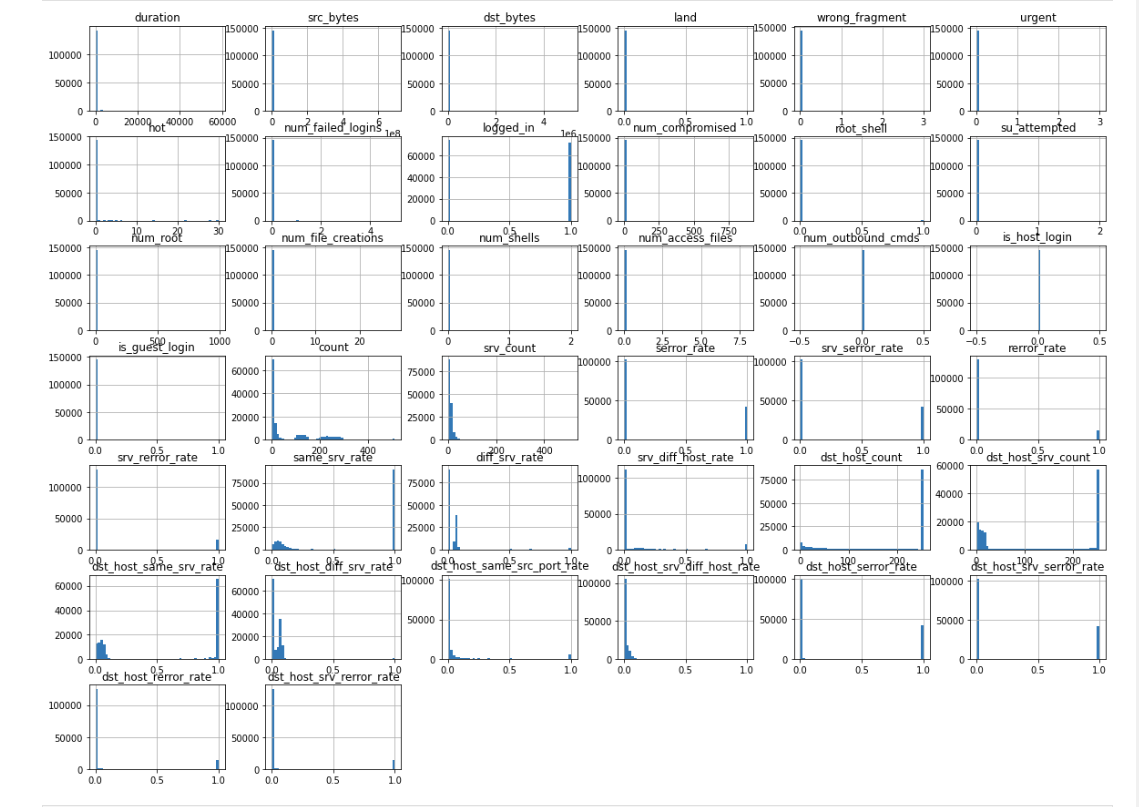
**5.4 PROTOTYPE SUBMISSION**

For Prototype and code please visit the given link of GitHub and Kaggle profile

[KDD-99-Intrusion-detection-solution/kdd-cup-99-dataset-visualisation-pipeline.ipynb at main · aayushkumar20/KDD-99-Intrusion-detection-solution (github.com)](https://github.com/aayushkumar20/KDD-99-Intrusion-detection-solution/blob/main/using%20pipelines/kdd-cup-99-dataset-visualisation-pipeline.ipynb)

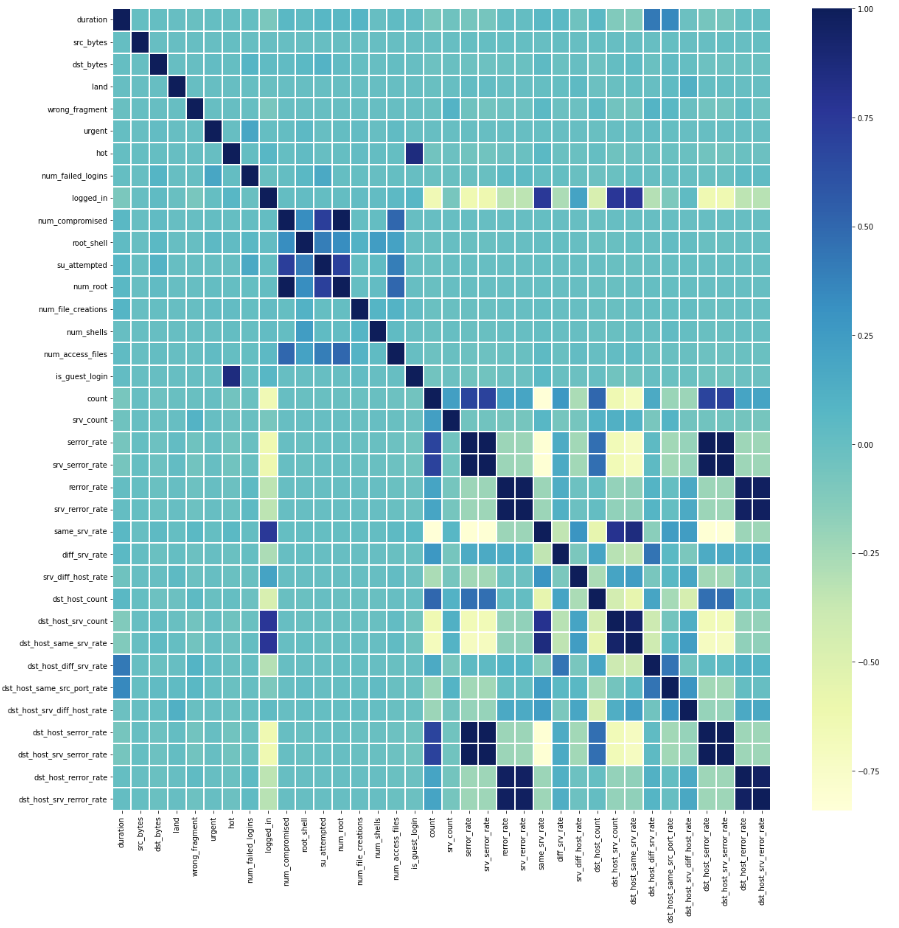
<https://www.kaggle.com/code/aayushkumar20bcy/kdd-cup-99-dataset-visualisation-pipeline>

**5.5 TEST AND VALIDATION**



*fig 5.5 a Graphs Generated*

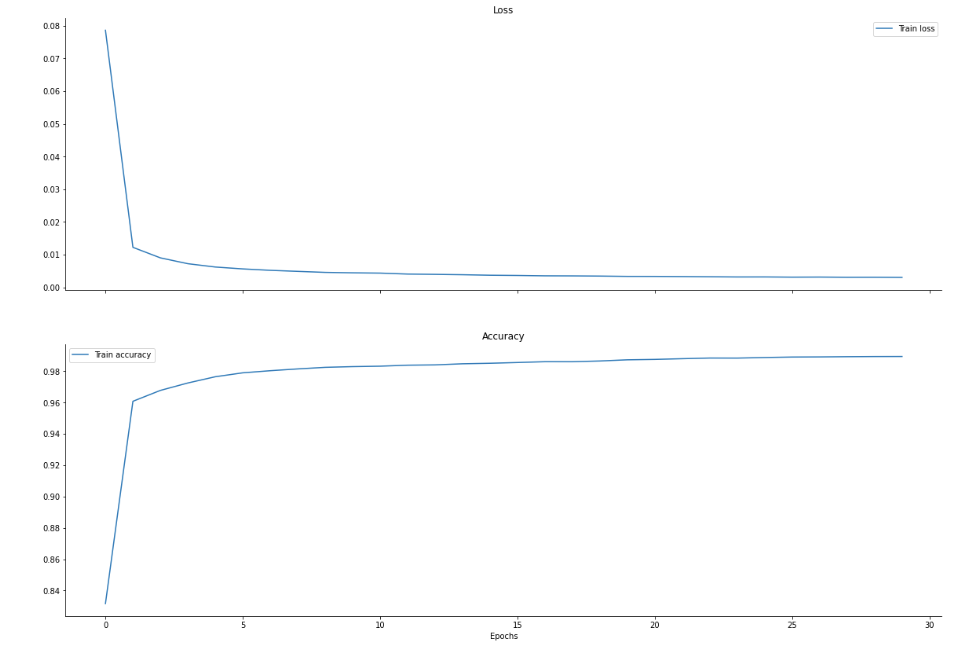
This graph shows various aspects of the up-link and down-link of the connections across various networks.



*fig 5.5 b Up-link and Down-link of the connections*

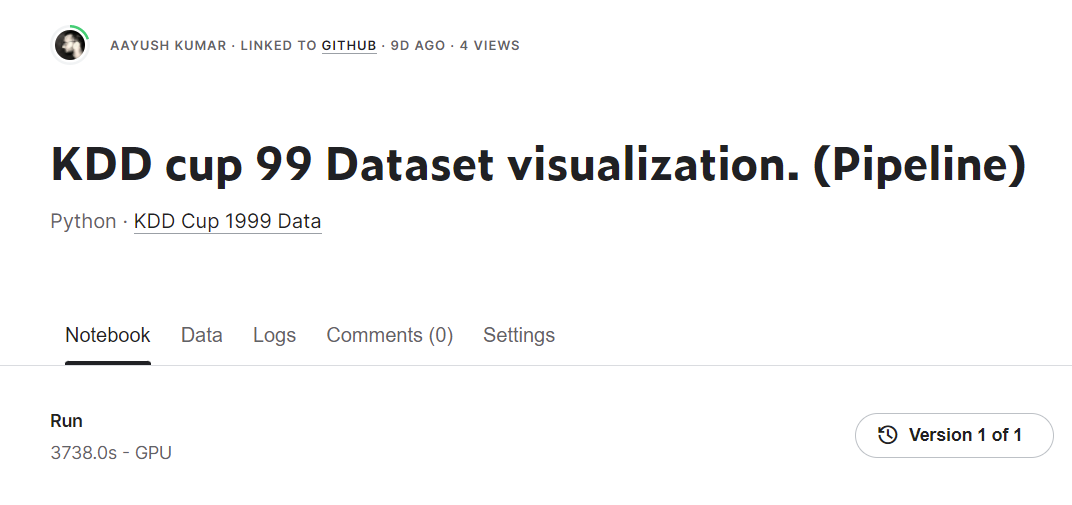
**5.6 PERFORMANE GRAPHS**

Following graphs shows our Performance in terms of percentage and plot.





*fig 5.6 a Performance Graph*





*fig 5.6 b Final Output*

**5.7 SUMMARY**

The dataset "KDD CUP 99" was used in the 3rd International Knowledge Discovery and Data Mining Tool Competition, which was conducted concurrently with the 5th KDD99. It is derived from the DARPA98 dataset after data extraction and pre-processing. Over a 9-week period, the DARPA98 dataset collects TCP dump network connections and system test data from a US Air Force LAN simulation facility. The raw data collected by TCP dumps is split into two sections: There are around 5,000,000 network connections in the 7-week training data. The remaining two weeks of test data comprise around 2,000,000 network connection records. We extracted the KDD99 dataset by doing feature analysis and data pre-processing on data from the DARPA98 dataset.

The KDD99 dataset contains 5 million records in total, as well as 10% of the training and testing subsets. Each network connection in the data set is labelled as normal (normal) or abnormal (attack), with the abnormal type classified into four categories and a total of 39 attack types, comprising 22 types of attacks that occur in the training set and 17 unknown types. The test set contains attack patterns.

* DOS, or denial of service assault, is one of the four exception kinds. Pingofdeath, flood syn, smurf, and so on;
* R2L, that is. illegal access from a remote server. Example of password guessing
* U2R, Unauthorized Super Local User Privilege Access. such as buffer overflow attacks;
* PROBING, as well as monitor and scan ports for example, portcan, pingsweep, and so on. The records in the KDD99 dataset are in csv format, and each connection is described by 41 characteristics plus the final label (label), for a total of 42 elements.
* TCP connection fundamental characteristics (9 items), TCP connection content characteristics (13 items), time-based network traffic statistics characteristics (9 items), and host-based network traffic statistics characteristics are the first 41 characteristics (10 items)

**6. PROJECT OUTCOME AND APPLICABILITY**

**6.1 OUTLINE**

The Fifth Generation of Communication Networks (5G) proposes a bigger variety of services than previous generations, with a greater number of use cases and applications supported. The expansion of the application domain leads to more user use and, as a result, increased hacker activity.

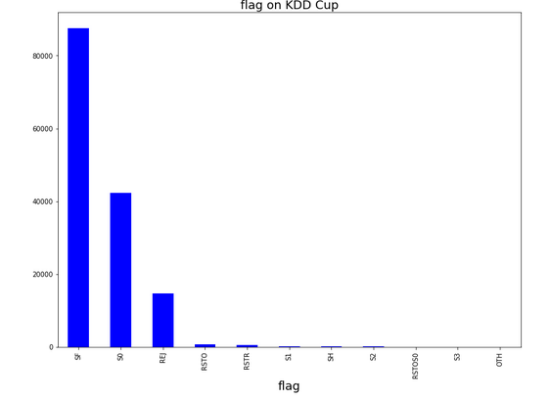
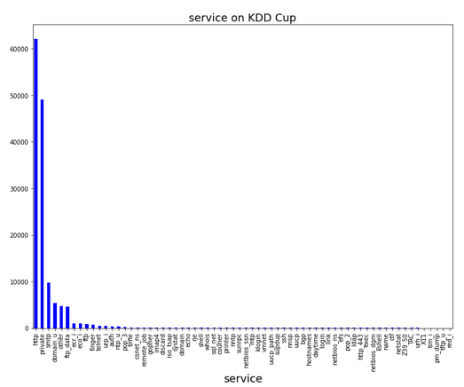
In this project, we use KDD Datasets to present an impartial overview of 5G security concerns in OSI layers, as well as current and newly suggested methods for securing the 5G transmission environment.

We are attempting to forecast the weaknesses and security difficulties in 5G technology by doing research on OSI layers and performing Intrusion Detection on the KDD 1999 DATASET. Because 5G has yet to be launched in India, the config file is not yet accessible, although 5G is an improved version of 4G with a fast transmission rate.

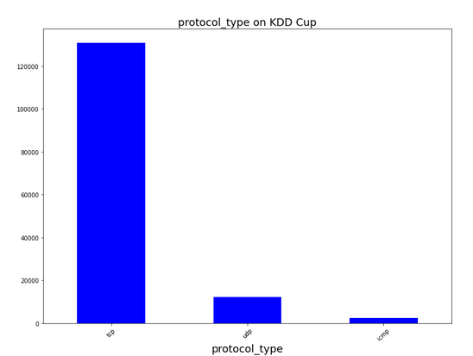
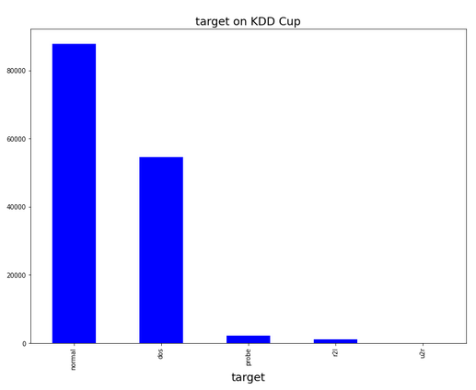
**6.2 SIGNIFICANT PROJECT OUTCOME**

After investigating data transmission in OSI layers, we observed that there are some security difficulties at each tier, with the physical layer being one of the most significant in terms of security challenges. We investigated several strategies that can be used to address these difficulties. SegWit, Public Key Cryptography, EPS-AKA and SPEKE, IPSec, firmware updating, Machine Learning IDS, RSS, Power Control, Beamforming, Clustering, and many more are examples of these approaches.

The results of trying to implement Softmax and KNN Algorithms for Intrusion Detection in LTE networks are shown in the graphs below.

** **

*fig 6.2 a Flag on KDD Cups fig 6.2 b Service on KDD Cups*

** **

*fig 6.2 c Target on KDD**Cups fig 6.2 d Protocol type**on KDD Cups*

**6.3 PROJECT APPLICABILITY ON REAL WORLD APPLICATION**

This project is applicable as soon as the 5G will launched in India. After visualizing the KDD Dataset on LTE networks through different algorithm we have seen that most data and Ransomware are transferred through

HTTP and FTP protocol. Maximum attacks happen through TCP, UDP, ICMP protocol and making them secure for future is the biggest concern. Also, the research on vulnerabilities in OSI Layer has shown us that which layer is the most vulnerable layer and what are different vulnerabilities are involved in it.

Once we will know what are the existing problems, we will try to mitigate them in our future technology and make the things easier and more secure for us.

**6.4 INFERENCE**

The Fifth Generation of Communication Networks (5G) offers a broader range of services than previous generations, with a higher number of use cases and applications supported. The extension of the application domain leads to higher user use and, as a result, increased hacker activity.

We are aiming to foresee the flaws and security challenges in 5G technology by doing research on OSI layers and performing Intrusion Detection on the KDD 1999 DATASET. Because 5G has yet to be deployed in India, the configuration file is not yet available, despite the fact that 5G is an emerging technology.

After researching data transmission in OSI layers, we discovered that each tier has various security concerns, with the physical layer being one of the most critical in terms of security challenges. We looked at a number of approaches that may be utilized to address these issues. These techniques include SegWit, Public Key Cryptography, EPS-AKA and SPEKE, IPSec, firmware update, Machine Learning IDS, RSS, Power Control, Beamforming, Clustering, and many others.

**7. CONCLUSIONS AND RECOMMENDATION**

**7.1 OUTLINE**

5G technology is the promising technology aims to increase the rate at which data gets transmitted. Billions of sensors and millions of devices will be linked by 5G technology. Overcoming security vulnerabilities is a critical challenge due to its widespread use. The expected number of connected devices expands the opportunities for hackers to attack networking at multiple tiers of the OSI layer model. We came to know that Physical layer is most vulnerable to attacks in transmission medium. It needs to be more secured as data loss may occur in physical medium. Other OSI layers are also vulnerable to various types of attacks and needs to get secured.

Nothing can be completed unless real implementations take place. The Intrusion Detection System on LTE Networks provides us with information about vulnerability patterns that we can use to address security issues in the 5G environment. We analyzed that maximum data and ransomware are transferred through FTP and TCP by means of HTTP. We also concluded that Softmax Algorithm gives maximum efficiency of 98% in detecting trends about various attacks.

**7.2 LIMITATION AND CONSTRAINTS**

Talking about limitation of our work, as our work is a research work along with practical implementation limitation is not that it cannot fulfill any particular function. Research work is said to be totally successful if it can help the people to get better understanding about the topic. When we talk about research in OSI layers, there were more attacks left out which needs to be addressed so that people become more aware. Since 5G is not yet completely developed in India, we need to be more precise with the work so that it gets fully developed we already know the domains where we need to work upon.

For Intrusion Detection System, there is no availability of 5G configuration files which we can used to analyze vulnerability trends in 5G network. Unavailability of 5G dataset is the biggest limitation of the work but we tried to figure out issues from the available dataset and config files that can be used to get an idea what are loopholes and we need to work upon.

**7.3 FUTURE ENHANCEMENTS**

In future, we are accepting to add 5G config files when it will be available so that more exact outcomes can be analyzed. We are also trying to add more and more recent case studies related to 5G networks along with challenges, the risk incorporated and how to mitigate them.

These are our future work which we are looking forward to complete and update it with our work so that it can help more and more end users and companies.

**7.4 INFERENCE**

Each new generation of mobile networks has tended to reduce information security risks. 5G network is developing technology which is setting its roots to get more and more strong. Because the primary selling point of 5G networks has been increased data rates and speed, it will be difficult to combat a wide range of threats from various points using standard or traditional protective measures. With the growing demands of 5G, hacker’s activity also increase. These problems caused considerable damage and affected the secrecy, integrity and availability of information. To overcome these activities, people and companies need to be aware and always been prepared to tackle threats if something happens.

**REFERENCES**

*1.* S. Sullivan *:5G Security Challenges and Solutions: A Review by OSI Layers*

*2.* Noman Haider, Zeeshan Baig, Muhammad Imran*: Artificial Intelligence and Machine Learning in 5G Network Security: Opportunities, advantages, and future research trends.*

3. Constatinos Kolias, Georgios Kambourakis, Angelos Stavrou and Stefanos Gritzalis: *Intrusion Detection in 802.11 Networks: Empirical Evaluation of Threats and a Public Dataset.*

4.<https://www.router-switch.com/faq/session-layer-the-functions-and-protocols.html>

5.<https://itstillworks.com/disadvantages-password-authentication-protocol-2895.html>