**MAJOR PROJECT END SEM VIVA REPORT**

**ON**

**CLOUD HYBRID NETWORK INTRUSION DETECTION**

SUBMITTED IN PARTIAL FULFILMENT FOR THE AWARD OF DEGREE OF

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE AND ENGINEERING**

Enrol No. (s) - 17103157, 17103020, 17103012  
Name of Students (s) - Aman Ahuja, Abhinav, Prashant Shekhar  
Name of supervisor (s) - Dr. Satish Chandra



**May 2021**

**Submitted in partial fulfillment of the Degree of**

**Bachelor of Technology**

**in**

**Computer Science Engineering**

**DEPARTMENT OF COMPUTER SCIENCE ENGINEERING & INFORMATION TECHNOLOGY**

**JAYPEE INSTITUTE OF INFORMATION TECHNOLOGY, NOIDA**

|  |  |  |
| --- | --- | --- |
| **TABLE OF CONTENTS** | | |
| **Chapter No.** | **Topics** | **Page No.** |
| **Chapter 1.**  **Introduction** | 1.1 General Introduction  1.2 Problem Statement  1.3 Significance/Novelty of the problem  1.4 Empirical Study  1.5 Brief Description of the Solution Approach  1.6 Comparison of existing approaches to the solution framed | 10-13 |
| **Chapter 2.**  **Literature Survey** | 2.1 Summary of papers studied | 14-25 |
| **Chapter 3.**  **Requirement**  **Analysis and**  **Solution Approach** | 3.1 Overall description of the project  3.2 Requirement Analysis  3.3 Solution Approach | 26-29 |
| **Chapter 4.**  **Modelling and**  **Implementation**  **Details** | 4.1 Design diagrams  4.1.1 Use Case diagrams  4.1.2 Class diagrams/Control  Flow diagrams  4.1.3 Sequence Diagram/Activity diagrams  4.2 Implementation details and issues  4.3 Risk Analysis and Mitigation | 30-35 |

(I)

|  |  |  |
| --- | --- | --- |
| **Chapter 5.**  **Testing(Focus on**  **Quality of**  **Robustness and**  **Testing)** | 5.1 Testing Plan  5.2 Component decomposition and type  of testing  5.3 List all test cases in prescribed  format  5.4 Errors and Exceptional Handling  5.5 Limitations of the Solution | 36-39 |
| **Chapter 6.**  **Findings,**  **Conclusions and**  **Future Work** | 6.1 Findings  6.2 Conclusions  6.3 Future Work | 40-42 |
|  | **References IEEE Format** |  |

(II)

**DECLARATION**

We hereby declare that the project entitled ‘Cloud Hybrid Network Intrusion Detection**’** has been carried away and completed by our group members only, under the guidance of our supervisor ‘Dr. Satish Chandra’. The contents of the project report are our own and that, to the best of our knowledge and belief, it does not contain any copied written material previously published by any person or organisation of higher learning. We also declare that the contents written in the project report has been based on our learning from various sources i.e. books, articles,journals. The project reported in this work has not been submitted earlier to our institution or any other institution, for the award of any degree, and will not be submitted post our submission.

Place: Noida

Date: June 6, 2021

|  |  |
| --- | --- |
| Signature : prashant shekhar  Name : Prashant Shekhar  Enrollment No : 17103012 | Signature : Aman  Name : Aman Ahuja  Enrollment No : 17103157 |
| Signature : abhinav  Name : Abhinav  Enrollment No : 17103020 |  |

(III)

**CERTIFICATE**

This is to certify that the work titled **“CLOUD HYBRID NETWORK INTRUSION DETECTION”** submitted by Prashant Shekhar, Abhinav and Aman Ahuja in partial fulfilment for the award of the degree of B. Tech from JIIT, Noida has been carried out under my supervision. Their dedication for which they carried away the project so easily is quite appreciating. Their knowledge towards the subject was the one of the causes for which they completed their project work perfectly, before their allotted time.As a supervisor, I only guided them wherever, they found it hard to get through. No personal help, either in project completion or report making, had been provided to them from my side.

Signature of Supervisor

Name of Supervisor Dr. Satish Chandra

Designation Associate Professor

Date June 6, 2021

(IⅤ)

**ACKNOWLEDGEMENT**

We are extremely thankful to Dr. Satish Chandra for his constant guidance and encouragement to accomplish this project within the allotted time period. We are indebted by the ocean of knowledge that our supervisor possesses,regarding the subject. His constant back support was the main reason why we did not face much difficulty in carrying forward the project and completing it within the allotted time frame. We are also thankful to the almighty god, who was always providing us strength to stand tall amidst all the hindrances that we faced during the project.

We extend our gratitude to the Department of Computer Science, Jaypee Institute of Information Technology for giving us this opportunity to apply our theoretical knowledge, regarding the subject, into the project work. We would also like to express our gratitude towards our parents, members of family and friends for their moral support.

|  |  |
| --- | --- |
| Signature : abhinav  Name : Abhinav  Enrollment No : 17103020 | Signature : Aman  Name : Aman Ahuja  Enrollment No : 17103157 |
| Signature : prashant shekhar  Name : Prashant shekhar  Enrollment No : 17103012 |  |

(Ⅴ)

**SUMMARY**

As more and more organizations are moving towards cloud based solutions, the need for securing the cloud network has become very important. The huge network infrastructure on cloud creates room for a huge attack surface for the intruders and imposes a great risk on the organizations using cloud solutions. Headway of Machine Learning Algorithms has created room for an intelligent approach to ensure security of the large networks. With our work we intend to develop an Intelligent Hybrid Cloud Intrusion Detection System that can help organisations secure their network from attacks.

Most of the Intrusion Detection systems use network packet signatures or network packet anomaly detection. But individual packet analysis in cloud networks is not possible due to huge network flow. Our approach is based using Machine Learning to recognise patterns in network streams that are available from cloud network providers. The Machine Learning model is prepared utilizing different supervised Machine Learning Algorithms like KNN, Random Forest and XGBoost and the model giving best precision is utilized to detect network intrusion. Model is validated utilizing cross validation. Organisations have isolated private networks for the internal workloads and databases and public networks for Internet facing assets such as Websites and Mobile Applications. With our work we strengthen the security with hybrid nature of such networks, thus providing multi layer enhanced security and a visual dashboard based on Elasticsearch, Logstash and Kibana to help the system admins in companies to easily manage network flows.

(ⅤI)

**LIST OF FIGURES**

|  |  |
| --- | --- |
| **Figure** | **Page number** |
| Figure 1 ELK Dashboard from AWS setup | 14 |
| Figure 2 Distribution of Attack Prediction on Network Stream | 15 |
| Figure 3 IP Addresses and Port Used By Attackers | 17 |
| Figure 4 Dataset after Pre Processing | 22 |
| Figure 5 Use Case Diagram | 23 |
| Figure 6 Control Flow Diagram | 23 |
| Figure 7 Sequence Diagram | 24 |
| Figure 8 Weighted Interrelationship Graph | 24 |
| Figure 9 Metrics of Lambda Function Tests | 25 |
| Figure 10 Attackers From External Network | 26 |
| Figure 11 Attackers From Internal Network | 26 |

(Ⅶ)

**LIST OF TABLES**

|  |  |
| --- | --- |
| **Table** | **Page No** |
| Table 1 | 12 |
| Table 2 | 19 |

(Ⅷ)

**LIST OF SYMBOLS AND ACRONYMS**

|  |  |
| --- | --- |
| **Acronym** | **Full form** |
| AWS | Amazon Web services |
| ELK | Elasticsearch Logstash Kibana |
| XGBoost | eXtreme Gradient Boosting |
| KNN | K Nearest Neighbours |
| CIDDS | Coburg Intrusion Detection Data Set |

(Ⅸ)

**Chapter - 1**

**1.1 General Introduction :**

The huge network infrastructure on cloud creates room for a huge attack surface for the intruders and imposes a great risk on the organizations using cloud solutions. Headway of Machine Learning Algorithms has created room for an intelligent approach to ensure security of the large networks. With our work we intend to develop an Intelligent Hybrid Cloud Intrusion Detection System that can help organisations secure their network from attacks..

Our approach is based using Machine Learning to recognise patterns in network streams that are available from cloud network providers. The Machine Learning model is prepared utilizing different supervised Machine Learning Algorithms like KNN, Random Forest and XGBoost and the model giving best precision is utilized to detect network intrusion. Model is validated utilizing cross validation. Organisations have isolated private networks for the internal workloads and databases and public networks for Internet facing assets such as Websites and Mobile Applications. With our work we strengthen the security with hybrid nature of such networks, thus providing multi layer enhanced security and a visual dashboard based on Elasticsearch, Logstash and Kibana to help the system admins in companies to easily manage network flows.

**1.2 Problem Statement :**

The aim of the project is to build an application which will detect all the malicious packets which may cause harm to the system. The increasing use of cloud computing has put the security of nodes present within a network at stake. To overcome this, we have made an attempt to build a model by using machine learning algorithms to detect the suspicious packets transferred within a network. The model built so has been tested on a dataset to compute the accuracy of the model. In addition, the application is supposed to be built not as a detector but also as a prevention against intrusion.

The dataset used is CIDDS. It is a labelled data set of Cisco netflow v9 format for evaluation of anomaly-based network IDS. This project will provide classification of attacks viz. into normal, unknown, suspicious, victim and attacker. The project aims at providing a hybrid setup of IDS (internal network and external network) for enhanced protection.

(X)

**1.3 Novelty of the problem:**

Large scale enterprises have a huge attack surface due to complex network structures and multiple cloud deployments. The current Network Intrusion Detection systems do not cover the entirety of the network over cloud. The intrusion detection system that we have built aims to detect the malicious or suspicious network streams which may harm enterprise networks and compromise confidential data to intruders. Our system provides a hybrid NIDS setup based on anomaly detection that supports internal network security and external network intrusion detection in the whole organisation scope. Thus even if any attacker is able to get through outside facing assets of the cloud network yet they would be detected by internal network NIDS. As a step further the solution that we provide considers not only detecting the possibility of an attack on a network but also to prevent the system from failing into attacks. This would ensure the zero-trust security in cloud and would also allow sensitive data of banks, payment systems, governments to reside on cloud with ease on the security of any network intrusion and ensure that security is decoupled from the cloud service provider.

**1.4 Empirical Study:**

**SNORT: (**Names- Zhou Zhimin, Chen Zhongwen, References - Yang Li. Research of intrusion detection systems based on Snort, The Technology and Application of Network Security **:** Snort is an NIDS based on static ruleset to define allowed and restricted entities of a network. It is widely adopted by large organisations and companies and has been a standard NIDS system. It is an open source project maintained by the Snort community. Major advantage is the easy integration of SNORT with 3rd party tools and existing network architecture.

**OWASP Intelligent Intrusion Detection System: (**Names- Alie El-Din Mady, Menouer Boubekeur, References- B. Donohue, Jan. 2014**):** Aintrusion Detection System by leveraging Machine Learning techniques to build a system which detects the intrusion and alerts the respective network administrator. OWASP Intelligent Intrusion Detection System is open source and maintained by the OWASP organization members. It is under fast development and thus every release has new features.

(XI)

**1.5 Brief Description Of Solution Approach**

Our Hybrid Cloud Network Intrusion Detection system allows organizations to reduce the security risk factor associated with cloud networks and Intelligently protect the system from any intruders. Our systems ensure multi level security by distinct internal and external network security. Even if an attacker happens to break through the external or Internet Facing still they will not be able to access the Internal Network that usually contains confidential information such as user databases. This two layer network security strengthens in security hardening and compliance. Thus even if any attacker is able to get through outside facing assets of the cloud network yet they would be detected by internal network NIDS. This would ensure the zero-trust security in cloud and would also allow sensitive data of banks, payment systems, governments to reside on cloud with ease on the security of any network intrusion and ensure that security is decoupled from the cloud service provider.

**1.6 Comparison of existing approaches to the solution framed :**

While the earlier studies and work done on cloud intrusion detection system illustrated various means to detect the presence of suspicious intruders in the computer network, we have tried to build a cloud intrusion detection hybrid model that will prevent the existence of malicious intruders in the computer network.The proposed solution is based on the idea of double level security. This double level security is related to the fact that attackers will have to breach the security of the outer network as well as the inner network. Most of the Intrusion Detection systems use network packet signatures or network packet anomaly detection. But individual packet analysis in cloud networks is not possible due to huge network flow. Our approach is based using Machine Learning to recognise patterns in network streams that are available from cloud network providers.

(ⅪI)

**Chapter 2: Literature Survey**

**2.1 Summary of papers studied :**

|  |  |  |
| --- | --- | --- |
| **Title** | **Author** | **Summary** |
| **Intrusion Detection and Prevention Systems For IOT, 2016[1]**  **(IEEE)** | Abishek verma, Babu Banarasi Das university, Lucknow, U.P., India Virender Ranga NIT, Kurukshetra, Haryana, India | CIDDS-001 dataset is used for its complex and statistical examination.Unsupervised and supervised types of techniques are used. Decision trees, Random forests are used as classifiers. This proves that this dataset is apt for Anomaly networks. |
| **Comparison of Machine Learning and Deep Learning models for Network Intrusion Detection System[2]**  **(IEEE)** | Dukka B. KC, Department of Computer science,North Carolina A&T State University, USA  Jipeng Liu, Department of Computational Data Science and Engineering,North Carolina A&T State University, USA | The Machine Learning model performed well due to the less ambiguous nature of future vectors in the CIDDS-001 dataset. It is also advantageous over the Deep Learning model in feature engineering. |
| **A survey of Network-based intrusion Detection Data sets[3]**  **(IEEE)** | Makus Ring, Deter Landes Department of Electrical Engineering and Computer science, Coburg university | This paper has overviewed 34 datasets including CIDDS-001 and each data set has its own peculiarities. |

(ⅫI)

|  |  |  |
| --- | --- | --- |
| **A Survey on CIDDS- 001 Dataset for Network Intrusion Detection Systems using Kernel based K-Means Algorithm[4]**  **(MAT JOURNALS)** | Gulafsha, Prakas Misra  Department of Computer Science and Engineering, Rajiv Gandhi Proudyogiki vishwavidyalaya, Bhopal, M.P., India | In this paper, both KNN and K-means clustering perform well over the CIDDS-001 dataset regarding utilized unmistakable measurements. So, the dataset can be utilized for anomaly based intrusion detection systems. |
| **A Cooperative Intrusion Detection System Framework for Cloud Computing Networks[5]** | Ci-Chun Lo, Institute of Information management National Chiao Tung University Hsinchu, Taiwan,Jo Ku,Institute of Information management National Chiao Tung University Hsinchu | Authors have propounded the IDS system to minimize the effect of DoS attack. In this system if one cloud region is affected by DoS attack the other performs well. |
| **LSTM for Anomaly Based Network Intrusion Detection[6]**  **(IEEE)** | Sara A.Althubiti, Department of Computer Science    Kausik Roy, North Carolina A&T State University Greensboro, NC | Authors have found LSTM has an edge over Naïve Bayes or SVM technique in multiple classification problems. |
| **Intrusion Detection in Cloud Computing[7]**  **(IEEE)** | Masimo Ficco, Luca Tasquer, Rocco Aversa Second University of Naples (SUN), Italy | The Research paper has propounded the model to build distributed IDS architecture in Cloud Computation. |

(ⅩⅣ)

|  |  |  |
| --- | --- | --- |
| **Efficient Approaches for Intrusion Detection in Cloud Environment[8]**  **(IEEE)** | Priti Mishra, Department of Computer Science and Engineering,  Emanuel S.Pilli, Malaviya National Institute of Technology, Jaipur | In this paper, the whole limelight was on machine learning with intrusion detection to deal with the security areas. Authors have assimilated such approaches in cloud computing and given security architectures. |
| **A Review of Intrusion Detection and Blockchain Applications in the cloud:Approaches,Challenges and Solutions[9]**  **(IEEE)** | Osama Alkedi, School of Engineering and Information Technology,  Nur Moustafa, University of New South Wales Canberra, Australia | Authors have explained the backdrop of IDS and Blockchain.They have also utilized and analyzed the scope of IDS and Blockchain techniques in such a framework. |
| **Hybrid Intrusion Detection System using Machine learning Techniques in cloudComputing Environments[10]**  **(IEEE)** | Ibrahim Aljamal  SUNY Polytechnic Institute, New-York, U.S.A | A hybrid IDS system is the use of unsupervised and supervised techniques to detect any kind of attacks in cloud atmospheres. This IDS system uses the K -means algorithm for clustering data. |

(ⅩⅤ)

|  |  |  |
| --- | --- | --- |
| **Intrusion Detection Based on Autoencoder and Isolation Forest in Fog Computing[11]**  **(IEEE)** | Kiswar Sadaf, Department of Computer Science    Jabin Sultana, College of Computer and Information Sciences, Majmaah University,  Saudi Arabia | For the fog environment authors have proposed a method for IDS using Autoencoder and Isolation Forest. The incoming packets are targeted for binary categorisation as fog are used for distinguishing normal and other attacks. |
| **Performance Evaluation of Lazy and Decision Tree Classifier: A Data Mining Approach for Global Celebrity's Death Analysis[12]** **(IEEE)** | [Vivek Kumar](https://ieeexplore.ieee.org/author/37086019144),Dept. of Automated Control System, National University of Science & Technology-MiSiS, Moscow, Russian Federation  [Roman Zinovyev](https://ieeexplore.ieee.org/author/37086492903), National University of Science & Technology-MiSiS, Moscow, Russian Federation | This Research paper was more concerned about utilizing decision trees and tried to find the death of celebrities in a decade based on these algorithms. |
| **A comparative analysis of different classification techniques for cloud intrusion detection systems' alerts and fuzzy classifiers[13]** **(IEEE)** | [Saed M. Alqahtani](https://ieeexplore.ieee.org/author/37086144741),  ASAP and LUCID Groups, University of Nottingham  [Robert John](https://ieeexplore.ieee.org/author/37271269700), ASAP and LUCID Groups, University of Nottingham | Authors have used various techniques and algorithms such as KNN, Decision tree etc. for better results to develop an intrusion detection system for better classification and identification of attacks |

(ⅩⅥ)

|  |  |  |
| --- | --- | --- |
| **Intrusion detection system combining misuse detection and anomaly detection using Genetic Network Programming[14]** **(IEEE)** | [Yunlu Gong](https://ieeexplore.ieee.org/author/37668938300), Department of Mathematics, Shanghai University, Shanghai, China Graduate School of Information, Production and Systems  [Shinjo Mabu](https://ieeexplore.ieee.org/author/37272290100), Waseda University, Fukuoka, Japan, Graduate School of Information, Production and Systems, Waseda University, Fukuoka, Japan | Genetic Network programming is used by the authors for developing intrusion detection systems for the combination of both anomaly and misuse identification. |
| **Distributed Intrusion Detection System using Blockchain and Cloud Computing Infrastructure[15]** **(IEEE)** | [Manis Kumar](https://ieeexplore.ieee.org/author/37401564500), M S Ramaiah Institute of Technology, Bangalore, India  [Asish Kumar Singh](https://ieeexplore.ieee.org/author/37086485856), M S Ramaiah Institute of Technology, Bangalore, India | Authors have developed an intrusion system basically distributed IDS by utilising the platform of blockchain and cloud environment. |
| **Building Multiclass Classification Baselines for Anomaly-based Network Intrusion Detection Systems[16]** **(IEEE)** | [Ajay Shah](https://ieeexplore.ieee.org/author/37088559592),Seton Hall University, South Orange, New Jersey, US  [Sophin Clachar](https://ieeexplore.ieee.org/author/37088559741), Seton Hall University, South Orange, New Jersey, US | Authors attempted to elucidate multiclass classification baselines that require more than one algorithm utilized in machine learning. These algo are used for changing valid network traffic from direct as well as unclear network intrusions that are illustrated. |

(ⅩⅦ)

|  |  |  |
| --- | --- | --- |
| **IDSaaS: Intrusion Detection System as a Service in Public Clouds[17]** **(IEEE)** | [Turki Alarkan](https://ieeexplore.ieee.org/author/38246738500), School of Computing, Queen's University, Kingston,  [Patric Martin](https://ieeexplore.ieee.org/author/37278040500), ONT, Canada,School of Computing, Queen's University, Kingston, ONT, Canada | Intrusion Detection as a Service targeted security of the infrastructure level of a public cloud has been proposed. Consumers cannot always just rely on the cloud provider's security infrastructure in a public cloud computing environment. Invigilating and securing their virtual existence by enforcing their own intrusion detection capabilities in addition to other security technologies within the cloud network is a must. To implement this,authors provide intrusion detection technology which is highly elastic, portable and fully controlled by the cloud consumer. |
| **Unified, Multi-level Intrusion Detection in Private Cloud Infrastructures[18]** **(IEEE)** | [Marti Humphry](https://ieeexplore.ieee.org/author/37276435000), Department of Computer Science, University of Virginia, Charlottesville, VA, USA  [Robert Emerson](https://ieeexplore.ieee.org/author/37086060624), Department of Computer Science, University of Virginia, Charlottesville, VA, USA | Authors presented an integrated, multi-level IDS architecture that illustrated modern approaches with protection and detection setups that briefed knowledge of right private cloud operations. |

(ⅩⅧ)

|  |  |  |
| --- | --- | --- |
| **Scalable intrusion detection systems log analysis using cloud computing infrastructure[19]** **(IEEE)** | [Manis Kumar](https://ieeexplore.ieee.org/author/37401564500), Department of Computer Science and Applications, Bangalore University, Bangalore, India  [M. Hanumanthappa](https://ieeexplore.ieee.org/author/38234740700),  Department of Computer Science and Applications, Bangalore University, Bangalore, India | Enhancing the throughput and scalability of the IDS Log analysis was the main goal of this literary work. After data collection, rapid analysis was required. Determining the presence of any attacks or malicious activities was the next objective, which was the reason that impacted the performance of IDS. |
| **Detection Framework for Monitoring the Traffic of Cloud Environments[20]** **(IEEE)** | [Bo L](https://ieeexplore.ieee.org/author/37085834113)e, Guangxi Key Lab of Multi-source Information Mining & Security, Guangxi, China  [Ping Liu](https://ieeexplore.ieee.org/author/37085839806), Guangxi Key Lab of Multi-source Information Mining & Security, Guangxi, China | For the purpose of monitoring the network traffic present within cloud environments, authors proposed cIDS. Enabling the invigilation of Intra-VM network traffic, was also achieved. The proposed framework is a novel cluster-based intrusion detection setup built to monitor the traffic inside cloud networks.. |

(ⅩⅨ)

|  |  |  |
| --- | --- | --- |
| **Securing Cloud from DDOS Attacks Using Intrusion Detection System in Virtual Machine[21]** **(IEEE)** | [Aman Baksi](https://ieeexplore.ieee.org/author/37399466300), [Yoges B. Dujodwala](https://ieeexplore.ieee.org/author/37087934686) SRM University, Chennai, TN, India  [Yoges B. Dujodwala](https://ieeexplore.ieee.org/author/37087934686), B.TECH-Computer Sci., SRM Univ., Chennai, India | In this paper, authors illustrated the utilisation of an IT strategy that could be implemented as a feedback to a contradiction of service attack. After selecting an inappropriate atypical hike in inbound congestion, transfer of selected applications to virtual machines can be done easily, hosted in distant data centers. |
| **Protocol Specific Multi-Threaded Network Intrusion Detection System (PM-NIDS) for DoS/DDoS Attack Detection in Cloud[22]** **(IEEE)** | [Rajindra Patil](https://ieeexplore.ieee.org/author/37085810070), NIT, Goa, India  [Harsa Dudeja](https://ieeexplore.ieee.org/author/37086084353), NIT, Goa, India | The growth of Cloud computing has brought an evolution in the arena of computing network technology. In addition to it, security and privacy concerns are noticeable and unavoidable.. Denial of service and Distributed Denial of Service attacks are the main threats to cloud technology in the past few years. Authors showcased an efficient security framework through this paper. It is a protocol peculiar Multithreaded Network Intrusion Detection  system. |

(ⅩⅩ)

|  |  |  |
| --- | --- | --- |
| **A Streaming Intrusion Monitoring and Classification System for IaaS Cloud[23]** **(IEEE)** | [Jason Nikoli](https://ieeexplore.ieee.org/author/37085999555), College of Business and Information Systems, Dakota State University, Madison, SD, USA  [Yong Wang](https://ieeexplore.ieee.org/author/38541588400), College of Computing, Dakota State University, Madison, SD, USA | The proposed work supported a security in depth method by gripping several distributed intrusion detection and security system sensors in an IaaS cloud computing setup. The Authors attempted to explain a streaming cloud intrusion invigilating and classification system to provide a back support to cloud service providers with numerous security systems by filtering noisy alert messages and classifying previously recognized attacks. |
| **Cloud Intrusion Detection Method Based on Stacked Contractive Auto-Encoder and Support Vector Machine** **[24]**  **(IEEE)** | Wenjon Wang ,PLA Strategic Support Force Information Engineering University  Xuhui Du, PLA Strategic Support Force Information Engineering University, Zhengzhou, Henan China | The potential for intrusion detection lies in Deep learning technology.Therefore,  extraction of essential feature representations automatically by using deep learning is the main objective of the proposed work. An effective assembled styptic autoencoder method is elucidated for unsupervised features . |

(ⅩⅩⅠ)

|  |  |  |
| --- | --- | --- |
| **Intrusion Detection for Grid and Cloud Computing[25]** **(IEEE)** | [Klever Vieira](https://ieeexplore.ieee.org/author/38557454000), Federal University of Santa Catarina, Brazil  [Alexandr Schulter](https://ieeexplore.ieee.org/author/37887863400), Federal University of Santa Catarina, Brazil | The objective of the study was to provide security in a distributed system that demands multiple users authentication along with passwords and security in transmission of data.The Grid and Cloud Computing Intrusion Detection System combines knowledge and behavior analysis for the purpose of intrusion detection |
| **The Research of Intrusion Detection System in Cloud Computing Environment[26]** **(Springer, Berlin)** | Huaeben Wang, Key Laboratory of Computer Vision and System, Ministry of Education  Haeyun Zho , Tianjin University of Technology,Tianjin, China | Multiple intrusion detection systems are used in each layer of cloud infrastructure for the safeguard of each Virtual Machine against threats. Authors also proposed the concept of cloud alliance by the communication agents having an exchange of mutual alerts between them, to cease the single point attack. The testing results point to a higher detection rate, lower false-positive rate,that has been achieved by the proposed system. |

(ⅩⅩⅡ)

|  |  |  |
| --- | --- | --- |
| **New balancing technique for green cloud computing and environmental Sustainability[27]** **(International Journal of Advanced Research 2015)** | Dr. Aman E. Khedr, Dr. Mona Nasr, Hesam Elmasri  Faculty of Computers & Information Helwan University | The need to propose a green cloud load balancing (GCLB) solution that aims to lower energy utilization in cloud data centers while maintaining the service level agreement between the customer and the cloud service provider was felt by the experts. Designing such solutions requires deep analysis of the Cloud network. . |
| **A hybrid web log based intrusion detection model[28]** **(IEEE)** | [Jig Yu](https://ieeexplore.ieee.org/author/37086143398), School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China  [Dan Ta](https://ieeexplore.ieee.org/author/37085856059)u, School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China | In particular, authors constructed a normal access model based on request feature vectors by using K-means clustering algorithm. The test data indicate that compared to a single intrusion detection model, the hybrid intrusion detection model proposed can effectively improve the rate of detection and lower the false alarm rate. |

(ⅩⅩⅢ)

|  |  |  |
| --- | --- | --- |
| **Network security for Hybrid Cloud[29]** **(IEEE)** | [Pasqual Donadio](https://ieeexplore.ieee.org/author/38276584100), Alcatel-Lucent Italia, Milan, Italy, Dip. di Ingegneria Elettrica e delle Tecnologie dcl l'Informazione  [Giovani B. Fioccola](https://ieeexplore.ieee.org/author/37085433473), Università degli Studi di Napoli Federico II, Napoli, Italy | In this paper, a solution for Hybrid Cloud security, focusing on a Virtual Intrusion Detection System (V-IDS), has been proposed. Also a briefing on a new building design that considers the basic principles of the Cloud computing, virtualization and GMPLS Control Plane and implies them to the intrusion detection systems,has been elucidated. The purpose behind the work was to protect Cloud networks characterized by constantly fluctuating the underlying infrastructure. |
| **Hypervisor-based cloud intrusion detection system[30]** **(IEEE)** | [Jaso Nikola](https://ieeexplore.ieee.org/author/37085999555)e, [Yung Wang](https://ieeexplore.ieee.org/author/37086002816)  College of Business and Information Systems Dakota State University, Madison, SD, USA, College of Business and Information Systems Dakota State University, Madison, SD, USA | The authors proposed an architecture and method to leverage the virtualization technology at the core of cloud computing. It performs intrusion detection security using hypervisor performance metrics Hypervisor-based cloud intrusion detection system explained in this paper does not need extra software pre installed in virtual machines. |

(ⅩXIV)

**Chapter - 3**

**Requirement Analysis and Solution Approach**

**3.1 Overall description of the project**

****

Figure 1 :- ELK Dashboard from AWS setup

Network security in the cloud is a big concern due to no direct physical control over servers/ resources. Our Hybrid Cloud Network Intrusion Detection system allows organizations to reduce the security risk factor associated with cloud networks. Our systems ensure multi level security. Even if an attacker happens to break through the external or Internet Facing still they will not be able to access the Internal Network that usually contains confidential information such as user databases. This two layer network security strengthens in security hardening and compliance. Thus even if any attacker is able to get through outside facing assets of the cloud network yet they would be detected by internal network NIDS. This would ensure the zero-trust security in cloud and would also allow sensitive data of banks, payment systems, governments to reside on cloud with ease on the security of any network intrusion and ensure that security is decoupled from the cloud service provider.

(ⅩXV)

**3.2 Requirement Analysis**

* Dataset of Network streams based on isolated internal network and Internet facing external network.
* Machine Learning Model to detect intrusion in cloud networks
* Cross-validation and accuracy of Random Forest and XGBoost Model.
* Deployment of ML Model to detect intrusion in live network on a cloud setup such as AWS.
* Elasticsearch, Logstash and Kibana to visualize the level of possibility of intrusion in cloud networks.

**3.3 Solution Approach**

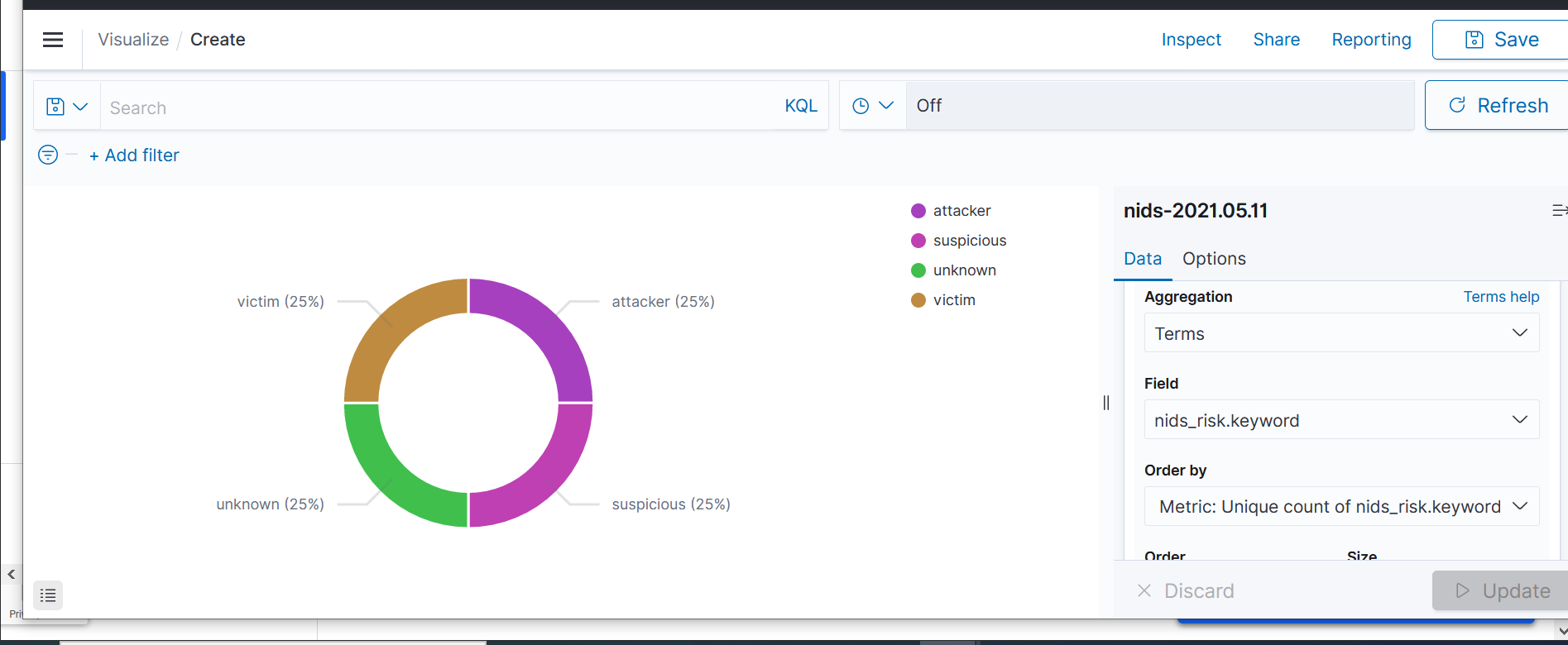


Figure 2 :- Distribution of Attack Prediction on Network Stream

Hybrid Cloud NIDS - A Network Intrusion Detection based on anomaly detection in network streams in cloud networks. Enhanced security with isolated protection of internal network resources. Security Analyst can easily monitor and analyse potential attacking attempts with the provided Elasticsearch, Logstash and Kibana Dashboard. Incident Response Administrators are instantly notified about any attack in the cloud network so as to take rapid action.

(ⅩXVI)

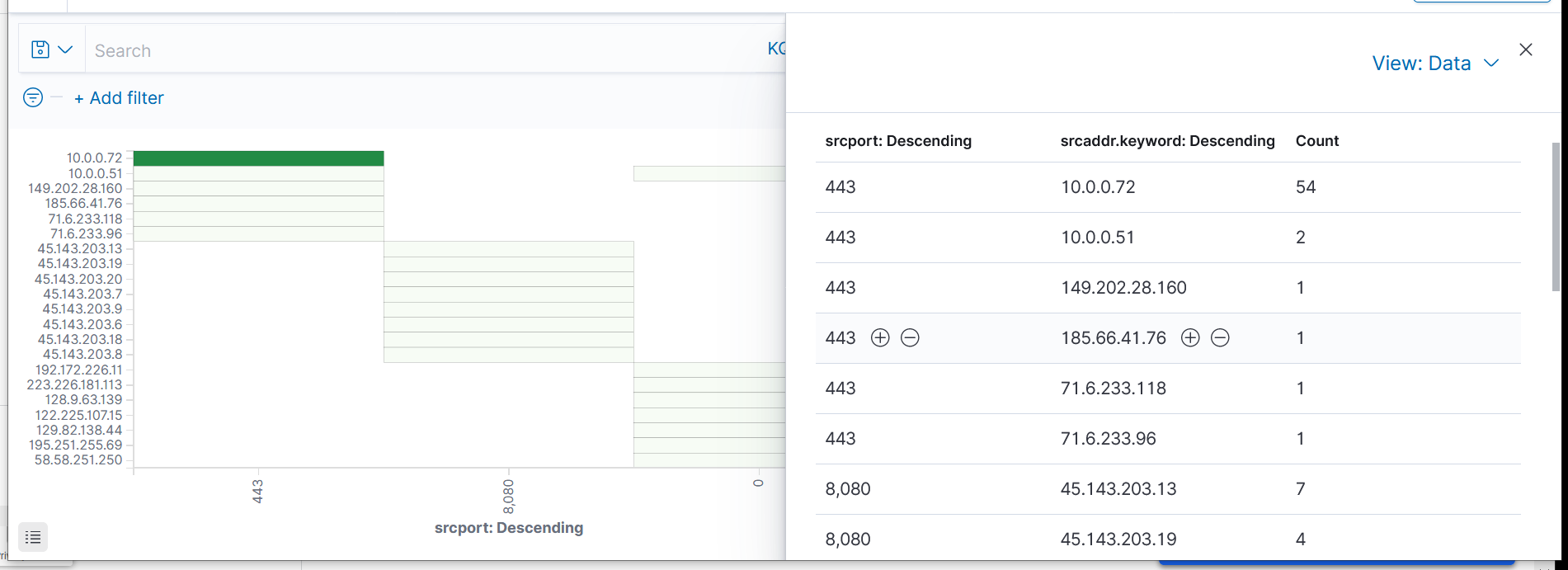
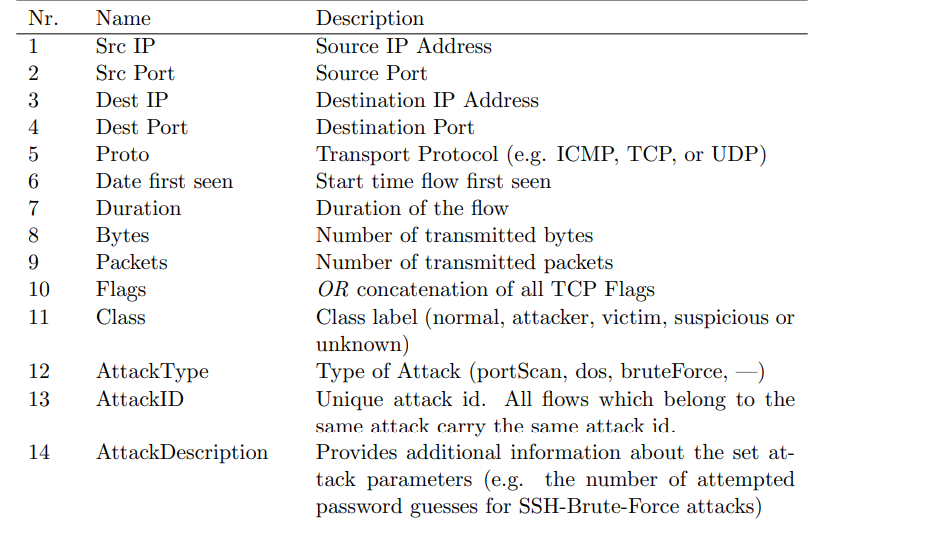
****

Figure 3 :- IP Addresses and Port Used By Attackers

**Intrusion Detection System**

**Dataset**

For our study we have chosen CIDDS-001 Dataset. It is a fairly recent labelled dataset of network flow streams. For creation of the CIDDS-001 data set, a small environment was emulated using the OpenStack environment. Following are the attributes of Dataset:



(ⅩVII)

**Data Preparation**

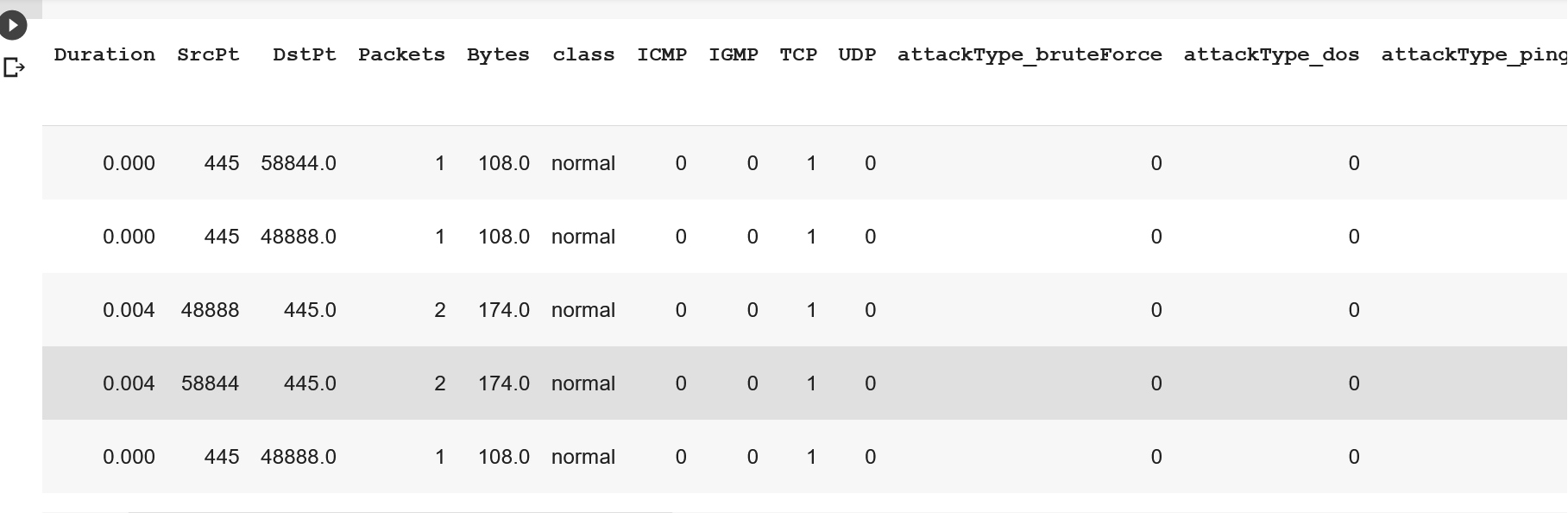


Figure 4 :- Dataset after Pre Processing

We performed the following preprocessing techniques.

* One Hot Encoding - Integer encoded variable is removed and a replacement binary variable is added for every unique integer value.
* Data Factorize - Encodes labels into categorical variables

**Models:**

We applied various supervised algorithms to our dataset

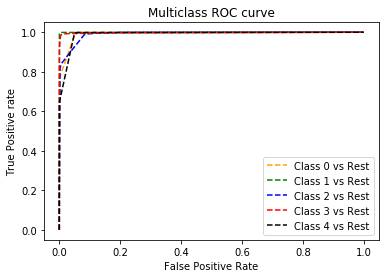
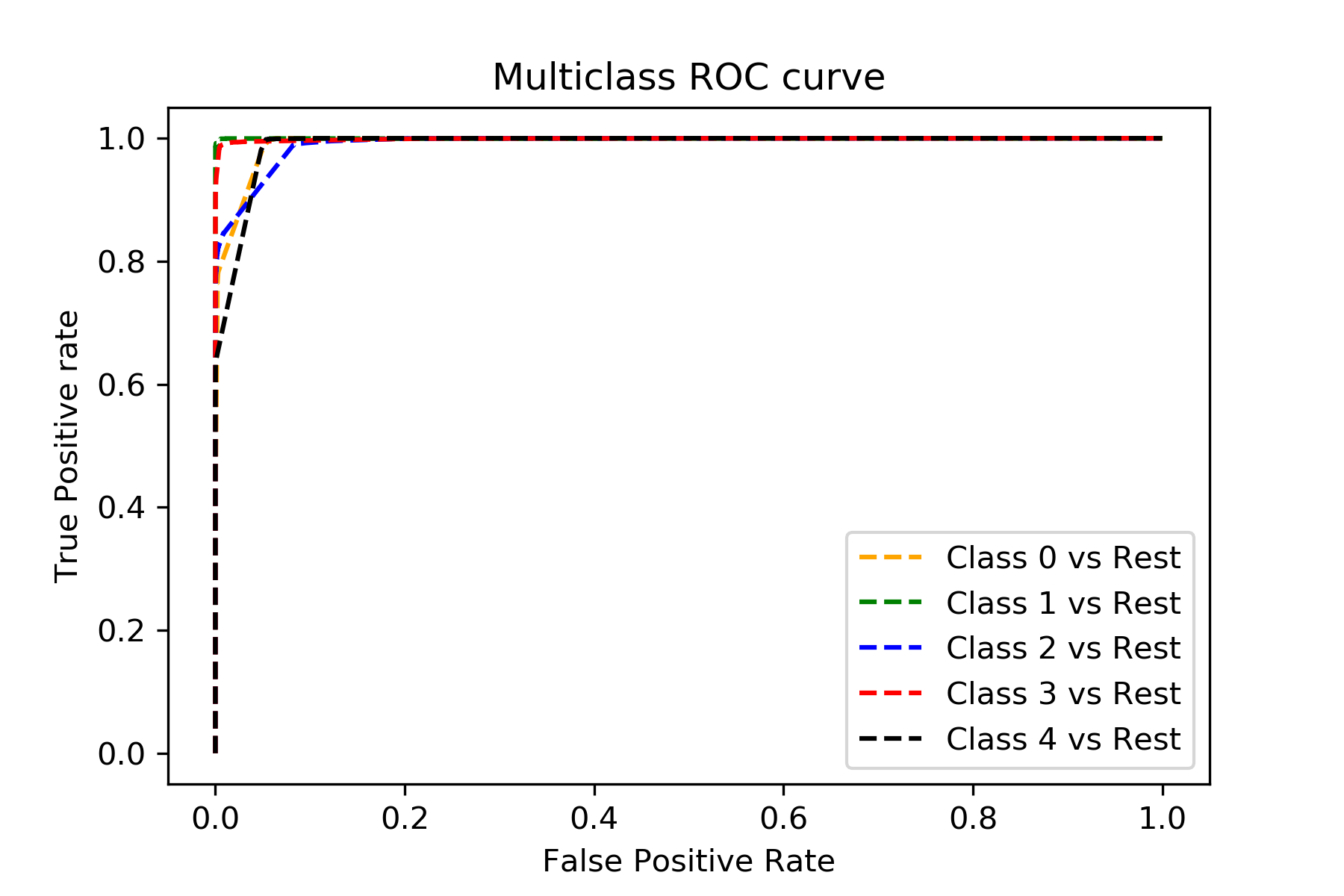
* KNN
* Random Forest
* XGBoost

**Validation:**

XGBoost gave the best results with the following validation metrics.

k Fold Cross Validation - Results showed 97.25% for external network and 97.4% for internal network accuracy with 5 fold cross validation technique.

**ROC - AUC**

****

(ⅩVⅢ)

**Chapter 4**

**Modeling and Implementation Details**

**4.1 Design Diagrams**

**4.1.1Use Case diagrams**

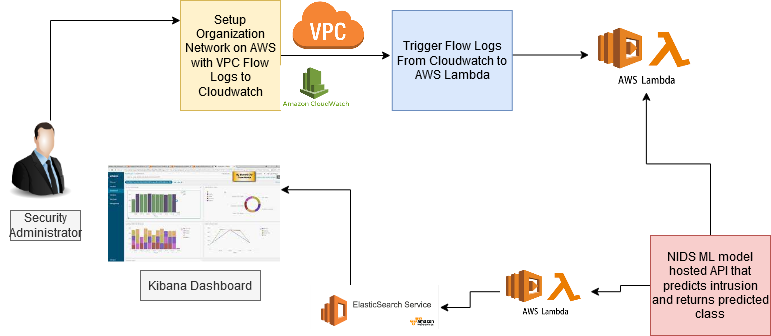
****

Figure 5:- Use Case Diagram

(ⅩⅩIX)

**4.1.2 Control Flow Diagrams**

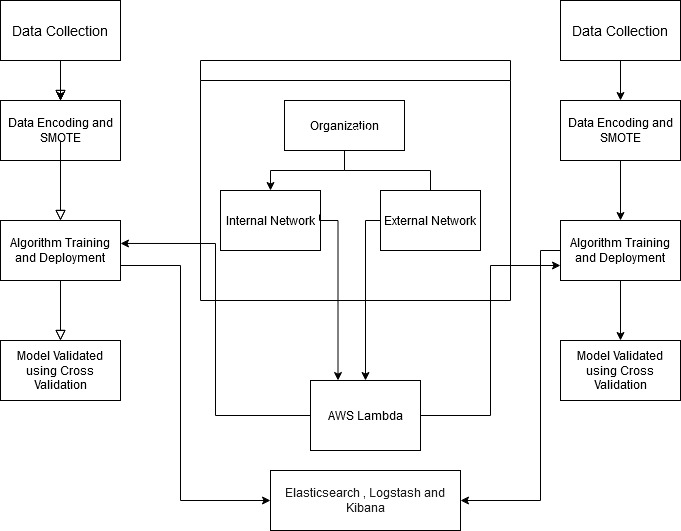


Figure 6:- Control Flow Diagram

(ⅩⅩX)

**4.1.3 Sequence Diagram**

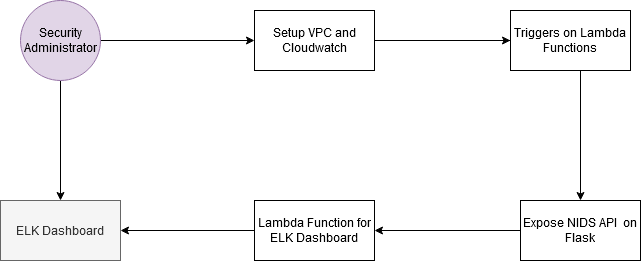
****

Figure 7:- Sequence Diagram

**4.2 Implementation details and issues**

The Dashboard is made using Elasticsearch, Logstash and Kibana. Elasticsearch is an open source, advanced search engine and analytics platform that supports various languages. Logstash is a tool that allows us to collect and manage data from various sources. Kibana is an open source dashboard that has various visualization features and query options. The dataset used is CIDDS-001 Dataset available from. AWS Lambda is used for retrieving predicted class value from Machine Learning Model.The machine learning model is created using python Jupyter. Various python libraries have been used for processing data, creating models and validating models. Pickle library of python has been used for storing the Machine Learning Model which has been used in the web application for ranking the candidates. Pandas library has been used for reading data and numpy has been used to store data in numpy arrays. Scikit - learn library has been used for preprocessing data and creating models.

(ⅩⅩXI)

**4.3 Risk Analysis and Mitigation**

Table 1: Risk Analysis

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Risk\_ID** | **Classification** | **Description of Risk** | **Risk Area** | **Impact** |
| Risk\_1 | Design | The possibility of low accuracy as we are using traditional machine learning algorithms. | Performance | High (H) |
| Risk\_2 | Engineering Specialties | The project scope demands maximum possible reliability on the predicted outcomes, because if the attacked network stream is marked as non attacked then it can be a dangerous situation for the organisation as it concerns security of the company. | Reliability | Medium (M) |
| Risk\_3 | Requirements | Risk of availability of complete, robust and reliable dataset with proper labels for training our models candidate. | Completeness | Low (L) |

(ⅩⅩXII)

Table 2 : Risk Area Wise Total Weighting Factor

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S. No** | **Risk Area** | **Weights (Input)** | **Total Weights** | **Priority** |
| 1 | Performance | 9+3+3+1 | 16 | 1 |
| 2 | Budget | 9+3+1 | 13 | 2 |
| 3 | Hardware Constraints | 9+3+1 | 13 | 3 |
| 4 | Reliability | 9+3 | 12 | 4 |
| 5 | Requirements | 3+1 | 4 | 5 |

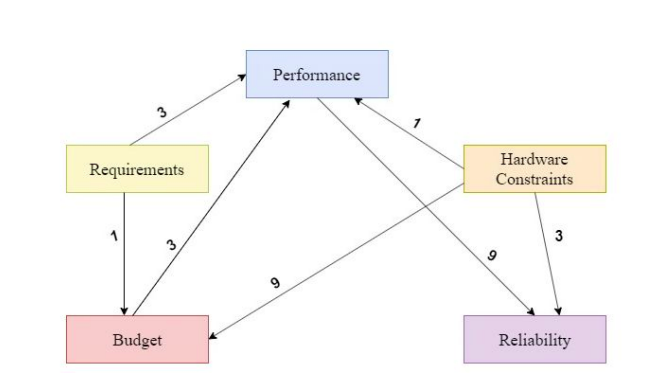
****

Figure 8 Weighted Interrelationship Graph

(ⅩⅩXⅢ)

**Chapter 5**

**Testing**

Software testing is a very important part of the software development life cycle as it verifies the system under test and also validates whether it works as per requirements stated in BRS and satisfies the stakeholders' needs. With reference to the Hybrid cloud based network intrusion detection system also, software testing, validating, verifying & evaluation is significant; as it is vital to check whether the system is working as expected before deployment. Since the system is required to heavily process very sensitive data and thus must be reliable in production as this may define security of the organization's data which is extremely crucial. To assess the system output, efficiency of the system, performance benchmark and appropriate quality assessment techniques should be adopted for comparison to the benchmark level and with the similar sorts of different products

**5.1 Testing Plan**

We used various supervised regression algorithms and then validated using the K-Fold Cross Validation Technique. The procedure features a single parameter called k that refers to the amount of groups that a given data sample is to be split into.In our setup we've taken k as 5 thus 5 fold cross validation tests.

**5.2 Component decomposition and type of testing required**

The objectives behind the testing of our developed model are:

● Evaluation of Parameters of the developed system

● Calculating accuracy

● Integrated features of AWS and ELK Dashboard

|  |  |  |
| --- | --- | --- |
| Type of test | Explanation | Software component |
| Requirement Testing | Validation checks were made to ensure that hardware and software specifications meet the minimum requirements. Certain libraries were required to be specially installed and the minimum CPU/GPU requirements for our architecture were also checked. | Anaconda/Google Colab |

(ⅩⅩⅩⅣ)

|  |  |  |
| --- | --- | --- |
| Performance Testing | Performance testing is the process of determining the speed, accuracy, and consistency of the proposed model. This was achieved by creating, training, and testing the Random Forest Model. | Anaconda/Google Colab |
| Experimental Testing | Our model was checked against various experimental tests to fine-tune the hyperparameters. The Random Forest model was experimented with different learning rates and training methodologies. | Anaconda/Google Colab |

.

**5.3 List of test cases in prescribed format**

**Sanity of Lambda Function:**

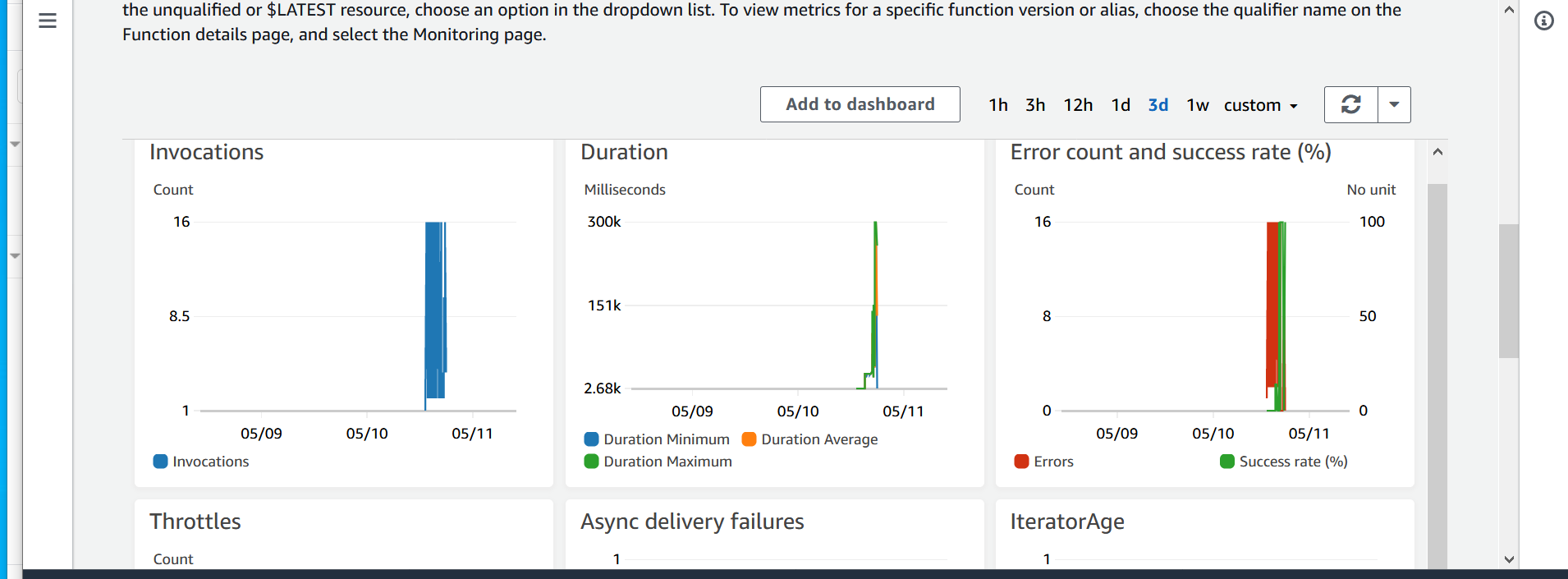
****

Figure 9 Metrics of Lambda Function Tests

**Predicted Network Flow Stream:**

****

Figure 10 Attackers From External Network

(ⅩⅩXV)

**IP Addresses of Attackers in Internal Network:**

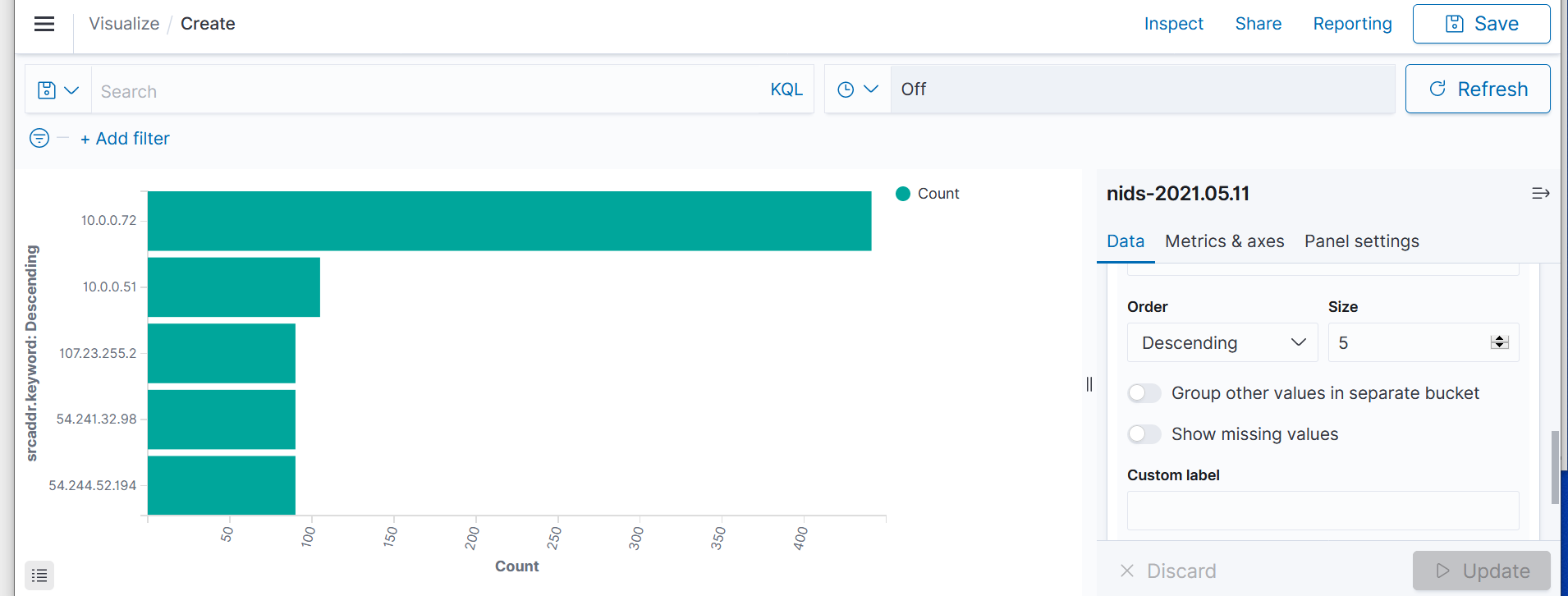


Figure 11 Attackers From Internal Network

**5.4 Limitations of the solution**

The system is created to efficiently predict intrusion in a real time cloud network such as on AWS. A Supervised Machine Learning algorithm has been used to Classify the network flow stream into a particular class viz attacker, suspicious, normal, victim and unknown. Training of Machine Learning Model to create an accurate model requires a varied dataset. We were able to find an appropriate dataset but it did not take into account several other factors related to the network such as header length of packets in the stream flow. Also the model is limited to the scenario where the information of network stream flow from the router can be extracted in Cisco flow version 9 specification. Due to this factor the model and approach is a bit less flexible as not every Cloud Service Provider shares flow logs in Cisco Netflow v9 format. Real time prediction of intrusion in Cloud networks is not fully compliant if based only on Anomaly Detection Models as there can be false negative results that can be the cause of a data breach in the organisation. Also this method requires extra resources such as ELK dashboard and serverless computation functions such as AWS Lambda that has a large cost associated with it because of the large amount of AWS Cloudwatch Logs. Thus cost is a big limitation, however, log retention policies can help achieve lower cost and also comply with standards such as GDPR.

(ⅩXXVI)

**Chapter 6**

**6.1 Findings**

**Internal Network**

|  |  |  |
| --- | --- | --- |
| **Algorithm** | **Model Accuracy** | **Cross Validation Average Accuracy** |
| **KNN** | 97.92% | 96.21% |
| **Random Forest** | 99.61% | 97.4% |
| **XGBoost** | 99.58% | 97.23% |

**External Network**

|  |  |  |
| --- | --- | --- |
| **Algorithm** | **Model Accuracy** | **Cross Validation Average Accuracy** |
| **KNN** | 97.61% | 95.84% |
| **Random Forest** | 98.91% | 97.25% |
| **XGBoost** | 98.92% | 97.1% |

We used Validation to evaluate the ML models. It can be seen that the Random Forest and XGBoost model produce the best results. On the other hand, KNN performs poorly, suggesting that the selection criteria are not separable based on Classifying as Neighbour Distance.

(ⅩXXVII)

**6.2 Conclusion**

In this project, we have presented a novel approach for detecting intrusion in cloud networks such as AWS. We have developed the system for hybrid scenarios, that is, a system taking into account both the internal as well as external attacks. The Dataset used is CIDDS-001. The accuracy with random forest reaches near 99.61% and that with XGBoost is 99.58% for internal networks. For Model Validation we have used a 5-fold cross validation technique. We can conclude that our system can be easily integrated and our approach gives Security Administration an overall picture of the Network Security of their organization and help them to take informed decisions and protect their data. Security Analyst can easily monitor and analyse potential attacking attempts with the provided Elasticsearch, Logstash and Kibana Dashboard. Incident Response Administrators are instantly notified about any attack in the cloud network so as to take rapid action.

**6.3 Future Work**

There have been significant improvements in Network Intrusion Detection Systems. As a future scope we would like to build automatic decision making capabilities in the system based making it a Detection and Prevention System that can ensure network security and is self sufficient. Also such a system can be extended to other cloud service providers such as Google Cloud and Azure Cloud by Microsoft or other setups that do not have Cisco NEtflow v9 based logs. There needs to be extensive work done to reduce log management and reduce the pricing of overall Security Operation Centres and to automate all the processes as it will help companies to detect and prevent any cyber attack at early stages.

(ⅩⅩXVⅢ)

**REFERENCES**

[1] C. Esposito, M. Ficco, F. Palmieri, and A. Castiglione, Interconnecting Federated Clouds by Using Publish-Subscribe Service, In Cluster Computing, 2013, pp. 1–17.

[2] A. Patela, M. Taghavia, K. Bakhtiyaria, JC. Jniorc. An intrusion detection and prevention system in cloud computing: A systematic review. In Journal of Network and Computer Applications, vol. 36, no. 1, Jan. 2013, pp. 25–41.

[3] N. Moustafa, K.-K.-R. Choo, I. Radwan, and S. Camtepe, ‘‘*Outlier Dirichlet mixture mechanism”.* IEEE Trans. Inf. Forensics Security, vol. 14, no. 8, pp. 1975–1987, Aug. 2019.

[4] B. B. Zarpelao, R. S. Miani, C. T.

Kawakani, and S. C. de Alvarenga, ¯ ‘‘*A survey of intrusion detection in Internet of Things,*’’ J. Netw. Comput. Appl., vol. 84, pp. 25–37, Apr. 2017.

[5] D. Basu, X. Wang, Y. Hong, H. Chen, and S. Bressan, ‘*‘Learn-as-you-go with Megh: Efficient live migration of virtual machines,*’’ in Proc. IEEE 37th Int. Conf. Distrib. Comput. Syst. (ICDCS), Jun. 2017, pp. 1786–1801.

[6] D. Vaghela, Analysis and security testing of android systems by malware detection in networks. international journal of engineering sciences & research technology, pp.274-278, 2015.

[7] Aminanto, M. E., Choi, R., Tanuwidjaja, H. C., Yoo, P. D., & Kim, K. (2018). Deep abstraction and weighted feature selection.

[8] Ajay Kumara M. A and Jaidhar C. D, “*Hypervisor and virtual machine dependent Intrusion Detection and Prevention System for virtualized cloud environment*” Kuala Lumpur, 2015, pp. 28-33.

[9] Z chiba N. Abghour, K. Moussaid, A. El omri, M. Rida “*A Cooperative and Hybrid Network Intrusion Detection Framework in Cloud Computing Based on Snort and Optimized Back Propagation Neural Network*”.

[10] B. I. Santoso, M. R. S. Idrus and I. P. Gunawan, ‘‘*Designing Network Intrusion Detection System .*6th International Annual Engineering Seminar (InAES), Yogyakarta, 2016,pp. 61-66

[11] Fog computing and the Internet of Things: Extend the cloud to where the things are San Jose CA USA Apr. 2015 [online] Available: https://www.Cisco.Com, C. Bhatt and C. K. Bhensdadia "*Fog computing: Applications concepts and issues*" Int. J. Grid High Perform. Comput. vol. 9 no. 4 pp. 105-113 Oct. 2017, F. Bonomi R. Milito P. Natarajan and J. Zhu "*Fog computing: A platform for Internet of Things and analytics*".

[12] Vijay kumar and Vijender Kumar Solanki Singh "*Data Mining approach of Accident Occurrences Identification with Effective Methodology and Implementation in International*".

[13] X. Zhang and G Premkumar "*Data mining for network intrusion detection: a comparison of alternative methods*" .

.

[14] C. Chen S. Mabu K. Shimada and K. Hirasawa "*Network Intrusion Detection using Class Association Rule Mining Based on Genetic Network Programming*" IEEJ Transactions on Electrical and Electronic Engineering, S. Mabu K. Hirasawa and J. Hu "*A GraphBased Evolutionary Algorithm: Genetic Network Program-ming(GNP) and Its Extension Using Reinforcement Learning*" in Evolutionary Computation MIT press vol. 15 no. 3 pp. 369-398 2007.

[15] A. A. Titorenko and A. A. Frolov "*Analysis of modern intrusion detection system*"

[16] Ivan Homoliak and Petr Hanacek "*ASNM Datasets: A Collection of Network Traffic Data for Testing of Adversarial Classifiers and Network Intrusion Detectors*" ,Management pp. 350-358 2013 ISBN 1-60132-259-3.

[17] C. Burns. (2011 Oct. 10) "*Public cloud security remains MISSION IMPOSSIBLE*" Network World [Online] Website: http://www.networkworld.com/ supp / 2011 / enterprise5/101011-ecs-cloud-security-250973.html.

[18] Mary R. Thompson,Humphrey Marty and Keith R. Jackson "*Security for grids*", Marty Humphrey and Feng Ju "*Eliminating replica selection-using multiple replicas to accelerate data transfer on grids*" Parallel and Distributed Systems 2004. ICPADS 2004. Proceedings. Tenth International Conference on 2004.

[19]A. Wespi "Towards a taxonomy of Intrusion Detection Systems" and Axelsson Stefan "Intrusion Detection Systems: A Taxonomy and Survey" Technical Report No 99-15 Dept. of Computer Engineering Chalmers University of Technology Sweden March 2000, H.Debar M. Dacier

[20] Amazon Elastic Compute Cloud (Amazon EC2)[EB/OL] 12 2015 [online] Available: http://aws.amazon.com/ec2/, Sefraoui Omar Mohammed Aissaoui and Mohsine Eleuldj "*OpenStack: toward an open-source solution for cloud computing*" International Journal of Computer Applications 55 vol. 3 2012.

[21] V.V Das "Honeypot Scheme for Distributed Denial-of-Service", Y Kobayashi "Intrusion Detection Systems: technology and development"., S. Khattab and R. Mosse Melhem "*Honeypot backpropagation for mitigating spoofing distributed Denial-of-Service attacks*"

[22] S. Stelios J. Perkins and M. Rinard Cloud Intrusion Detection and Repair (CIDAR) Cambridge United States:MIT Computer Science and Artificial Intelligence Laboratory 2016, R. Sebastian F. Cheng and C. Meinel "*Intrusion detection in the cloud*"

.

[23] S. Cobb Target breach 12 months on a year of lessons learned 2014 [online] Available:http://www.welivesecurity.com/2014/12/18/target-breach-lessons-learned/#lesson, M-Trends: Beyond the Breach - 2014 Threat Report 2014 [online] Available: https://dl.mandiant.com/EE/library/WP\_M-Trends2014\_140409.pdf.

[24] Zhendong Wang, Yong Zeng, Yaodi Liu, Dahai Li, "*Deep Belief Network Integrating Improved Kernel-Based Extreme Learning Machine for Network Intrusion Detection*", *Access IEEE*, vol. 9, pp. 16062-16091, 2021.

[25] F-Y. Leu "*Integrating Grid with Intrusion Detection*" Proc. Int'l Conf. Advanced Information Networking and Applications (AINA 05) vol. 1 pp. 304-309 2005, H. Debar M. Dacier and A. Wespi "*Towards a Taxonomy of Intrusion Detection Systems*" Int'l J. Computer and Telecommunications Networking vol. 31 no. 9 pp. 805-822 1999.

[26] Kandukuri, B.R., Paturi, V., Rakshit, A.: Cloud Security Issues, Laureano, M., Maziero, C., Jamhour, E.: Protecting Host-based Intrusion Detectors through Virtual Machines. Computer Networks 51.

[27] P. Mell and T. Grance. “*The NIST definition of cloud computing*”., pp.1-106, (2011) , M .AL-Roomi, S. Al -Ebrahim, S. rais and I. Ahmad, “*Cloud Computing Pricing Models: A Survey*”

International Journal of Grid and Distributed Computing Vol.6, No.5 (2013), pp.93-106.

[28] Sungdeok Cha Junsup Lee Sangrok Kim and Sanghyun Cho "*ADAM: Web Anomaly Detection Assistant Based on Feature Matrix*" Ninth International Conference on Quality Software pp. 123-128 2009.

[29] R. Buyya et al. Cloud computing and emerging IT platforms: Vision hype and reality for delivering computing as the 5th utility Future Generation Computer Systems (2009) doi:10.1016/j.future.2008.12.001, Cloud Futures 2010-Cloud Computing.

[30] S. Biggs and S. Vidalis "*Cloud Computing: The Impact on Digital Forensic Investigations* " 2010 pp. 1-6, G. Brunette and R. Mogull "*Security guidance for critical areas of focus in cloud computing v3.0* " Cloud Security Alliance 2011.