

CONTENTS

| | |
|------------------------|------------|
| <i>Abstract</i> | <i>i</i> |
| <i>Acknowledgment</i> | <i>ii</i> |
| <i>Contents</i> | <i>iii</i> |
| <i>List of Figures</i> | <i>vi</i> |
| <i>List of Tables</i> | <i>vii</i> |

| <u>Chapters</u> | <u>Page No.</u> |
|--|-----------------|
| Chapter 1: Introduction | 01 |
| 1.1 Overview | 01 |
| 1.2 Network Security | 01 |
| 1.2.1 Anomaly in Network Security | 01 |
| 1.2.2 Intrusion Detection System | 01 |
| 1.2.3 Attack Classification | 02 |
| 1.3 Machine Learning and its Applications | 03 |
| 1.4 Types of Intrusion Detection Systems | 04 |
| 1.4.1 Signature Based Detection | 04 |
| 1.4.2 Statistical Anomaly Based Detection | 04 |
| 1.5 Machine Learning | 04 |
| 1.5.1 Supervised Learning | 05 |
| 1.5.2 Unsupervised Learning | 06 |
| 1.5.3 Reinforcement Learning | 06 |
| 1.6 Types of Anomaly Based Detection | 07 |
| 1.6.1 Statistical Anomaly Detection | 07 |
| 1.6.2 Data Mining Based Approach | 09 |
| 1.6.3 Knowledge Based Detection Techniques | 10 |
| Chapter 2: Literature Survey | 12 |
| 2.1 Statistical Modeling | 12 |
| 2.1.1 Markov Models | 12 |
| 2.1.2 Markov Chains | 13 |

| | |
|--|-----------|
| 2.1.3 Hidden Markov Models | 13 |
| 2.1.4 HMM Based Related Work | 15 |
| 2.2 Naive Bayes Classifier | 16 |
| 2.2.1 Bayes Theorem | 16 |
| 2.2.2 Formal Definition and Background | 17 |
| 2.2.3 Properties of Naive Bayes | 18 |
| 2.2.4 Advantages and Limitations | 18 |
| 2.2.5 Recent Research in Network Anomaly Detection using NBC | 19 |
| 2.3 Proposed System | 22 |
| Chapter 3: System Analysis | 23 |
| 3.1 Requirements | 23 |
| 3.1.1 Hardware Requirements | 23 |
| 3.1.2 Software Requirements | 23 |
| 3.1.3 Functional Requirements | 23 |
| 3.2 Datasets | 24 |
| Chapter 4: System Design | 27 |
| 4.1 Design Considerations | 27 |
| 4.1.1 Modules | 27 |
| 4.1.2 Use Case Diagram | 29 |
| 4.1.3 Data Flow Diagram | 30 |
| 4.1.4 State Diagram | 30 |
| Chapter 5: Implementation | 32 |
| 5.1 Language Used for Implementation | 32 |
| 5.1.1 Python | 32 |
| 5.2 Libraries Used for Implementation | 33 |
| 5.2.1 NumPy | 33 |
| 5.2.2 Flask | 34 |
| 5.2.3 Matplotlib | 34 |
| 5.3 Pseudo Code | 35 |
| 5.3.1 Determine Optimal Bands Procedure | 35 |
| 5.3.2 Update Probabilities Procedure | 36 |
| 5.3.3 Train Procedure | 37 |

| | |
|---|-----------|
| 5.3.4 Train Phase Procedure | 38 |
| 5.3.5 Determine Probability Procedure | 39 |
| 5.3.6 Determine Threshold Probability Procedure | 39 |
| 5.3.7 Deploy Procedure | 40 |
| Chapter 6: Testing | 41 |
| 6.1 Testing Methods | 41 |
| 6.1.1 Static Testing | 41 |
| 6.1.2 Dynamic Testing | 41 |
| 6.1.3 White Box Testing | 41 |
| 6.1.4 Black Box Testing | 41 |
| 6.2 Testing Levels | 42 |
| 6.2.1 Unit Testing | 42 |
| Chapter 7: Results | 44 |
| 7.1 Screenshots | 44 |
| 7.2 Results Obtained | 48 |
| Chapter 8: Conclusion | 50 |
| Chapter 9: Challenges and Future Work | 51 |
| Appendix | 52 |
| Bibliography | 57 |