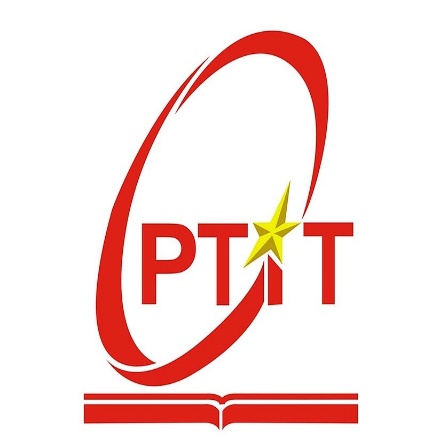
**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

**KHOA AN TOÀN THÔNG TIN**



**ĐỒ ÁN TỐT NGHIỆP**

**ĐỀ TÀI**

**XÂY DỰNG HỆ THỐNG PHÁT HIỆN XÂM NHẬP**

**PHÁT HIỆN TẤN CÔNG DDOS DỰA TRÊN HỌC MÁY**

**Họ và tên: Nguyễn Đức Linh**

**Mã sinh viên: B20DCAT109**

**Nhóm lớp: 01**

**Giảng viên hướng dẫn: Đinh Trường Du**

**HÀ NỘI – 2023**

# **LỜI CẢM ƠN**

**a**

# **NHẬN XÉT, ĐÁNH GIÁ VÀ CHO ĐIỂM**

**(Của Giảng viên hướng dẫn)**

................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................

**Điểm**: ......................... (**Bằng chữ**: .......................................................................................)

**Đồng ý/Không đồng ý** cho sinh viên bảo vệ trước hội đồng chấm đồ án tốt nghiệp?

*Hà Nội, ngày ... tháng ... năm ...*

CÁN BỘ - GIẢNG VIÊN HƯỚNG DẪN

(Ký và ghi rõ họ tên)

# **NHẬN XÉT, ĐÁNH GIÁ VÀ CHO ĐIỂM**

**(của)**

................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................

**Điểm**: ......................... (**Bằng chữ**: .......................................................................................)

**Đồng ý/Không đồng ý** cho sinh viên bảo vệ trước hội đồng chấm đồ án tốt nghiệp?

*Hà Nội, ngày ... tháng ... năm ...*

CÁN BỘ - GIẢNG VIÊN HƯỚNG DẪN

(Ký và ghi rõ họ tên)

# **MỤC LỤC**

# **DANH MỤC TỪ VIẾT TẮT**

# **DANH MỤC HÌNH VẼ**

# **DANH MỤC BẢNG**

# **MỞ ĐẦU**

Ngày nay, Internet mang lại cho chúng ta những lợi ích vô cùng to lớn, và dần trở thành một phần không thể thiếu trong cuộc sống của con người. Tuy nhiên, bên cạnh những lợi ích và sự thuận tiện mà Internet mang lại đó là những hiểm họa, nguy cơ tiềm tàng từ tin tặc, những kẻ luôn lợi dụng những sở hở để tấn công, nhằm chuộc lợi riêng. Trong số các hiểm hoạ, nguy cơ thường trực thì tấn công từ chối dịch vụ DoS/DDoS là một trong các nguy cơ lớn đối với an toàn của hệ thống, mạng và người dùng. Các nạn nhân của DoS/DDoS không ngừng gia tăng và số lượng các mạng Botnet phục vụ DDoS cùng những quy mô của chúng ngày càng được mở rộng và sử dụng kỹ thuật công nghệ cao. Do đó, nhiều các công trình nghiên cứu cũng như các tổ chức đã tìm hiểu và đưa ra những giải pháp nhằm phát hiện sớm nhất tấn công DoS/DDoS hay giảm thiểu thiệt hại do chúng gây ra.

Chính vì vậy, đồ án lựa chọn sử dụng phương pháp học máy để có thể phát hiện sớm tấn công DDoS. Đồ án này tập trung vào nghiên cứu và ứng dụng các kỹ thuật học máy để xây dựng mô hình phát hiện sớm DDoS dựa trên lưu lượng mạng. Đề tài “Nghiên cứu phát hiện tấn công DDoS dựa trên học máy” bao gồm các nội dung chính sau:

Chương 1: Tổng quan về các dạng tấn công mạng phổ biến trên thế giới hiện nay. Từ đó đi sâu vào tìm hiểu về các dạng và cách phòng chống tấn công DoS/DDoS.

Chương 2: Tổng quan về học máy, phân loại các thuật toán học máy và một số thuật toán học máy sẽ được dùng để thử nghiệm trong đồ án. Ngoài ra, chương này cũng sẽ trình bày lý thuyết về cách ứng dụng học máy vào phát hiện tấn công DDoS.

Chương 3: Trong chương này, đồ án sẽ trình bày về cách xây dựng mô hình, tập dữ liệu, tiền xử lý dữ liệu và cách huấn luyện mô hình. Sau khi huấn luyện sẽ sử dụng các thông số tính điểm để đánh giá độ hiệu quả của các thuật toán được sử dụng và chọn ra thuật toán hiệu quả nhất cho mô hình.

# **CHƯƠNG 1: TỔNG QUAN VỀ TẤN CÔNG MẠNG, TẤN CÔNG DOS/DDOS VÀ PHÒNG CHỐNG**

*Ở chương 1, đồ án sẽ trình bày về các dạng tấn công mạng phổ biến trên thế giới hiện nay. Từ đó đi sâu vào tìm hiểu về các dạng và cách phòng chống tấn công DoS/DDoS, một trong những hình thức tấn công mạng phổ biến và đã xuất hiện từ những ngày đầu của Internet.*

1. **Khái quát về tấn công mạng**
2. **Giới thiệu tấn công xâm nhập**
3. **Các dạng tấn công xâm nhập**
4. **Một số dạng tấn công mạng điển hình**
5. **Tổng quan về tấn công DoS/DdoS**
6. **Tấn công DoS**
7. **Tấn công DDoS**
8. **Các biện pháp phòng chống tấn công DoS/DdoS**
9. **Kết luận chương**

**C1: Tổng quan về tấn công mạng, tấn công dos/ddos**

**- Các cuộc tấn công mạng**

**- Tấn công DoS/DdoS**

**+/ Lý thuyết**

**+/ Phân loại**

**- Phòng chống**

**CÁC ĐẶC TRƯNG CỦA KDD99**

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Feature** | **Meaning** | **Type** |
| 1 | duration | length (number of seconds) of the connection | continuous |
| 2 | **protocol\_type** | type of the protocol (tcp, udp, icmp) | discrete |
| 3 | service | network service on the destination (http, telnet, ...) | discrete |
| 4 | **flag** | normal or error status of the connection | discrete |
| 5 | **src\_bytes** | number of data bytes from source to destination | continuous |
| 6 | **dst\_bytes** | number of data bytes from destination to source | continuous |
| 7 | land | 1 if connection is from/to the same host/port; 0 otherwise | discrete |
| 8 | wrong\_fragment | number of “wrong” fragments | continuous |
| 9 | urgent | number of urgent packets | continuous |
| **Basic features of individual TCP connections** | | | |
| 10 | **hot** | number of “hot” indicators | continuous |
| 11 | **num\_failed\_logins** | number of failed login attempts | continuous |
| 12 | **logged\_in** | 1 if successfully logged in; 0 otherwise | discrete |
| 13 | **num\_compromised** | number of “compromised” conditions | continuous |
| 14 | **root\_shell** | 1 if root shell is obtained; 0 otherwise | discrete |
| 15 | **su\_attempted** | 1 if “su root” command attempted; 0 otherwise | discrete |
| 16 | **num\_root** | number of “root” accesses | continuous |
| 17 | **num\_file\_creations** | number of file creation operations | continuous |
| 18 | **num\_shells** | number of shell prompts | continuous |
| 19 | **num\_access\_files** | number of operations on access control files | continuous |
| 20 | **num\_outbound\_cmds** | number of outbound commands in an ftp session | continuous |
| 21 | **is\_host\_login** | 1 if the login belongs to the “hot” list; 0 otherwise | discrete |
| 22 | **is\_guest\_login** | 1 if the login is a “guest” login; 0 otherwise | discrete |
| **Content features within a connection suggested by domain knowledge** | | | |
| 23 | **count** | number of connections to the same host as the current connection in the past two seconds | continuous |
| *\** | *Note: The following features refer to these same-host connections.* | | |
| 24 | **srv\_count** | number of connections to the same service as the current connection in the past two seconds | continuous |
| 25 | serror\_rate | % of connections that have SYN errors | continuous |
| 26 | rerror\_rate | % of connections that have REJ errors | continuous |
| 27 | **same\_srv\_rate** | % of connections to the same service | continuous |
| 28 | **diff\_srv\_rate** | % of connections to different services | continuous |
| *\** | *Note: The following features refer to these same-service connections* | | |
| 29 | srv\_serror\_rate | % of connections that have SYN errors | continuous |
| 30 | srv\_rerror\_rate | % of connections that have REJ errors | continuous |
| 31 | srv\_diff\_host\_rate | % of connections to different hosts | continuous |
| **Traffic features computed using a two-second time window** | | | |
| 32 | dst\_host\_count |  |  |
| 33 | dst\_host\_srv\_count |  |  |
| 34 | dst\_host\_same\_srv\_rate |  |  |
| 35 | dst\_host\_diff\_srv\_rate |  |  |
| 36 | dst\_host\_same\_src\_port\_rate |  |  |
| 37 | dst\_host\_srv\_diff\_host\_rate |  |  |
| 38 | **dst\_host\_serror\_rate** |  |  |
| 39 | **dst\_host\_srv\_serror\_rate** |  |  |
| 40 | dst\_host\_rerror\_rate |  |  |
| 41 | dst\_host\_srv\_rerror\_rate |  |  |
| 42 |  |  |  |
| 43 |  |  |  |

"duration",

"protocol\_type",

"service",

"flag",

"src\_bytes",

"dst\_bytes",

"land",

"wrong\_fragment",

"urgent",

"hot",

"num\_failed\_logins",

"logged\_in",

"num\_compromised",

"root\_shell",

"su\_attempted",

"num\_root",

"num\_file\_creations",

"num\_shells",

"num\_access\_files",

"num\_outbound\_cmds",

"is\_host\_login",

"is\_guest\_login",

"count",

"srv\_count",

"serror\_rate",

"srv\_serror\_rate",

"rerror\_rate",

"srv\_rerror\_rate",

"same\_srv\_rate",

"diff\_srv\_rate",

"srv\_diff\_host\_rate",

"dst\_host\_count",

"dst\_host\_srv\_count",

"dst\_host\_same\_srv\_rate",

"dst\_host\_diff\_srv\_rate",

"dst\_host\_same\_src\_port\_rate",

"dst\_host\_srv\_diff\_host\_rate",

"dst\_host\_serror\_rate",

"dst\_host\_srv\_serror\_rate",

"dst\_host\_rerror\_rate",

"dst\_host\_srv\_rerror\_rate",

# **CHƯƠNG 2: TỔNG QUAN VỀ HỆ THỐNG IDS**

*Ở chương 2, đồ án sẽ trình bày tổng quan về học máy, phân loại các thuật toán học máy và một số thuật toán học máy sẽ được dùng để thử nghiệm trong đồ án. Ngoài ra, chương này cũng sẽ trình bày lý thuyết về cách ứng dụng học máy vào phát hiện tấn công DDoS.*

1. **Tổng quan về hệ thống IDS**
2. **Đặc điểm hệ thống IDS**
3. **Ứng dụng hệ thống IDS**
4. **Kết luận chương**

C2: Tổng quan về hệ thống IDS

* Lý thuyết
* Phân loại
* Cách hoạt động
* Ưu, nhược điểm

# **CHƯƠNG 3: PHÁT HIỆN TẤN CÔNG DOS/DDOS DỰA TRÊN HỌC MÁY**

*Ở chương 2, đồ án sẽ trình bày tổng quan về học máy, phân loại các thuật toán học máy và một số thuật toán học máy sẽ được dùng để thử nghiệm trong đồ án. Ngoài ra, chương này cũng sẽ trình bày lý thuyết về cách ứng dụng học máy vào phát hiện tấn công DDoS.*

1. **Tổng quan về học máy**
2. **Xây dựng mô hình phát hiện tấn công DoS/DdoS dựa trên học máy**
3. **Kết luận chương**

C3: Phát hiện tấn công DdoS dựa trên học máy

* Tổng quan về học máy

+/ Lý thuyết

+/ Một số kỹ thuật học máy

* Xây dựng mô hình học máy phát hiện ddos

+/ Tổng quan mô hình

+/ Các giai đoạn xử lý

# **CHƯƠNG 4: THỬ NGHIỆM VÀ KẾT QUẢ**

*Ở chương 3, đồ án sẽ trình bày về cách xây dựng mô hình, tập dữ liệu, tiền xử lý dữ liệu và cách huấn luyện mô hình. Sau khi huấn luyện sẽ sử dụng các thông số tính điểm để đánh giá độ hiệu quả của các thuật toán được sử dụng và chọn ra thuật toán hiệu quả nhất cho mô hình.*

1. **Xây dựng kịch bản**
2. **Huấn luyện và thử nghiệm mô hình**
3. **Thu thập và xử lý gói tin**
4. **Giả lập cuộc tấn công DoS/DdoS**
5. **Kết luận chương**

C4: Kiểm thử và kết quả

* Xây dựng kịch bản
* Huấn luyện và kiểm thử mô hình

+/ Mô tả dataset

+/ Các giai đoạn tiền xử lý

* Tiền xử lý
* Tích chọn đặc trưng quan trọng

+/ Huấn luyện và thử nghiệm 1 vài mô hình

* Decision tree
* Naive bayes
* Random forest
* Thu thập và xử lý gói tin từ cuộc tấn công

+/ Tool KDD99 extractor

+/

* Giả lập cuộc tấn công DdoS

+/ Mô tả các bước thực hiện

* Demo
* Kết luận

# **KẾT LUẬN**

# **DANH MỤC TÀI LIỆU THAM KHẢO**

# **PHỤ LỤC**