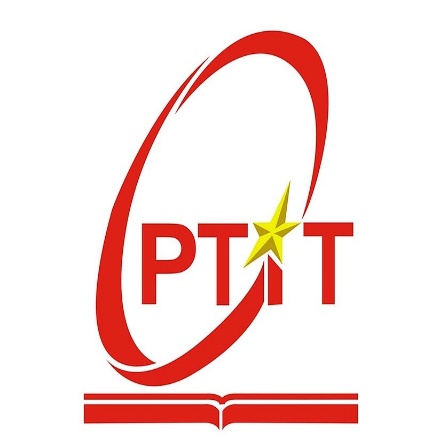
**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

**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 Duy**

**HÀ NỘI – 2023**

**LỜI CẢM ƠN**

**Trang nhiệm vụ**

**MỤC LỤC**

**DANH MỤC TỪ VIẾT TẮT**

**DANH MỤC HÌNH ẢNH**

**DANH MỤC BẢNG BIỂU**

**MỞ ĐẦU (1)**

**C1**

**C2**

**C3**

**C4**

**Kết luận**

**Tài liệu tham khảo**

**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",

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

C2: Tổng quan về hệ thống IDS

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

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

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