

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI  
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG  
\_\_\_\_\_ \* \_\_\_\_\_

ĐỒ ÁN  
**TỐT NGHIỆP ĐẠI HỌC**  
NGÀNH CÔNG NGHỆ THÔNG TIN

**PHÂN TÍCH PHÁT HIỆN DOS SỬ DỤNG  
NETFLOW TÍCH HỢP TRONG CISCO IOS**

Sinh viên thực hiện : **Hoàng Công Thắng**  
Lớp : **IS1 – K55**  
Giáo viên hướng dẫn : **PGS. TS. Ngô Hồng Sơn**

HÀ NỘI 5-2015

# PHIẾU GIAO NHIỆM VỤ ĐỒ ÁN TỐT NGHIỆP

## 1. Thông tin về sinh viên

Họ và tên sinh viên: HOÀNG CÔNG THẮNG

Điện thoại liên lạc: 01649 591 158

Email: thanghc2810@gmail.com

Lớp: IS1 – K55

Hệ đào tạo: Đại học chính quy

Đồ án tốt nghiệp được thực hiện tại: Bộ môn Truyền thông và Mạng máy tính, Viện CNTT&TT, Đại học Bách Khoa Hà Nội.

Thời gian làm ĐATN: Từ ngày 21/02/2015 đến 29/05/2015

## 2. Mục đích nội dung của ĐATN

- Ứng dụng các giải thuật vào phân tích phát hiện bất thường mạng trong tấn công SYN-flood sử dụng công cụ NetFlow tích hợp trong Cisco IOS của các thiết bị mạng Cisco.

## 3. Các nhiệm vụ cụ thể của ĐATN

- Tìm hiểu công cụ nhúng NetFlow trong các thiết bị mạng Cisco.
- Cấu hình và cài đặt thiết bị định tuyến của Cisco trong hệ thống mạng.
- Sử dụng dữ liệu NetFlow vào việc phân tích bất thường mạng.
- Tìm kiếm giải thuật phân tích phát hiện bất thường mạng trong tấn công SYN-flood.
- Cài đặt giải thuật phân tích phát hiện bất thường mạng trong tấn công SYN-flood sử dụng dữ liệu từ NetFlow.
- Tiến hành chạy mô phỏng, đánh giá thuật toán và so sánh kết quả.

## 4. Lời cam đoan của sinh viên:

Tôi – *Hoàng Công Thắng* - cam kết ĐATN là công trình nghiên cứu của bản thân tôi dưới sự hướng dẫn của *PGS. TS. Ngô Hồng Sơn*.

Các kết quả nêu trong ĐATN là trung thực, không phải là sao chép toàn văn của bất kỳ công trình nào khác.

*Hà Nội, ngày tháng năm*

Tác giả ĐATN

*Hoàng Công Thắng*

## 5. Xác nhận của giáo viên hướng dẫn về mức độ hoàn thành của ĐATN và cho phép bảo vệ:

*Hà Nội, ngày tháng năm*

Giáo viên hướng dẫn

*PGS. TS. Ngô Hồng Sơn*

## TÓM TẮT NỘI DUNG ĐỒ ÁN TỐT NGHIỆP

Trong các mạng doanh nghiệp lớn hoặc các hệ thống máy chủ, việc phân tích và giám sát lưu lượng mạng trong hệ thống có ý nghĩa rất quan trọng để đảm bảo việc duy trì tính ổn định và khả năng hoạt động của hệ thống mạng. Khả năng phân tích và giám sát được lưu lượng mạng trong hệ thống là hoạt động rất quan trọng đối với việc quản trị mạng. Việc kiểm tra dòng lưu lượng mạng sẽ đảm bảo tài nguyên mạng được sử dụng một cách hợp lý hơn. Nó giúp người quản trị nhận biết được chất lượng dịch vụ (QoS), nhận biết được dấu hiệu hay nguy cơ của những cuộc tấn công từ chối dịch vụ (DoS), việc phát tán virus, cũng như hàng loạt các sự cố khác.

NetFlow là tính năng của Cisco IOS cho phép thống kê lưu lượng gói qua bộ định tuyến. NetFlow thực hiện giám sát, phân tích, tính toán lưu lượng gói. Đáp ứng hầu hết các yêu cầu trong việc giám sát hệ thống mạng.

Đồ án này trình bày về cách thức hoạt động của NetFlow tích hợp trong các thiết bị Cisco. Qua đó thiết lập và cài đặt NetFlow trên một hệ thống sever nhỏ. Trích xuất ra dữ liệu NetFlow để sử dụng cho quá trình phân tích bất thường trên lưu lượng mạng.

## **ABSTRACT OF THESIS**

In large enterprise networks or server systems, the analysis and network traffic monitoring system is very important to ensure the maintenance of the stability and performance of the network. Ability to analyze and monitor the network traffic in the operating system is very important for network administrators. The inspection of network traffic flows to ensure network resources are used more sensibly. It helps administrators to recognize the quality of service (QoS), recognize the signs or risk of attacks Denial of Service (DoS), virus propagation, as well as numerous other incidents.

NetFlow is a Cisco IOS feature that allows traffic statistics through the router package. NetFlow performance monitoring, analysis, flow calculation package. Meet most of the requirements in the monitoring network.

This project provides information about how to operate the integrated NetFlow in Cisco devices. Thereby setting and install NetFlow on a small server systems. Extract data for use NetFlow analysis on network traffic anomaly.

## LỜI CẢM ƠN

Trước hết, tôi xin gửi lời cảm ơn chân thành tới PGS. TS. Ngô Hồng Sơn, người đã tận tình dạy dỗ và hướng dẫn tôi trong quá trình hoàn thành đồ án cũng như trong học tập.

Đồng thời, tôi xin bày tỏ lòng biết ơn đến các thầy cô giáo trong Viện Công nghệ thông tin và Truyền thông – trường Đại học Bách Khoa Hà Nội, những người đã tận tâm giảng dạy, truyền đạt cho tôi những kiến thức cơ bản làm nền tảng cho việc thực hiện đồ án cũng như trong quá trình công tác sau này.

Tôi cũng xin gửi lời cảm ơn tới các anh chị tại trường Đại học Bách Khoa Hà Nội, các bạn, các tôi cùng làm việc trong phòng thí nghiệm, những người luôn ở bên cạnh giúp đỡ, động viên tôi trong quá trình hoàn thành đồ án.

Cuối cùng, với tất cả sự kính trọng, con xin bày tỏ lòng biết ơn sâu sắc tới bố mẹ và anh chị tôi trong gia đình đã luôn là chỗ dựa tinh thần vững chắc và tạo mọi điều kiện cho con ăn học nên người.

*Tôi xin chân thành cảm ơn !*

# MỤC LỤC

PHIẾU GIAO NHIỆM VỤ ĐỒ ÁN TỐT NGHIỆP .....	2
TÓM TẮT NỘI DUNG ĐỒ ÁN TỐT NGHIỆP .....	3
ABSTRACT OF THESIS .....	4
LỜI CẢM ƠN .....	5
MỤC LỤC.....	6
DANH MỤC CÁC BẢNG.....	8
DANH MỤC HÌNH VẼ.....	9
DANH MỤC CÁC TỪ VIẾT TẮT VÀ THUẬT NGỮ .....	10
GIỚI THIỆU ĐỀ TÀI VÀ ĐỊNH HƯỚNG GIẢI QUYẾT .....	11
1. Tổng quan .....	11
2. Mục tiêu đề tài .....	11
3. Định hướng giải quyết .....	12
4. Giới thiệu nội dung chính .....	13
CHƯƠNG I. CƠ SỞ LÝ THUYẾT .....	14
1.1 Giới thiệu NetFlow .....	14
1.1.1 Ứng dụng của NetFlow.....	15
1.1.2 Tầm quan trọng của việc nhận thức hoạt động mạng.....	16
1.1.3 Cách thức hoạt động của NetFlow.....	17
1.1.4 Thông tin của luồng trong một bản ghi NetFlow .....	17
1.1.5 Cách truy cập dữ liệu tạo ra bởi NetFlow.....	18
1.1.6 Chi tiết thực hiện một báo cáo của NetFlow .....	18
1.1.7 Vị trí của NetFlow trong mạng.....	19
1.1.8 Định dạng của dữ liệu gửi đi bởi NetFlow .....	19
1.2 Giới thiệu bộ công cụ flow-tools.....	19
1.3 Tìm hiểu tấn công từ chối dịch vụ DoS.....	20
1.3.1 Giới thiệu về DoS .....	20
1.3.2 Các mục đích của tấn công DoS .....	20
1.3.3 Mục tiêu mà kẻ tấn công thường sử dụng tấn công DoS.....	20
1.3.4 Dấu hiệu khi bị tấn công DoS.....	20
1.3.5 Các kỹ thuật tấn công DoS .....	21
1.4 Giới thiệu các giải thuật phát hiện tấn công SYN-flood .....	28
1.4.1 Thuật toán AT (Adaptive Threshold) .....	29
1.4.2 Thuật toán CUSUM ( Cumulative SUM).....	29
CHƯƠNG II. MÔ HÌNH HỆ THỐNG.....	32

2.1	Xây dựng hệ thống mạng giám sát bằng NetFlow .....	32
2.2	Mô phỏng tấn công SYN-flood .....	33
2.2.1	Sử dụng công cụ flow-tools để xử lý dữ liệu cho NetFlow .....	33
2.2.2	Cài đặt thuật toán .....	34
2.2.3	Tổ chức chương trình.....	37
CHƯƠNG III. ĐÁNH GIÁ THUẬT TOÁN .....		38
3.1	Bài toán mô phỏng.....	38
3.2	Đánh giá hiệu quả thuật toán .....	38
3.2.1	Tấn công cường độ cao.....	38
3.2.2	Tấn công với cường độ thấp .....	41
3.2.3	Sự cân bằng giữa tỉ lệ DP và FAR trong thuật toán AT .....	44
3.2.4	Ảnh hưởng của tham số biên độ $\alpha$ .....	44
3.2.5	Ảnh hưởng của các yếu tố $\beta$ trong EWMA .....	45
3.3	Nhận xét về giải thuật.....	46
KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN .....		47
1.	Kết quả đạt được .....	47
2.	Hướng phát triển .....	47
PHỤ LỤC.....		49
1.	Cài đặt minicom trên máy giám sát.....	49
2.	Cấu hình NetFlow cho bộ định tuyến.....	50
TÀI LIỆU THAM KHẢO.....		56

## DANH MỤC CÁC BẢNG

Bảng 1: Sự khác nhau giữa SNMP và NetFlow.....	17
Bảng 2: Ví dụ đồ thị phân phối chuẩn .....	30
Bảng 3: Thông tin của flow header format .....	34
Bảng 4: Thông tin trong flow record format .....	35
Bảng 5: Tỷ lệ DP của hai thuật toán với tần công cường độ cao.....	41
Bảng 6: Tỷ lệ DP của hai thuật toán với tần công cường độ thấp .....	44



## DANH MỤC HÌNH VẼ

Hình 1: Ví dụ về mô hình mạng sử dụng NetFlow .....	12
Hình 2: Hình ảnh của bộ định tuyến Cisco 1921 .....	13
Hình 3: Các thành phần chính trong mô hình hoạt động của NetFlow.....	14
Hình 4: Ví dụ về một bộ đệm NetFlow.....	17
Hình 5: Gói tin TCP Header được Attacker sử dụng.....	21
Hình 6: Mô hình tấn công Ping Of Death .....	22
Hình 7: Cơ chế tấn công bằng Teardrop .....	22
Hình 8: Mô hình tấn công bằng SYN Flood Attack .....	24
Hình 9: Cấu tạo gói tin TCP.....	24
Hình 10: Mô hình tấn công bằng Land Attack.....	25
Hình 11: Mô hình tấn công Smuft Attack.....	25
Hình 12: Mô hình tấn công bằng Fraggle Attack.....	26
Hình 13: Mô hình tấn công bằng UDP Flood .....	27
Hình 14: Mô hình tấn công DDoS .....	28
Hình 15: Mô hình mạng thử nghiệm.....	32
Hình 16: Ví dụ các gói UDP NetFlow gửi đến máy giám sát.....	33
Hình 17: Lược đồ phân tích tìm gói tin SYN trong bản ghi NetFlow .....	35
Hình 18: Lược đồ mô tả chương trình .....	36
Hình 19: Giao diện chương trình .....	37
Hình 20: Biểu đồ phát hiện tấn công SYN-flood sử dụng thuật toán AT với time intervals lần lượt là 10, 20, 30, 40 giây.....	39
Hình 21: Biểu đồ phát hiện tấn công SYN-flood sử dụng thuật toán CUSUM với time intervals lần lượt là 10, 20, 30, 40 giây.....	40
Hình 22: Biểu đồ phát hiện tấn công SYN-flood sử dụng thuật toán AT trong tấn công cường độ thấp với time intervals lần lượt là 10, 20, 30, 40 giây .....	42
Hình 23: Biểu đồ phát hiện tấn công SYN-flood sử dụng thuật toán CUSUM trong tấn công cường độ thấp với time intervals lần lượt là 10, 20, 30, 40 giây.....	43
Hình 24: Biểu đồ đánh giá sự cân bằng giữa xác suất phát hiện và tỷ lệ báo động giả .....	44
Hình 25: Biểu đồ đánh giá ảnh hưởng của biên độ $\alpha$ .....	45
Hình 26: Biểu đồ đánh giá ảnh hưởng của tham số $\beta$ .....	45
Hình 27: Giao diện chính của minicom .....	50
Hình 28: Thông số cài đặt minicom để kết nối với cổng console.....	50
Hình 29: Hiển thị thông tin và trạng thái các cổng .....	51
Hình 30: Ví dụ kết nối với mạng ngoài.....	51
Hình 31: Hiển thị thông tin bảng NAT .....	52
Hình 32: Hiển thị trạng thái mạng VLAN .....	52
Hình 33: Ví dụ hiển thị thông tin NetFlow .....	54
Hình 34: Ví dụ hiển thị trạng thái Netflow .....	55

## DANH MỤC CÁC TỪ VIẾT TẮT VÀ THUẬT NGỮ

Chữ viết tắt	Viết đầy đủ	Ý nghĩa
EWMA	Exponential Weighted Moving Average	Công thức tính trung bình động
CUSUM	Cumulative SUM	Giải thuật dựa trên tổng tích lũy
AT	Adaptive Threshold	Giải thuật dựa trên ngưỡng giới hạn
SNMP	Simple Network Management Protocol	Tập hợp các giao thức cho phép kiểm tra các thiết bị mạng
IOS	Internetwork Operating System	Hệ điều hành mạng của Cisco
ICMP	Internet Control Message Protocol	Giao thức hoạt động trên layer 2
FAR	False Alarm Ratio	Tỉ lệ phát hiện cảnh báo sai
DP	Detection probability	Xác suất phát hiện tấn công

# GIỚI THIỆU ĐỀ TÀI VÀ ĐỊNH HƯỚNG GIẢI QUYẾT

## 1. Tổng quan

Khi càng nhiều công ty hoạt động dựa trên hệ thống mạng nội bộ, việc giám sát hệ thống cũng trở nên quan trọng hơn. Các sự cố gián đoạn hệ thống mạng, máy chủ không hoạt động, các dịch vụ và ứng dụng gặp vấn đề... đều gây ảnh hưởng nghiêm trọng đến hoạt động doanh nghiệp. Tổn thất có thể lên đến hàng ngàn, thậm chí hàng triệu USD tùy theo doanh thu và lĩnh vực hoạt động. Để duy trì độ ổn định và sẵn sàng cho hệ thống mạng nhằm phục vụ nhu cầu hoạt động liên tục, các doanh nghiệp ngày nay luôn phải đối mặt rất nhiều thách thức. Với những hệ thống mạng doanh nghiệp sử dụng các thiết bị của Cisco sẽ được tích hợp công cụ NetFlow đáp ứng được hầu hết các nhu cầu của người quản trị mạng. Phần mềm Cisco IOS NetFlow là một phần của họ các sản phẩm, tiện ích quản lý của Cisco. Các phần mềm được thiết kế đồng bộ kết hợp chặt chẽ với nhau để có thể quản lý kiểm tra giám sát hệ thống mạng một cách tốt nhất.

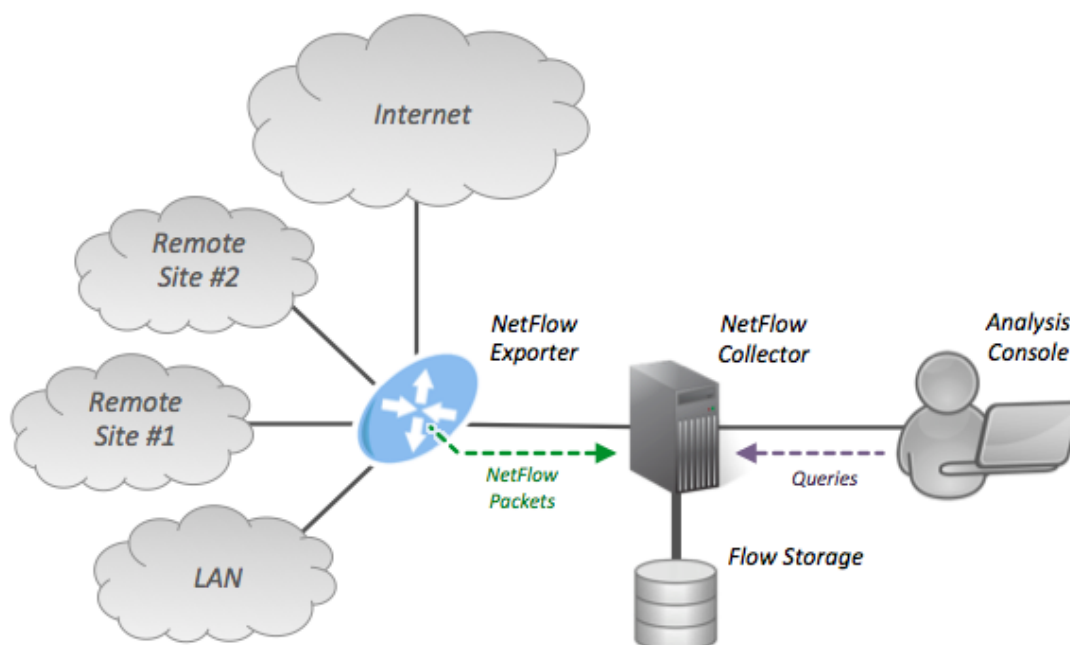
Đồ án thực hiện tìm hiểu cách thức hoạt động của NetFlow trong thiết bị của Cisco qua đó sử dụng dữ liệu trích xuất được để phân tích phát hiện bất thường mạng.

Đồ án được thực hiện trong quá trình thực tập và nghiên cứu tại phòng 901 tòa nhà B1, trường đại học Bách Khoa Hà Nội. Với môi trường thuận lợi và sự hướng dẫn nhiệt tình của PGS. TS. Ngô Hồng Sơn, tác giả đã thu được nhiều kiến thức quý báu trong thời gian thực tập tại đây.

## 2. Mục tiêu đề tài

NetFlow là tính năng của Cisco IOS cho phép thống kê lưu lượng gói qua bộ định tuyến. NetFlow thực hiện giám sát, phân tích, tính toán lưu lượng gói. Sử dụng phổ biến trong các yêu cầu sau:

- Giám sát mạng
- Giám sát ứng dụng
- Giám sát người dung
- Xây dựng kế hoạch phát triển mạng
- Phân tích an ninh mạng
- Tính toán lưu lượng



**Hình 1: Ví dụ về mô hình mạng sử dụng NetFlow**

Các nhiệm vụ cụ thể được tiến hành trong đồ án:

- Tìm hiểu công cụ nhúng NetFlow trong các thiết bị mạng Cisco
- Cấu hình và cài đặt thiết bị định tuyến Cisco trong hệ thống mạng
- Sử dụng dữ liệu NetFlow vào việc phân tích bất thường mạng
- Đề xuất giải thuật phân tích phát hiện bất thường mạng trong tấn công SYN-flood
- Cài đặt giải thuật phân tích phát hiện bất thường mạng trong tấn công SYN-flood sử dụng dữ liệu từ NetFlow
- Tiến hành chạy mô phỏng, đánh giá thuật toán và so sánh kết quả.

### **3. Định hướng giải quyết**

Sử dụng bộ định tuyến Cisco 1921 trong hệ thống mạng nhỏ, tiến hành cài đặt để trích xuất được dữ liệu NetFlow của lưu lượng mạng trong hệ thống mạng bởi bộ định tuyến.



**Hình 2: Hình ảnh của bộ định tuyến Cisco 1921**

Sử dụng bộ công cụ Flow-tool để trích xuất dữ liệu NetFlow được cài đặt trong máy giám sát.

Tiến hành thử nghiệm tấn công DoS hoặc DDoS vào hệ thống mạng sau đó lấy dữ liệu NetFlow để phân tích phát hiện bất thường mạng.

Áp dụng hai giải thuật AT và CUSUM thử nghiệm phát hiện tấn công SYN-flood và phân tích độ chính xác của hai thuật toán.

#### **4. Giới thiệu nội dung chính**

Đồ án được chia thành các phần chính như sau:

##### **CHƯƠNG I. CƠ SỞ LÝ THUYẾT LIÊN QUAN:**

Chương này trình bày kiến thức về NetFlow kiến thức về DoS cùng các thuật toán phát hiện bất thường mạng.

##### **CHƯƠNG II. MÔ HÌNH HỆ THỐNG VÀ KẾT QUẢ ĐẠT ĐƯỢC:**

Chương này mô tả hệ thống mạng đơn giản có sử dụng công cụ NetFlow và cài đặt thuật toán phân tích phát hiện SYN-flood.

##### **CHƯƠNG III. ĐÁNH GIÁ THUẬT TOÁN:**

Chương này đánh giá về hiệu quả các thuật toán sử dụng để phát hiện tấn công SYN-flood.

##### **KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN:**

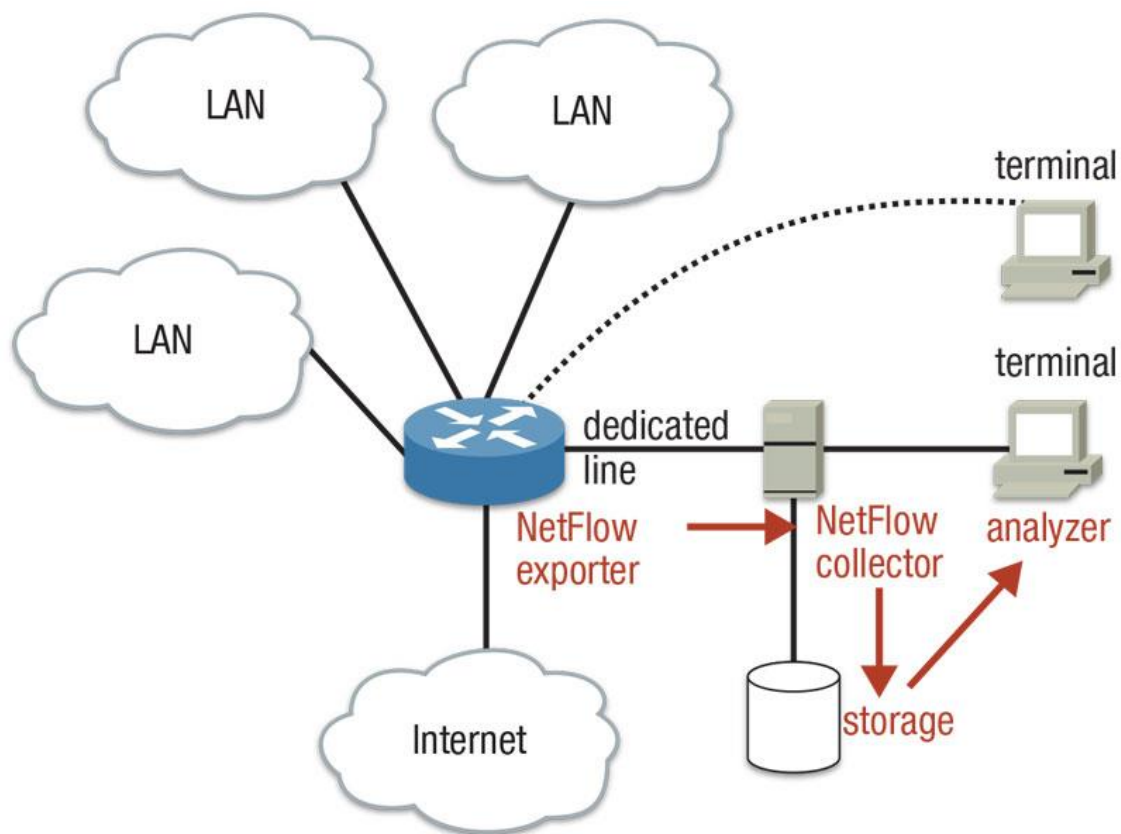
Trong phần này em xin phép trình bày đánh giá về hệ thống hiện tại cùng một vài hướng cải tiến mở rộng hệ thống chưa áp dụng được trong khuôn khổ đồ án.

# CHƯƠNG I. CƠ SỞ LÝ THUYẾT

## 1.1 Giới thiệu NetFlow

NetFlow là một công cụ nhúng trong các phần mềm IOS của Cisco để phân tích hoạt động của mạng. Đây là công cụ không thể thiếu cho người quản trị mạng chuyên nghiệp. Để đáp ứng lại các đòi hỏi và nhu cầu cấp bách của hệ thống mạng, việc tìm hiểu xem quá trình hoạt động của mạng ra sao là rất quan trọng.

NetFlow của Cisco có thể đáp ứng tất cả những yêu cầu trên. Nó tạo ra một môi trường mà người quản trị mạng có đầy đủ các công cụ để biết được thời gian, địa điểm, đối tượng cũng như cách thức lưu thông của lưu lượng mạng. Khi mà hoạt động của mạng được nắm vững thì hiệu quả của hệ thống mạng sẽ tăng lên rất nhiều.



**Hình 3: Các thành phần chính trong mô hình hoạt động của NetFlow**

Một mô hình giám sát lưu lượng mạng điển hình (sử dụng NetFlow) bao gồm ba thành phần chính:

**Flow exporter:** tập hợp các gói tin vào các dòng chảy và xuất dữ liệu thu thập được vào một hoặc nhiều Flow collector.

**Flow collector:** chịu trách nhiệm tiếp nhận, lưu trữ và thực hiện tiền xử lý dữ liệu thu thập được từ một hoặc nhiều Flow exporter.

**Analysis application:** phân tích dữ liệu được lưu trữ từ Flow collector.

### 1.1.1 Ứng dụng của NetFlow

NetFlow cho phép thống kê lưu lượng gói qua bộ định tuyến. NetFlow thực hiện giám sát, phân tích, tính toán lưu lượng gói và sử dụng phổ biến trong các yêu cầu sau:

**Giám sát mạng:** Cho phép giám sát hiện trạng mạng gần như thời gian thực. Giám sát mạng là kỹ thuật dựa vào flow (tập những gói có cùng 7 thông tin : IP nguồn, IP đích, Port nguồn, Port đích, ToS, loại giao thức lớp 3, cổng vào) nhằm thực hiện thu thập thông tin theo lưu lượng gói, theo luồng liên quan đến một thiết bị định tuyến, bộ chuyển mạch hoặc sự kết hợp của nhiều lưu lượng từ nhiều thiết bị giúp chủ động nhận diện được vấn đề, hiệu quả trong quá trình xử lý sự cố và đưa ra giải pháp giải quyết vấn đề một cách nhanh chóng.

**Giám sát ứng dụng:** Cho phép người quản trị nhìn thấy một cách chi tiết về hoạt động của ứng dụng trên mạng theo thời gian. Thông tin này được dùng để hiểu được những dịch vụ mới nhằm phân phối tài nguyên mạng (băng thông, chất lượng dịch vụ...) và tài nguyên cho ứng dụng cũng như là kế hoạch mở rộng.

**Giám sát người dùng:** Cho phép người vận hành mạng hiểu rõ tài nguyên mạng và tài nguyên ứng dụng mà người dùng sử dụng từ đó có kế hoạch phân phối tài nguyên một cách hợp lý cho người dùng, cũng như nhận diện được những vấn đề liên quan đến an ninh mạng hoặc vi phạm chính sách.

**Xây dựng kế hoạch phát triển mạng:** Do có khả năng giám sát và phân tích lưu lượng dữ liệu trong một khoảng thời gian dài, điều này cho phép người quản trị có cơ hội theo dõi, dự đoán sự phát triển của mạng để có kế hoạch nâng cấp như tăng số lượng bộ định tuyến, những cổng với băng thông lớn...

**Phân tích an ninh mạng:** NetFlow định danh và phân loại những loại tấn công như Dos, DDos, virus, worm theo thời gian thực dựa vào những sự hành vi thay đổi bất thường trong mạng.

**Tính toán lưu lượng:** Netflow cho phép người quản trị có được thông tin chi tiết lưu lượng dữ liệu như địa chỉ IP, ứng dụng, ToS, số lượng gói, số lượng byte, thời gian hoạt động giúp cho việc tính toán tài nguyên mạng được sử dụng theo người dùng, ứng dụng... trong khoảng thời gian cụ thể.

### 1.1.2 Tầm quan trọng của việc nhận thức hoạt động mạng

#### 1.1.2.1 Sự kiểm tra bằng SNMP truyền thống

Các khách hàng truyền thống thường hay sử dụng giao thức SNMP để kiểm tra hoạt động của băng thông. Mặc dù có một số ưu điểm nhất định nhưng SNMP khó có thể phân tích được thông tin các luồng trong lưu lượng mạng mà các ứng dụng sử dụng và mô hình trên mạng mà điều đó là rất cần thiết để biết được hệ thống mạng hỗ trợ công việc thế nào. Việc tìm hiểu cách thức sử dụng băng thông là một điều rất quan trọng trong hệ thống mạng IP ngày nay.

#### 1.1.2.2 Tầm quan trọng của NetFlow

Khả năng phân tích và giám sát được lưu lượng mạng trong hệ thống là hoạt động rất quan trọng đối với việc quản trị mạng. Việc kiểm tra dòng lưu lượng IP sẽ đảm bảo tài nguyên mạng được sử dụng một cách hợp lý hơn. Nó giúp người quản trị nhận biết được chất lượng dịch vụ (QoS), nhận biết được dấu hiệu hay nguy cơ của những cuộc tấn công từ chối dịch vụ (DoS), việc phát tán virus, cũng như hàng loạt các sự cố khác. NetFlow có thể giải quyết nhiều vấn đề đặt ra đối với một người quản trị chuyên nghiệp.

Phân tích các ứng dụng mới và ảnh hưởng của chúng lên hệ thống mạng: nhận dạng các ứng dụng mạng mới, ví dụ như VoIP chẳng hạn...

- Giảm sự quá tải của lưu lượng WAN.
- Xử lý sự cố và nhận biết được điểm yếu của hệ thống mạng.
- Phát hiện các lưu lượng WAN trái phép.
- Bảo mật hệ thống mạng, phát hiện được các sự cố bất thường.
- Phân chia băng thông hợp lý cho từng loại dịch vụ mạng khác nhau.

	SNMP	NetFlow
Cài đặt	Dễ dàng	Phức tạp (Bộ định tuyến có thể phải cấu hình lại)
Lọc traffic	Không	Có
Phân biệt bandwidth sử dụng giao thức hoặc IPs	Không	Có
Hiển thị toplists của PRTG (Top Talker, Top Connections, Top Protocols, etc.)	Không	Có
Lọc băng thông sử dụng IP	Không	Có
Lọc băng thông sử dụng MAC address	Không	Không
Lọc băng thông sử dụng physical network port	Có	Không



Giám sát các thông số mạng khác	Có	Không
CPU load trên các máy chạy PRTG	Thấp	Cao, phụ thuộc vào lượng traffic
Giám sát vượt quá băng thông	Nhỏ	Tùy thuộc vào traffic

**Bảng 1: Sự khác nhau giữa SNMP và NetFlow**

### 1.1.3 Cách thức hoạt động của NetFlow

NetFlow hoạt động bằng cách tạo ra một bộ đệm NetFlow trong đó chứa thông tin về tất cả các luồng đang hoạt động. Bộ đệm NetFlow được xây dựng trước hết bằng cách xử lý packet đầu tiên của một luồng thông qua một đường chuyển mạch chuẩn. Một bản ghi về luồng được duy trì bởi bộ đệm NetFlow cho tất cả luồng hoạt động. Mỗi một bản ghi luồng trong bộ đệm NetFlow chứa các trường thuộc tính có thể được sử dụng sau đó để xuất dữ liệu tới một thiết bị thu thập dữ liệu. Mỗi một bản ghi luồng được tạo ra bằng cách so sánh thuộc tính của các packet và đếm số packet và số byte của mỗi luồng.

#### 1. Create and update flows in NetFlow cache

Srdfs	Srd Padd	Dstlf	Dstl Padd	Protocol	TOS	Flgs	11000	00A2	Src Msk	Src AS	00A2	Dst Msk	Dst AS	Next Hop	Bytes/Pkt	Active	Idle
Fa1/0	173.100.21.2	Fa0/0	10.0.227.12	11	80	10	11000	00A2	/24	5	00A2	/24	15	10.023.2	1528	1745	4
Fa1/0	173.100.3.2	Fa0/0	10.0.227.12	6	40	0	2491	15	/26	196	15	/24	15	10.023.2	740	41.5	1
Fa1/0	173.100.20.2	Fa0/0	10.0.227.12	11	80	10	10000	00A1	/24	180	00A1	/24	15	10.023.2	1428	1145.5	3
Fa1/0	173.100.6.2	Fa0/0	10.0.227.12	6	40	0	2210	19	/30	180	19	/24	15	10.023.2	1040	1745	14

#### 2. Expiration

- Inactive timer expired (15 sec is default)
- Active timer expired (30 min (1800 sec) is default)
- NetFlow cache is full (oldest flows are expired)
- RST or FIN TCP Flag

Srdfs	Srd Padd	Dstlf	Dstl Padd	Protocol	TOS	Flgs	11000	00A2	Src Msk	Src AS	00A2	Dst Msk	Dst AS	Next Hop	Bytes/Pkt	Active	Idle
Fa1/0	173.100.21.2	Fa0/0	10.0.227.12	11	80	10	11000	00A2	/24	5	00A2	/24	15	10.023.2	1528	1800	4

#### 3. Aggregation

No Yes

#### 4. Export version

Non-Aggregated Flows—Export Version 5 or 9

e.g. Protocol-Port Aggregation Scheme Becomes

#### 5. Transport protocol

Export Packet

Payload (Flows)

Protocol	Pkts	SrcPort	DstPort	Bytes/Pkt
11	11000	00A2	DstPort	1528

Aggregated Flows — Export Version 8 or 9

**Hình 4: Ví dụ về một bộ đệm NetFlow**

### 1.1.4 Thông tin của luồng trong một bản ghi NetFlow

Một số thông tin cơ bản trong luồng NetFlow:

- Địa chỉ IP nguồn cho biết đối tượng đang phát sinh lưu lượng
- Địa chỉ IP đích cho biết đối tượng đang nhận lưu lượng
- Port cho biết loại ứng dụng đang sử dụng lưu lượng
- Lớp dịch vụ chiếm quyền ưu tiên lưu lượng
- Giao diện thiết bị cho biết cách thức sử dụng lưu lượng của thiết bị mạng
- Kiểm tra packet và byte cho biết độ lớn của lưu lượng
- Flow timestamps cho biết thời gian tồn tại của luồng; qua timestamps có thể biết tính toán được số packet và byte truyền đi trong mỗi giây.

- Địa chỉ IP hop kế tiếp
- Subnet mask của địa chỉ nguồn và đích
- TCP flag.

### 1.1.5 Cách truy cập dữ liệu tạo ra bởi NetFlow

Có 2 phương pháp chính để truy cập dữ liệu của NetFlow:

Thứ nhất đó là chế độ dòng lệnh CLI(Command Line Interface). Chế độ này có thể giúp phát hiện ra những thay đổi tức thì của mạng, điều đó rất có lợi cho việc xử lý sự cố xảy ra trong mạng.

Thứ hai đó là chế độ chuyển các thông tin NetFlow tới một máy chủ báo cáo gọi là “NetFlow collector”. NetFlow collector có nhiệm vụ thu thập thông tin về luồng và tổng hợp chúng lại để tạo ra một báo cáo về vấn đề lưu lượng cũng như phân tích an ninh mạng. Không giống như SNMP, NetFlow thường xuyên gửi thông tin một cách định kỳ tới NetFlow collector. Bộ đệm NetFlow liên tục được lấp đầy bởi các luồng và phần mềm trong bộ định tuyến hoặc bộ chuyển mạch sẽ tìm trong cache những luồng đã kết thúc và những luồng này sẽ được gửi ra máy chủ báo cáo. Luồng sẽ kết thúc khi giao tiếp mạng kết thúc. Lượng dữ liệu gửi tới NetFlow collector chỉ chiếm 1.5% lưu lượng chuyển mạch trong bộ định tuyến. Những ghi nhận chi tiết của NetFlow về từng gói cung cấp một cái nhìn đầy đủ và chi tiết về toàn bộ lưu lượng mạng đã chuyển qua bộ định tuyến hoặc thiết bị chuyển mạch.

### 1.1.6 Chi tiết thực hiện một báo cáo của NetFlow

Sau đây là các bước cơ bản để thực hiện một report của NetFlow:

- NetFlow được cấu hình để bắt luồng vào Bộ đệm NetFlow
- NetFlow export được cấu hình để gửi các luồng tới Collector
- Bộ đệm NetFlow tìm kiếm luồng đã kết thúc và gửi thông tin về luồng đó tới máy chủ báo cáo (NetFlow collector server).
- Có khoảng từ 1-30 luồng được đóng gói và gửi dưới dạng UDP tới máy chủ báo cáo.
- Phần mềm NetFlow collector tạo ra real-time và historical report từ dữ liệu

Cách thức bộ định tuyến hoặc bộ chuyển mạch quyết định luồng được gửi tới NetFlow Collector là một luồng sẵn sàng được xuất ra khi nó không hoạt động trong một khoảng thời gian nhất định hoặc luồng đã tồn tại (hoạt động) vượt quá thời gian hoạt động cho phép. Có một bộ đếm thời gian sẽ quyết định luồng là không hoạt động hay tồn tại quá lâu và thời gian mặc định cho luồng không hoạt động là trong 15s, còn thời gian mặc định giới hạn cho hoạt động của một luồng là 30 phút. Collector có thể kết hợp các luồng và đưa ra tổng hợp về lưu lượng mạng.

### 1.1.7 Vị trí của NetFlow trong mạng

NetFlow thường được sử dụng ở site trung tâm bởi tất cả lưu lượng mạng từ các site remote khác đều được phân tích và giám sát bởi NetFlow. Vị trí triển khai NetFlow phụ thuộc vào cấu trúc mạng. Nếu reporting collection server đặt ở vị trí trung tâm thì vị trí tối ưu nhất để cài đặt NetFlow chính là ở gần server đó.

### 1.1.8 Định dạng của dữ liệu gửi đi bởi NetFlow

Dữ liệu NetFlow gửi tới collector bao gồm một header và tuần tự các bản ghi tiếp theo. Phần header chứa thông tin về số thứ tự, số bản ghi và thời gian hệ thống. Các bản ghi luồng chứa thông tin về luồng, ví dụ như địa chỉ IP, port, thông tin định tuyến. Có các version khác nhau của Cisco NetFlow như 1, 5, 7, 8, 9 với các định dạng dữ liệu có sự khác nhau:

- Version 9: Sử dụng khi cần xuất dữ liệu từ nhiều công nghệ khác nhau như Multicast, DoS, IPv6, BGP nexthop, ... Định dạng có tính tương thích và mở rộng rất cao. Version 9 hỗ trợ xuất dữ liệu từ cả main cache và aggregation cache.
- Version 8: Chỉ hỗ trợ xuất dữ liệu từ aggregation cache
- Version 5: Chỉ hỗ trợ xuất dữ liệu từ main cache. Hầu hết các thiết bị truyền thống đều hỗ trợ định dạng này cũng được dùng phổ biến
- Version 1: Chỉ nên sử dụng khi hệ thống chỉ hỗ trợ version này. Nếu không nên sử dụng version 9 hoặc 5.

## 1.2 Giới thiệu bộ công cụ flow-tools

Flow-tools là bộ công cụ làm việc với dữ liệu của NetFlow. Flow-tools bao gồm các thư viện và tập hợp các chương trình được sử dụng để thu thập, gửi, xử lý, và tạo ra các báo cáo từ dữ liệu NetFlow. Bộ công cụ có thể được sử dụng trên cùng một máy chủ hoặc sử dụng cho nhiều máy chủ trong các hệ thống lớn. Các thư viện trong flow-tools cung cấp các hàm API để phát triển các ứng dụng tùy chỉnh cho version NetFlow 1, 5, 6 phù hợp.

Một số tool được sử dụng trong bộ công cụ NetFlow

- flow-capture: Thu thập, nén, lưu trữ và quản lý bộ nhớ đĩa cho dữ liệu xuất ra từ luồng NetFlow
- flow-cat: nối các file con. Thông thường các flow file sẽ bao gồm nhiều file nhỏ được xuất ra định kỳ cách nhau từ 5 đến 15 phút.
- flow-import: Import dữ liệu từ các định dạng ASCII hoặc cflowd.
- flow-export: Xuất dữ liệu ra dạng ASCII hoặc cflowd.
- flow-send: Gửi dữ liệu qua mạng sử dụng giao thức NetFlow.
- flow-split: Các tập tin flow đã lưu được chia thành các file nhỏ hơn dựa trên kích thước, thời gian, hoặc các thẻ tag.

Các bộ công cụ khác trong flow-tools bao gồm: flow-capture, flow-cat, flow-dscan, flow-expire, flow-export, flow-fanout, flow-filter, flow-nfilter, flow-gen, flow-header, flow-import, flow-merge, flow-print, flow-receive, flow-report, flow-send, flow-split, flow-stat, flow-tag, flow-xlate.

### **1.3 Tìm hiểu tấn công từ chối dịch vụ DoS**

#### **1.3.1 Giới thiệu về DoS**

Tấn công DoS là kiểu tấn công mà một người làm cho một hệ thống không thể sử dụng, hoặc làm cho hệ thống đó chậm đi một cách đáng kể với người dùng bình thường, bằng cách làm quá tải tài nguyên của hệ thống.

Nếu kẻ tấn công không có khả năng thâm nhập được vào hệ thống, thì chúng cố gắng tìm cách làm cho hệ thống đó sụp đổ và không có khả năng phục vụ người dùng bình thường đó là tấn công Denial of Service (DoS).

Mặc dù tấn công DoS không có khả năng truy cập vào dữ liệu thực của hệ thống nhưng nó có thể làm gián đoạn các dịch vụ mà hệ thống đó cung cấp. Như định nghĩa trên DoS khi tấn công vào một hệ thống mạng sẽ khai thác những cái yếu nhất của hệ thống để tấn công.

#### **1.3.2 Các mục đích của tấn công DoS**

- Cố gắng chiếm băng thông mạng và làm hệ thống mạng bị ngập (flood), khi đó hệ thống mạng sẽ không có khả năng đáp ứng những dịch vụ khác cho người dùng bình thường.
- Cố gắng làm ngắt kết nối giữa hai máy, và ngăn chặn quá trình truy cập vào dịch vụ.
- Cố gắng ngăn chặn những người dùng cụ thể vào một dịch vụ nào đó
- Cố gắng ngăn chặn các dịch vụ không cho người khác có khả năng truy cập vào.

#### **1.3.3 Mục tiêu mà kẻ tấn công thường sử dụng tấn công DoS**

Như chúng ta biết ở bên trên tấn công DoS xảy ra khi kẻ tấn công sử dụng hết tài nguyên của hệ thống và hệ thống không thể đáp ứng cho người dùng bình thường được vậy các tài nguyên chúng thường sử dụng để tấn công là gì:

- Tạo ra sự khan hiếm, những giới hạn và không đổi mới tài nguyên
- Băng thông của hệ thống mạng (Network Bandwidth), bộ nhớ, ổ đĩa, và CPU
- Time hay cấu trúc dữ liệu đều là mục tiêu của tấn công DoS.
- Phá hoại hoặc thay đổi các thông tin cấu hình.
- Phá hoại tầng vật lý hoặc các thiết bị mạng như nguồn điện, điều hoà...

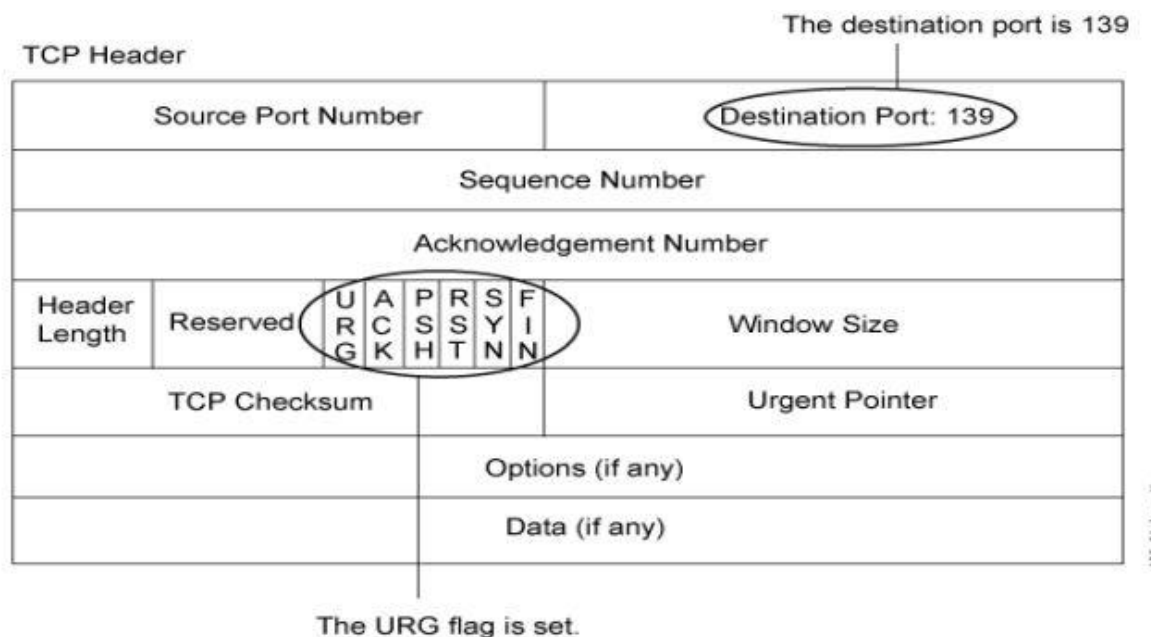
#### **1.3.4 Dấu hiệu khi bị tấn công DoS**

- Thông thường thì hiệu suất mạng sẽ rất chậm.

- Không thể sử dụng website.
- Không truy cập được bất kỳ website nào.
- Tăng lượng thư rác nhanh chóng.

### 1.3.5 Các kỹ thuật tấn công DoS

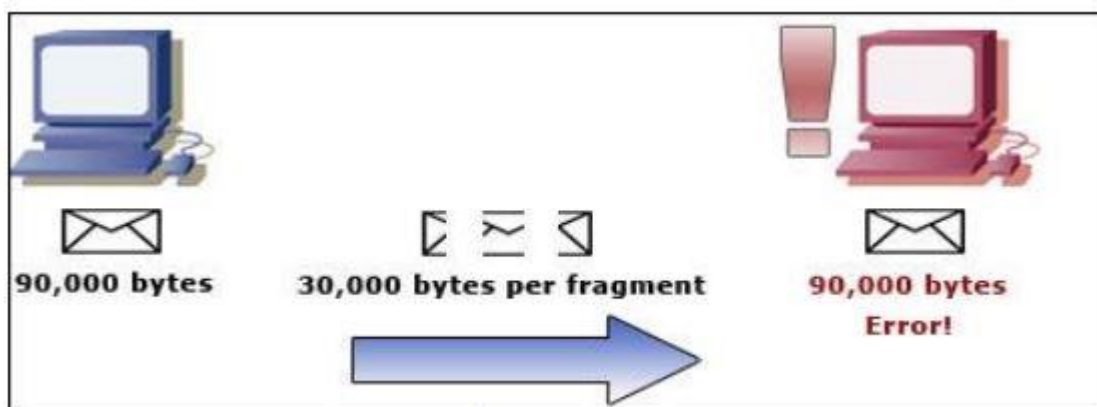
#### 1.3.5.1 Winnuke



**Hình 5: Gói tin TCP Header được Attacker sử dụng**

DoS attack loại này chỉ có thể áp dụng cho các máy tính đang chạy Windows9x. Hacker sẽ gửi các gói tin với dữ liệu "Out of Band" đến cổng 139 của máy tính đích. (Cổng 139 chính là cổng NetBIOS, cổng này chỉ chấp nhận các gói tin có cờ Out of Band được bật). Khi máy tính của nạn nhân nhận được gói tin này, một màn hình xanh báo lỗi sẽ được hiển thị lên với nạn nhân do chương trình của Windows nhận được các gói tin này nhưng nó lại không biết phản ứng với các dữ liệu Out Of Band như thế nào dẫn đến hệ thống sẽ bị crash.

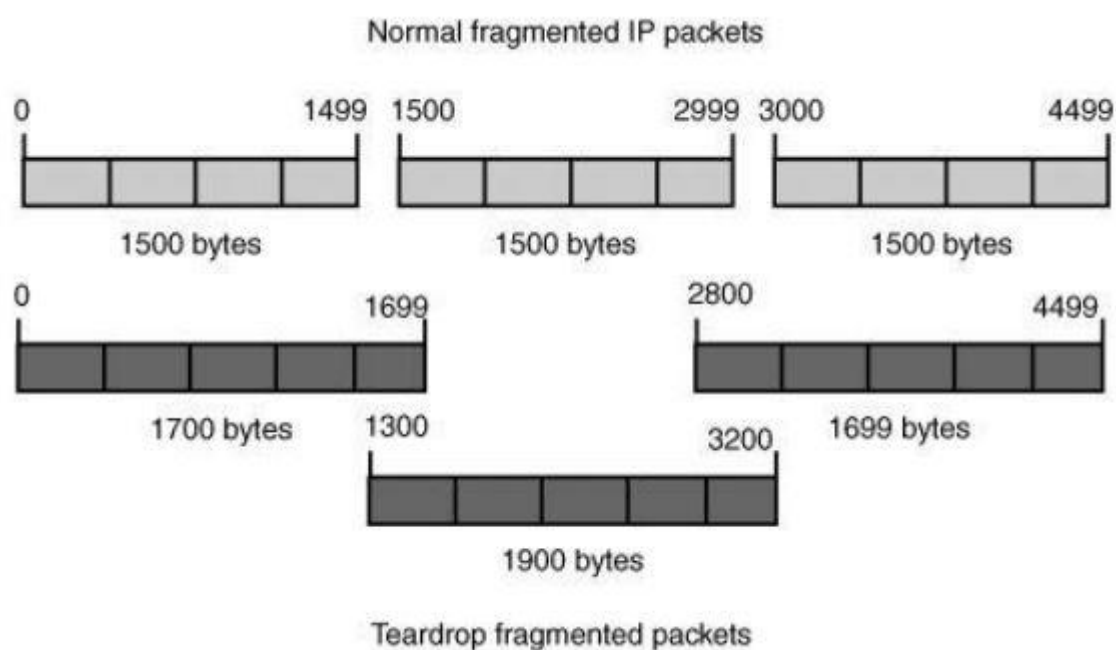
### 1.3.5.2 Ping of Death



**Hình 6: Mô hình tấn công Ping Of Death**

Tấn công Ping of Death (hay PoD) có thể làm tê liệt cả mạng lưới dựa trên lỗi hỏng của hệ thống TCP/IP. Kích thước tối đa cho 1 gói dữ liệu là 65,535 bytes. Nếu ta gửi các gói tin lớn hơn nhiều so với kích thước tối đa thông qua lệnh “ping” đến máy đích thì sẽ làm máy tính đích bị treo. Nhưng gửi 1 gói tin lớn hơn kích thước quy định là điều trái với luật của giao thức TCP/IP, vì vậy Hacker đã khéo léo gửi các gói tin trên các đoạn phân mảnh. Khi máy tính nạn nhân ráp các phân mảnh dữ liệu thì sẽ nhận thấy gói tin quá lớn. Điều này sẽ gây ra lỗi tràn bộ đệm và treo các thiết bị. Nhưng đến nay thì hầu hết các thiết bị được sản xuất sau năm 1998 đã miễn dịch với loại tấn công này.

### 1.3.5.3 Teardrop



**Hình 7: Cơ chế tấn công bằng Teardrop**

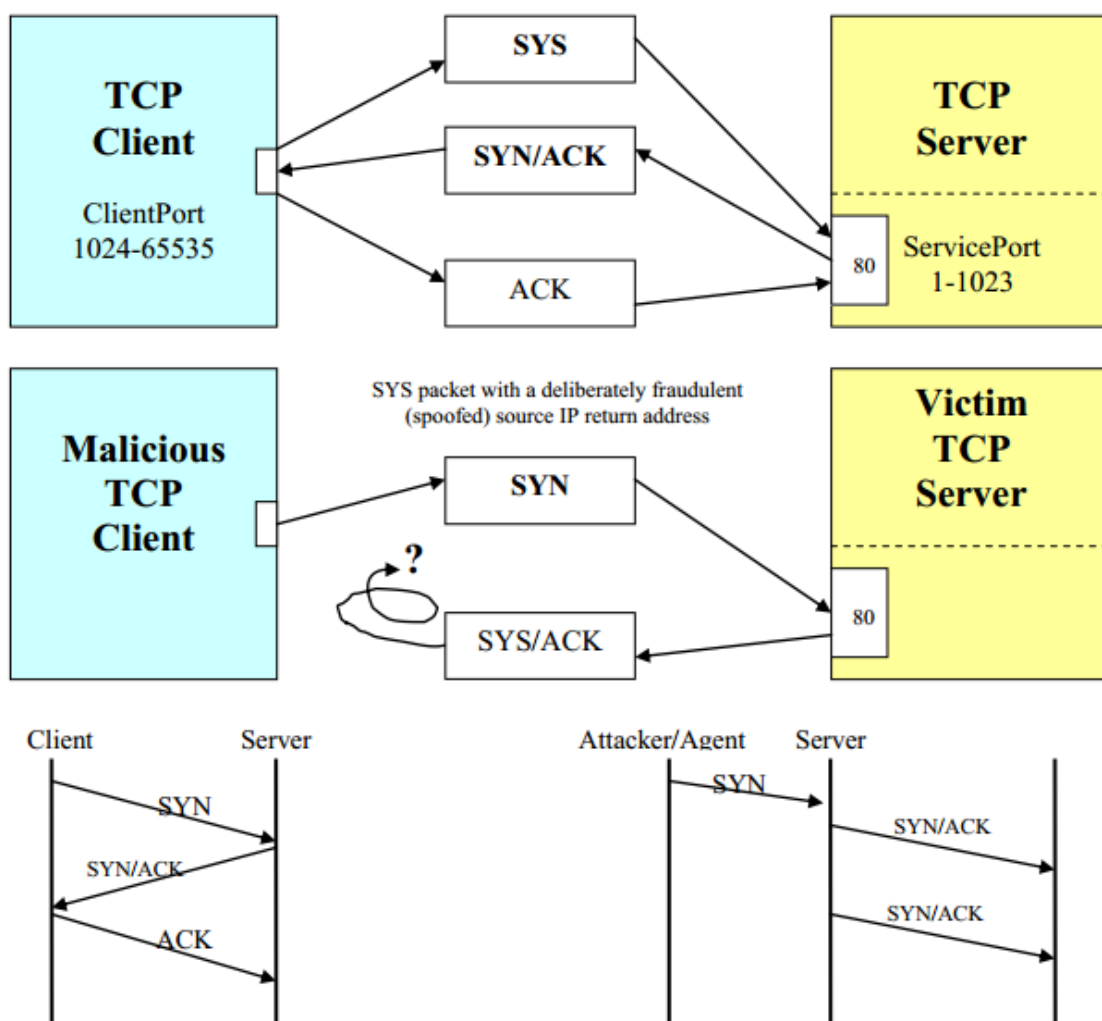
Như ta đã biết, tất cả các dữ liệu chuyển đi trên mạng từ hệ thống nguồn đến hệ thống đích đều phải trải qua 2 quá trình: dữ liệu sẽ được chia ra thành các mảnh nhỏ ở hệ thống nguồn, mỗi mảnh đều phải có một giá trị offset nhất định để xác định vị trí của mảnh đó trong gói dữ liệu được chuyển đi. Khi các mảnh này đến hệ thống đích, hệ thống đích sẽ dựa vào giá trị offset để sắp xếp các mảnh lại với nhau theo thứ tự đúng như ban đầu. Lợi dụng sơ hở đó, ta chỉ cần gửi đến hệ thống đích một loạt gói packets với giá trị offset chồng chéo lên nhau. Hệ thống đích sẽ không thể nào sắp xếp lại các packets này, nó không điều khiển được và có thể bị crash, reboot hoặc ngừng hoạt động nếu số lượng gói packets với giá trị offset chồng chéo lên nhau quá lớn.

#### ***1.3.5.4 SYN Attack***

Kiểu tấn công TCP SYN flood là một kiểu tấn công trực tiếp vào máy chủ bằng cách tạo ra một số lượng lớn các kết nối TCP nhưng không hoàn thành các kết nối này.

Trong SYN Attack, hacker sẽ gửi đến hệ thống đích một loạt SYN packets với địa chỉ IP nguồn không có thực. Hệ thống đích khi nhận được các SYN packets này sẽ gửi trở lại các địa chỉ không có thực đó và chờ đợi để nhận thông tin phản hồi từ các địa chỉ IP giả.

Vì đây là các địa chỉ IP không có thực, nên hệ thống đích sẽ chờ đợi vô ích và còn đưa các "request" chờ đợi này vào bộ nhớ, gây lãng phí một lượng đáng kể bộ nhớ trên máy chủ mà đúng ra là phải dùng vào việc khác thay cho phải chờ đợi thông tin phản hồi không có thực này. Nếu ta gửi cùng một lúc nhiều gói tin có địa chỉ IP giả như vậy thì hệ thống sẽ bị quá tải dẫn đến bị crash hoặc boot máy tính.



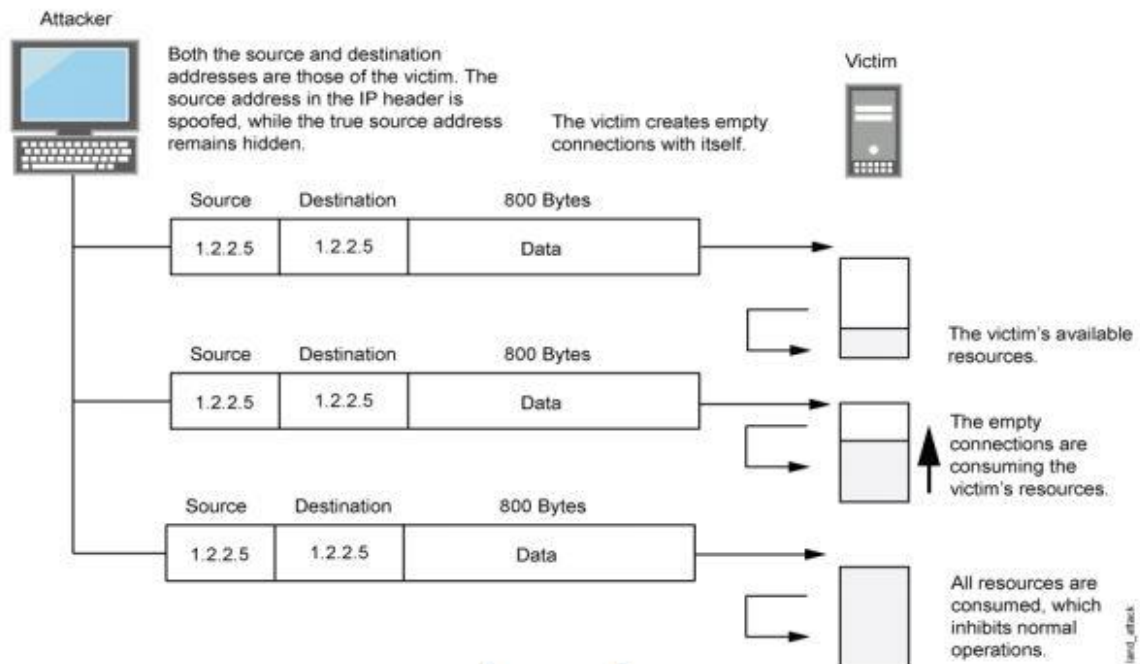
**Hình 8: Mô hình tấn công bằng SYN Flood Attack**

Bit offset	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	Source port																Destination port															
32	Sequence number																															
64	Acknowledgment number																															
96	Data offset				Reserved				C	E	U	A	P	R	S	F	Window Size															
									W	R	E	C	K	H	T	N																
128	Checksum																Urgent pointer															
160	Options (if Data Offset > 5)																															
...	...																															

**Hình 9: Cấu tạo gói tin TCP**



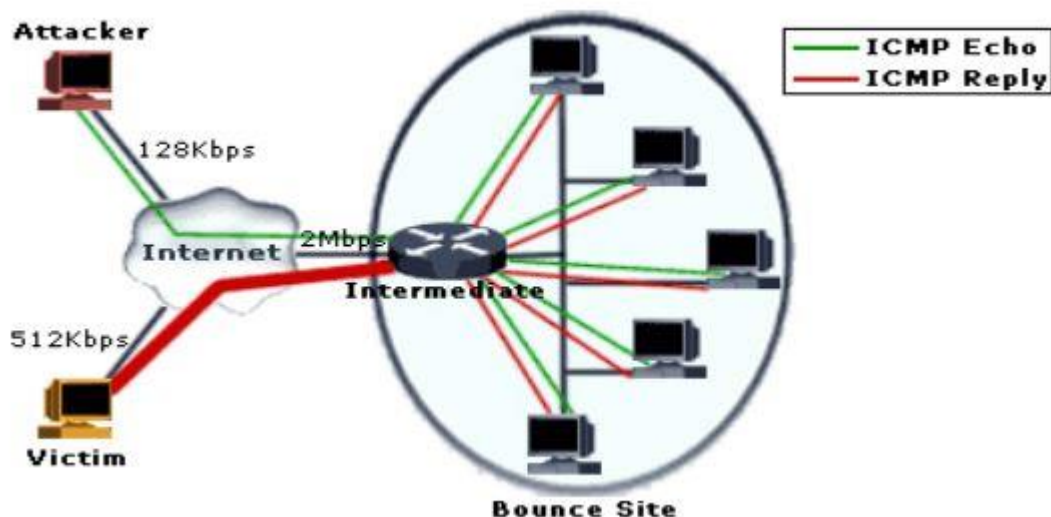
### 1.3.5.5 Land Attack



**Hình 10: Mô hình tấn công bằng Land Attack**

Land Attack cũng gần giống như SYN Attack, nhưng thay vì dùng các địa chỉ IP không có thực, hacker sẽ dùng chính địa chỉ IP của hệ thống nạn nhân. Điều này sẽ tạo nên một vòng lặp vô tận giữa trong chính hệ thống nạn nhân đó, giữa một bên cần nhận thông tin phản hồi còn một bên thì chẳng bao giờ gởi thông tin phản hồi đó đi cả.

### 1.3.5.6 SmurfAttack



**Hình 11: Mô hình tấn công Smuft Attack**

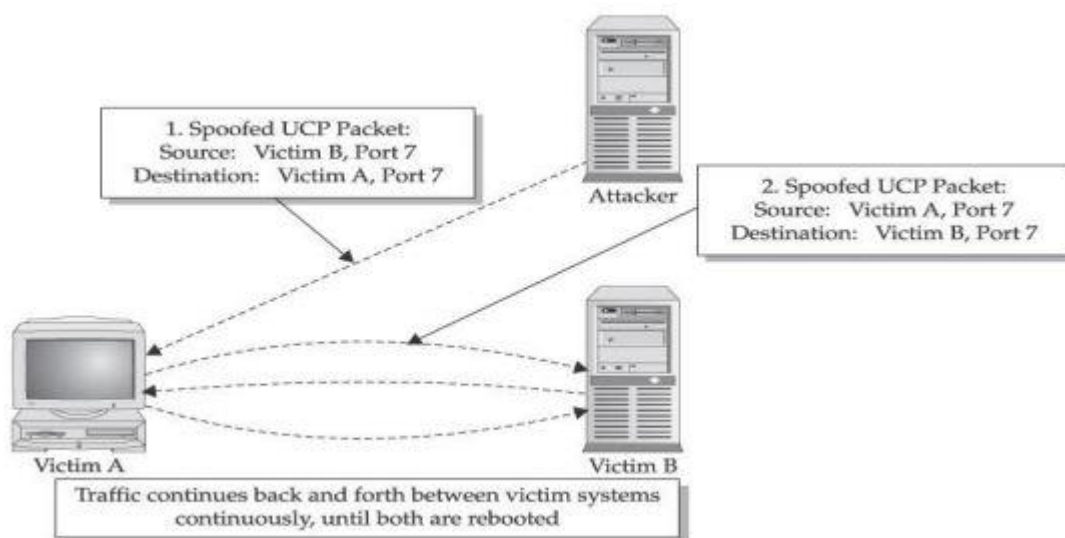
Trong Smurf Attack, cần có ba thành phần: hacker (người ra lệnh tấn công), mạng khuếch đại (sẽ nghe lệnh của hacker) và hệ thống của nạn nhân. Hacker sẽ gởi

các gói tin ICMP đến địa chỉ broadcast của mạng khuếch đại. Điều đặc biệt là các gói tin ICMP packets này có địa chỉ IP nguồn chính là địa chỉ IP của nạn nhân.

Khi các packets đó đến được địa chỉ broadcast của mạng khuếch đại, các máy tính trong mạng khuếch đại sẽ tưởng rằng máy tính nạn nhân đã gửi gói tin ICMP packets đến và chúng sẽ đồng loạt gửi trả lại hệ thống nạn nhân các gói tin phản hồi ICMP packets. Hệ thống máy nạn nhân sẽ không chịu nổi một khối lượng khổng lồ các gói tin này và nhanh chóng bị ngừng hoạt động, crash hoặc reboot.

Như vậy, chỉ cần gửi một lượng nhỏ các gói tin ICMP packets đi thì hệ thống mạng khuếch đại sẽ khuếch đại lượng gói tin ICMP packets này lên gấp bội. Tỷ lệ khuếch đại phụ thuộc vào số mạng tính có trong mạng khuếch đại. Nhiệm vụ của các hacker là cố chiếm được càng nhiều hệ thống mạng hoặc router cho phép chuyển trực tiếp các gói tin đến địa chỉ broadcast không qua chỗ lọc địa chỉ nguồn ở các đầu ra của gói tin. Có được các hệ thống này, hacker sẽ dễ dàng tiến hành Smurf Attack trên các hệ thống cần tấn công.

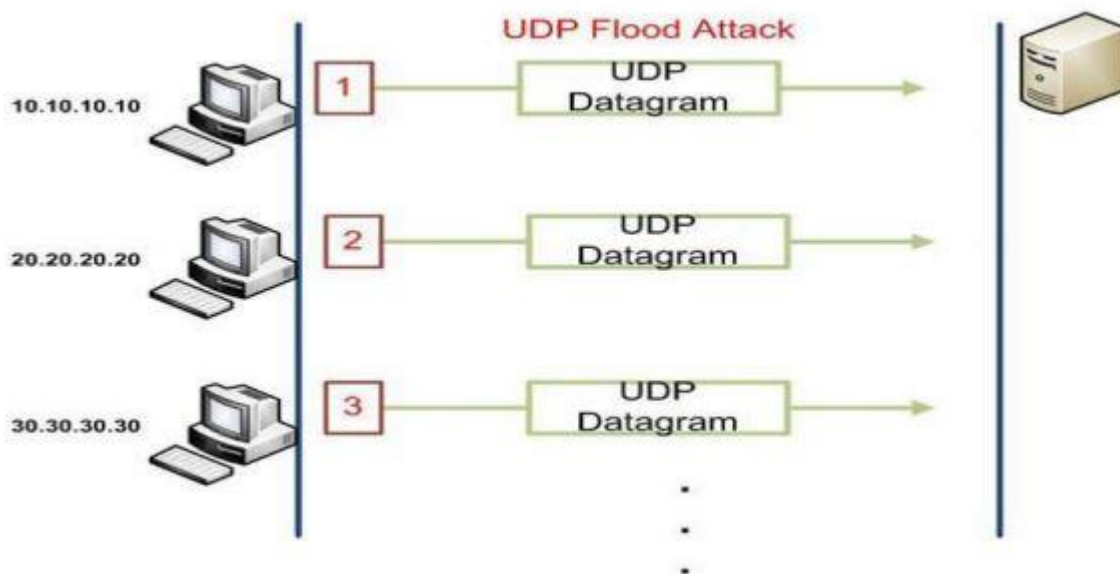
#### 1.3.5.7 Fraggle Attack



**Hình 12: Mô hình tấn công bằng Fraggle Attack**

Tương tự như Smurt attack nhưng thay vì dùng gói tin ICMP ECHO REQUEST thì sẽ dùng cách tấn công này sẽ dùng gói tin UDP ECHO gửi đến mục tiêu. Nhưng Flaggle Attack nguy hiểm hơn Smurt attack rất nhiều. Vì Attacker tấn công bằng một gói tin ECHO REQUEST với địa chỉ bên nhận là một địa chỉ broadcast, toàn bộ hệ thống thuộc địa chỉ này lập tức gửi gói tin REPLY đến port echo của nạn nhân, sau đó từ máy nạn nhân một gói tin ECHO REPLY lại gửi trở về địa chỉ broadcast, và quá trình cứ thế tiếp diễn.

### 1.3.5.8 UDP Flooding



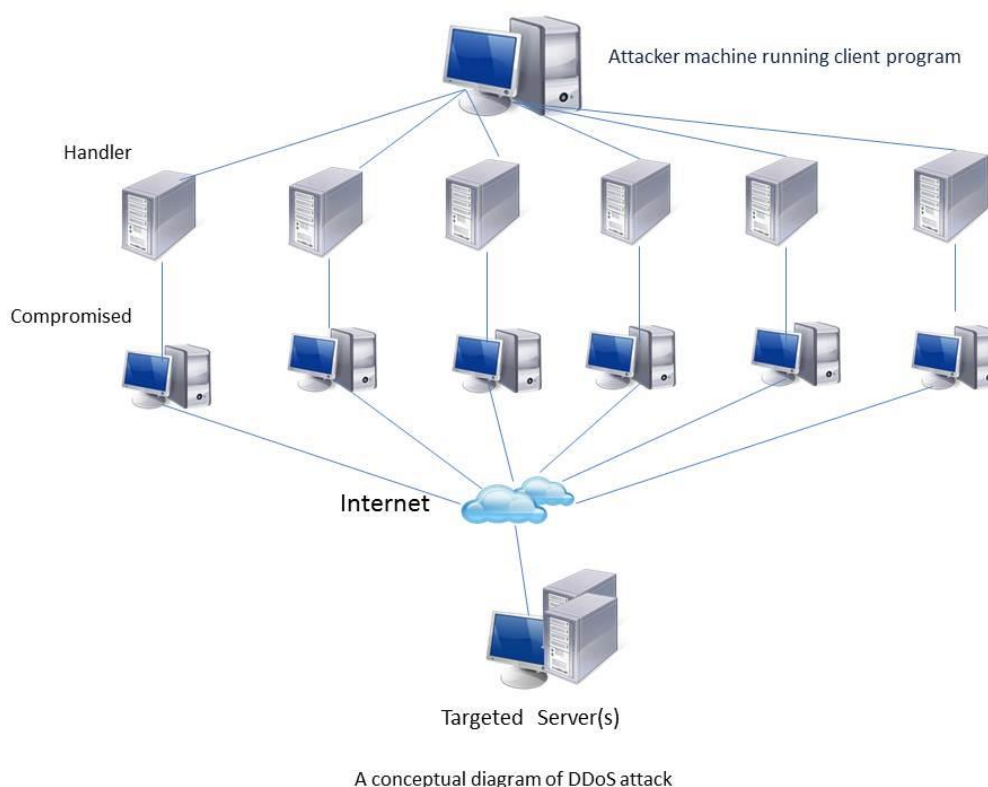
**Hình 13: Mô hình tấn công bằng UDP Flood**

Cách tấn công UDP đòi hỏi phải có 2 hệ thống máy cùng tham gia. Hackers sẽ làm cho hệ thống của mình đi vào một vòng lặp trao đổi các dữ liệu qua giao thức UDP. Và giả mạo địa chỉ IP của các gói tin là địa chỉ loopback (127.0.0.1), rồi gửi gói tin này đến hệ thống của nạn nhân trên cổng UDP echo. Hệ thống của nạn nhân sẽ trả lời lại các messages do 127.0.0.1 (chính nó) gửi đến, kết quả là nó sẽ đi vòng một vòng lặp vô tận. Tuy nhiên, có nhiều hệ thống không cho dùng địa chỉ loopback nên hacker sẽ giả mạo một địa chỉ IP của một máy tính nào đó trên mạng nạn nhân và tiến hành ngập lụt UDP trên hệ thống của nạn nhân. Nếu bạn làm cách này không thành công thì chính máy của bạn sẽ bị đầy.

### 1.3.5.9 Tấn công DNS

Hacker có thể đổi một lỗi vào trên Domain Name Server của hệ thống nạn nhân rồi cho chỉ đến một website nào đó của hacker. Khi máy khách yêu cầu DNS phân tích địa chỉ bị xâm nhập thành địa chỉ IP, lập tức DNS (đã bị hacker thay đổi cache tạm thời) sẽ đổi thành địa chỉ IP mà hacker đã cho chỉ đến đó. Kết quả là thay vì phải vào trang Web muốn vào thì các nạn nhân sẽ vào trang Web do chính hacker tạo ra. Một cách tấn công từ chối dịch vụ thật hữu hiệu.

### 1.3.5.10 Distributed DoS Attacks (DDoS)



**Hình 14: Mô hình tấn công DDoS**

DDoS yêu cầu phải có ít nhất vài kẻ tấn công cùng tham gia. Đầu tiên những kẻ tấn công sẽ cố thâm nhập vào các mạng máy tính được bảo mật kém, sau đó cài lên các hệ thống này chương trình DDoS server. Bây giờ những kẻ tấn công sẽ hẹn nhau đến thời gian đã định sẽ dùng DDoS client kết nối đến các DDoS servers, sau đó đồng loạt ra lệnh cho các DDoS servers này tiến hành tấn công DDoS đến hệ thống nạn nhân.

## 1.4 Giới thiệu các giải thuật phát hiện tấn công SYN-flood

Thực tế đã chứng minh, khi các cuộc tấn công Dos xảy ra. Lập tức phân tích sẽ thấy được lưu lượng mạng rất khác thường. Do đó hầu hết các thuật toán phân tích phát hiện tấn công Dos hiện nay đều dựa trên tính khác thường của lưu lượng mạng. Một số các công nghệ thống kê được áp dụng để tiến hành phân tích, thống kê những lưu lượng tải làm việc để phát hiện. Từ những kỹ thuật phân tích này, sẽ có những thuật toán phát hiện để đưa ra các tham số hoặc công nghệ thống kê, các mức độ nguy hiểm của cuộc tấn công.

Sau đây, tôi sẽ giới thiệu tổng quan về hai thuật toán phát hiện DoS hiện nay được ứng dụng nhiều trong phát hiện tấn công SYN-Flood. Nếu muốn tìm hiểu kỹ hơn về các thuật toán nêu ở dưới thì có thể tham khảo trong [1]

### 1.4.1 Thuật toán AT (Adaptive Threshold)

Thuật toán dựa trên việc kiểm tra xem các phép đo lưu lượng, số lượng gói tin SYN trong một khoảng thời gian nhất định vượt quá một ngưỡng cụ thể. Các giá trị của ngưỡng được thiết lập phù hợp dựa trên ước tính về số lượng trung bình của các gói tin SYN trong khoảng thời gian nhất định.

Gọi  $x_n$  là số lượng gói tin SYN trong khoảng thời gian  $n$ , và  $\bar{\mu}_{n-1}$  là giá trị trung bình động của các gói tin ở thời điểm  $n - 1$  ta có:

$$x_n \geq (\alpha + 1)\bar{\mu}_{n-1}$$

Nếu  $x_n \geq (\alpha + 1)\bar{\mu}_{n-1}$  tín hiệu cảnh báo sẽ xuất hiện tại thời điểm  $n$ , với  $\alpha > 0$  là một tham số cho biết ngưỡng giới hạn số lượng gói tin SYN  $x_n$  trong khoảng thời gian  $n$  xuất hiện mà không coi là có dấu hiệu tấn công SYN-flood. Giá trị trung bình động  $\mu_n$  có thể được tính toán dựa trên công thức EWMA (exponential weighted moving average)

$$\bar{\mu}_n = \beta\bar{\mu}_{n-1} + (1 - \beta)x_n$$

với  $\beta$  là tham số của công thức EWMA

Tuy nhiên, khi áp dụng trực tiếp thuật toán trên thì tỉ lệ phát hiện sai khá cao. Để cải thiện hiệu quả thuật toán ta sẽ thay đổi cách phát hiện tấn công sao cho chỉ phát hiện tấn công khi số lượng gói tin  $x_n$  vượt quá ngưỡng giới hạn cho phép trong  $k$  khoảng thời gian liên tiếp:

$$\sum_{i=n-k+1}^n 1_{\{x_i \geq (\alpha+1)\bar{\mu}_{i-1}\}} \geq k$$

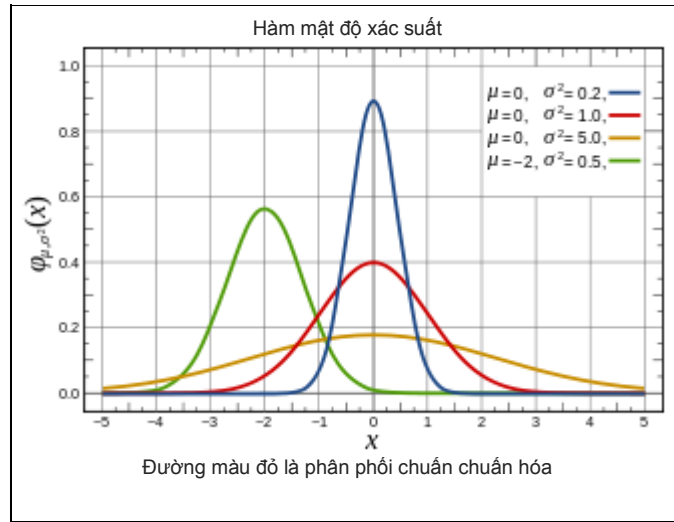
Với  $k > 1$  là một tham số cho biết số lượng của các khoảng thời gian liên tiếp mà ngưỡng giới hạn bị vi phạm. Khi đó phát hiện tấn công SYN-flood sẽ xuất hiện. Các thông số điều chỉnh của các công thức ở trên bao gồm biên độ  $\alpha$  để tính ngưỡng báo động, số lượng  $k$  khoảng thời gian, tham số trong công thức tính EWMA, và độ dài của khoảng thời gian đo lưu lượng (số lượng gói dữ liệu SYN) được thực hiện.

### 1.4.2 Thuật toán CUSUM (Cumulative SUM)

#### 1.4.2.1 Phân phối chuẩn

Phân phối chuẩn, còn gọi là phân phối Gauss, là một phân phối xác suất cực kì quan trọng trong nhiều lĩnh vực. Nó là họ phân phối có dạng tổng quát giống nhau, chỉ khác tham số vị trí (giá trị trung bình  $\mu$ ) và tỉ lệ (phương sai  $\sigma^2$ ).

Phân phối chuẩn chuẩn hóa (standard normal distribution) là phân phối chuẩn với giá trị trung bình bằng 0 và phương sai bằng 1 (đường cong màu đỏ trong hình bên phải). Phân phối chuẩn còn được gọi là đường cong chuông (bell curve) vì đồ thị của mật độ xác suất có dạng chuông.



**Bảng 2: Ví dụ đồ thị phân phối chuẩn**

#### 1.4.2.2 Giới thiệu thuật toán CUSUM

Thuật toán Cumulative sum (CUSUM) được sử dụng để phát hiện các cuộc tấn công SYN-flood dựa trên lưu lượng SYN được dự báo trước.

Trong quá trình phát hiện các cuộc tấn công SYN-flood, với mỗi gói tin SYN, CUSUM sẽ giám sát tập  $n$  gói SYN tạo thành mẫu gói tin  $\{y_1, \dots, y_n\}$ . Với  $y_n$  là tổng số gói SYN trong khoảng thời gian mẫu  $n$ -th (khoảng thời gian phát hiện). Giả sử sự biến đổi lưu lượng SYN  $\{y_i\}$  tuân theo phân phối Gauss độc lập với phương sai  $\sigma^2$  đã biết, và không thay đổi sau khi có bất thường mạng, ta có  $\mu_0$  và  $\mu_1$  là giá trị lưu lượng trung bình SYN trước và sau khi thay đổi.

Giải thuật CUSUM ( $f_n$ ) có thể được mô tả như sau:

$$f_n = \left[ f_{n-1} + \frac{\mu_1 - \mu_0}{\sigma^2} \left( y_n - \frac{\mu_1 + \mu_0}{2} \right) \right]^+$$

Với giả định phân phối Gaussian về  $\{y_i\}$  có thể không đúng đối với các phép đo gói tin TCP SYN, do những biến động số lượng gói tin hàng tuần hoặc hàng ngày, xu hướng và tương quan thời gian. Chính vì thế, trạng thái biến đổi không ổn định như vậy cần được loại bỏ trước khi áp dụng CUSUM. Hơn nữa để việc tính toán đỡ phức tạp và tốn thời gian, chúng ta xem xét một cách tiếp cận đơn giản để áp dụng CUSUM cho  $\tilde{x}_n$ , với:

$$\tilde{x}_n = x_n - \bar{\mu}_{n-1}$$

Ta có  $x_n$  là tổng các gói SYN trong khoảng thời gian mẫu  $n$ -th, và  $\bar{\mu}_n$  là giá trị trung bình động ước tính của gói tin SYN tại mẫu  $n$ , được tính bằng công thức exponential weighted moving average (EWMA) như sau:

$$\bar{\mu}_n = \beta \bar{\mu}_{n-1} + (1 - \beta)x_n$$

Lưu lượng trung bình của  $x_n$  trước thay đổi là 0, do đó giá trị trung bình ở (2) là  $\mu_0 = 0$ . Một vấn đề còn lại cần được giải quyết là giá trị của  $\mu_1$ , là số lượng trung bình của gói SYN sau khi thay đổi. Điều này có thể không được biết trước, vì thế ta ước lượng gần đúng với  $\alpha \bar{\mu}_n$ , với  $\alpha$  là tham số phần trăm biên độ, tương ứng với tỷ lệ mà ở đó gần như có thể xảy ra sự tăng về số lượng trung bình của gói tin SYN sau một sự thay đổi bất thường (tấn công).

Từ đó giải thuật CUSUM có thể viết lại như sau:

$$g_n = \left[ g_{n-1} + \frac{\alpha \bar{\mu}_{n-1}}{\sigma^2} (x_n - \bar{\mu}_{n-1} - \frac{\alpha \bar{\mu}_{n-1}}{2}) \right]^+$$

Nếu  $g_n \geq h$  ( $h > 0$  là tham số ngưỡng), cảnh báo sẽ xảy ra.

Các tham số điều chỉnh của thuật toán CUSUM là biên độ tỷ lệ tham số  $\alpha$ , ngưỡng báo động  $h$ , nhân tố  $\beta$  trong công thức EWMA, và độ dài của khoảng thời gian đo lưu lượng được thực hiện. Các tham số  $\alpha$  và  $\beta$  tương tự như tham số trong giải thuật Adaptive Threshold.

#### 1.4.2.3 Nhận xét chung về các thuật toán

Cả hai thuật toán được sử dụng nhiều trong các công trình nghiên cứu phát hiện bất thường mạng không chỉ áp dụng cho riêng tấn công SYN-flood mà còn trong nhiều dạng tấn công khác dựa vào sự thay đổi bất thường trong lưu lượng mạng.

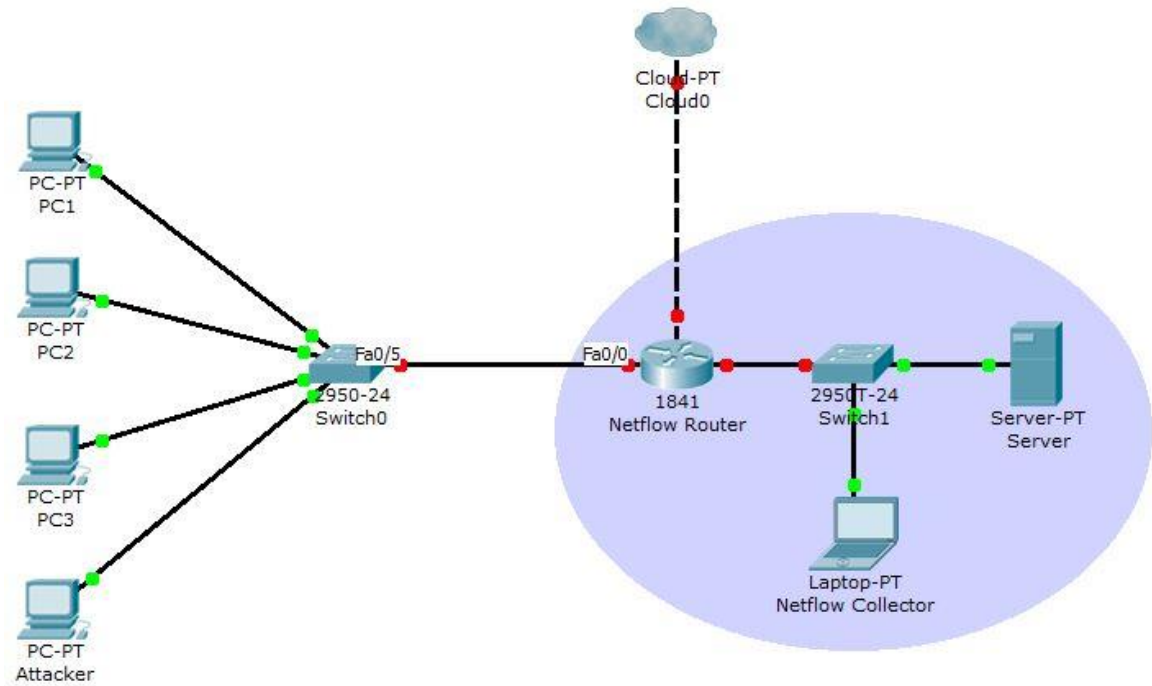
Thuật toán Adaptive Threshold nói chung khá đơn giản và dễ hiểu. Thuật toán phát hiện sự không bình thường dựa trên sự vi phạm của một ngưỡng khả năng đáp ứng của lưu lượng mạng trong thời gian gần. Thuật toán đặc biệt có khả năng phát hiện cao nhất khi kẻ tấn công tiến hành một cuộc tấn công TCP SYN. Thuật toán tin tưởng vào việc kiểm tra phép đo lưu lượng có vượt qua một ngưỡng giới hạn cụ thể hay không. Nếu vượt qua, chứng tỏ đã có một cuộc tấn công xảy ra.

Thuật toán CUSUM dựa trên giá trị trung bình của một quá trình xử lý thống kê. Sự phát hiện điểm thay đổi cần phải theo dõi trong các khoảng thời gian. Một công thức được xây dựng để theo dõi sự thay đổi này, khi vượt qua một ngưỡng giới hạn chứng tỏ đã xảy ra một cuộc tấn công.



## CHƯƠNG II. MÔ HÌNH HỆ THỐNG

### 2.1 Xây dựng hệ thống mạng giám sát bằng NetFlow



**Hình 15: Mô hình mạng thử nghiệm**

Các thành phần trong hệ thống mạng bao gồm:

- Bộ định tuyến Cisco Router 1921 có hỗ trợ NetFlow
- Máy chủ Web Server có thể trở thành host server trên Internet
- Máy giám sát NetFlow Collector
- Các máy client PC1, PC2, ... có thể nhiều hơn
- Máy tấn công Attacker

Chi tiết cài đặt, cấu hình hệ thống giám sát NetFlow được trình bày trong phần phụ lục phía cuối đồ án.



## 2.2 Mô phỏng tấn công SYN-flood

Tiến hành cài đặt thuật toán phát hiện bất thường trong tấn công SYN-flood, loại tấn công phổ biến nhất trong các kiểu tấn công từ chối dịch vụ (DoS).

Tiến hành đánh giá xác suất phát hiện, tỷ lệ báo động giả, và sự chậm trễ trong phát hiện. Đánh giá sự ảnh hưởng bởi các thông số của thuật toán và đặc điểm của các cuộc tấn công đến kết quả. Qua đó có thể điều chỉnh các thông số của các thuật toán sao cho phù hợp với đặc thù lưu lượng mạng.

### 2.2.1 Sử dụng công cụ flow-tools để xử lý dữ liệu cho NetFlow

Câu lệnh dùng để lưu trữ data NetFlow vào folder flowfilestore, khi luồng NetFlow được cấu hình xuất qua cổng 5000

```
./flow-capture -w /flowfilestore 0/0/5000
```

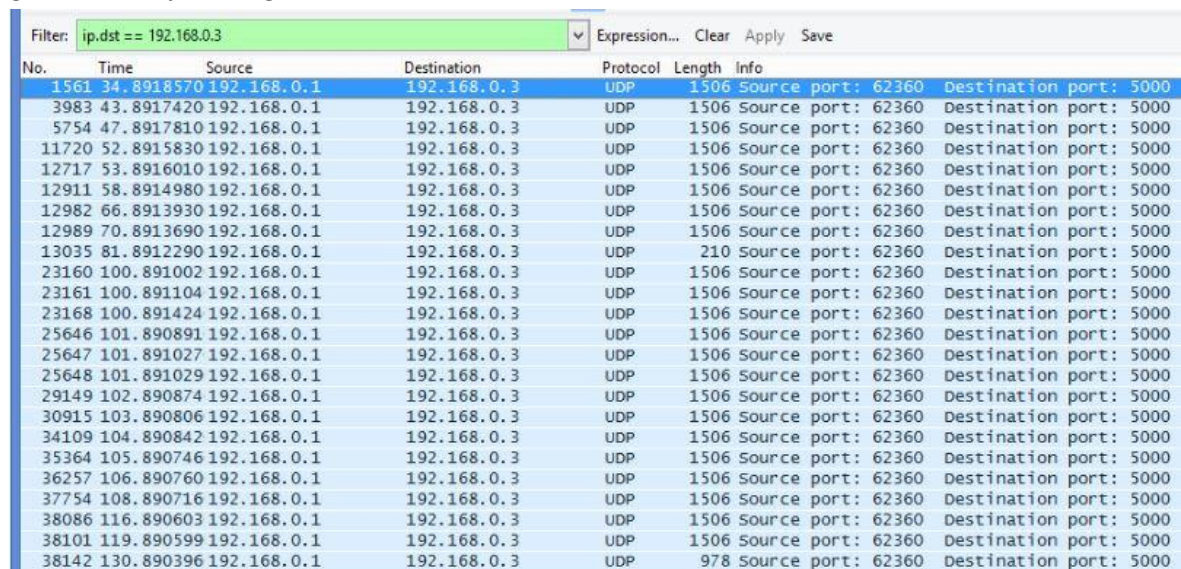
Câu lệnh dùng để ghép toàn bộ các file flow con tạo ra trong quá trình thu thập dữ liệu flow thành 1 file duy nhất có tên “newfile.flows”

```
./flow-cat -p -z9 /flowfilestore > newfile.flows
```

Sử dụng câu lệnh này để xuất ra định dạng ascii từ tệp NetFlow nhị phân, file định dạng ascii được lưu dưới tên “flows.ascii”

```
./flow-export -f2 -m0x3000 < flows > flows.ascii
```

Có thể kiểm tra bằng Wireshark để xem dữ liệu NetFlow gửi đến địa chỉ máy giám sát hay không:



No.	Time	Source	Destination	Protocol	Length	Info
1561	34.8918570	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
3983	43.8917420	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
5754	47.8917810	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
11720	52.8915830	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
12717	53.8916010	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
12911	58.8914980	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
12982	66.8913930	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
12989	70.8913690	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
13035	81.8912290	192.168.0.1	192.168.0.3	UDP	210	Source port: 62360 Destination port: 5000
23160	100.891002	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
23161	100.891104	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
23168	100.891424	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
25646	101.890891	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
25647	101.891027	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
25648	101.891029	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
29149	102.890874	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
30915	103.890806	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
34109	104.890842	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
35364	105.890746	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
36257	106.890760	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
37754	108.890716	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
38086	116.890603	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
38101	119.890599	192.168.0.1	192.168.0.3	UDP	1506	Source port: 62360 Destination port: 5000
38142	130.890396	192.168.0.1	192.168.0.3	UDP	978	Source port: 62360 Destination port: 5000

Hình 16: Ví dụ các gói UDP NetFlow gửi đến máy giám sát

### 2.2.2 Cài đặt thuật toán

Ta sử dụng dữ liệu từ định dạng NetFlow phiên bản v5 để phân tích lưu lượng mạng. Trong một bản ghi NetFlow ta dựa vào các trường “Layer 4 Protocol”, “TCP flags” và “Packet Count” để tiến hành đếm các gói tin SYN xuất hiện trong trong giao thức TCP.

Dưới đây là thông tin chi tiết các trường trong một packet Netflow phiên bản 5:

Bytes	Contents	Description
0-1	version	Version NetFlow sử dụng
2-3	count	Số lượng packet được xuất ra trong 1 flow (1-30)
4-7	sys_uptime	Thời gian hệ thống từ khi thiết bị bắt đầu export
8-11	unix_secs	Thời gian khi 1 flow được export tính theo hệ thống thời gian UTC 1970
12-15	unix_nsecs	Thời gian tính theo nano giây 0000 UTC 1970
16-19	flow_sequence	Trình tự truy cập của flow
20	engine_type	Loại flow-switching engine
21	engine_id	Số slot của flow-switching engine
22-23	sampling_interval	Hai bit đầu tiên giữ chế độ lấy mẫu; còn lại 14 bit giữ giá trị của khoảng thời gian lấy mẫu

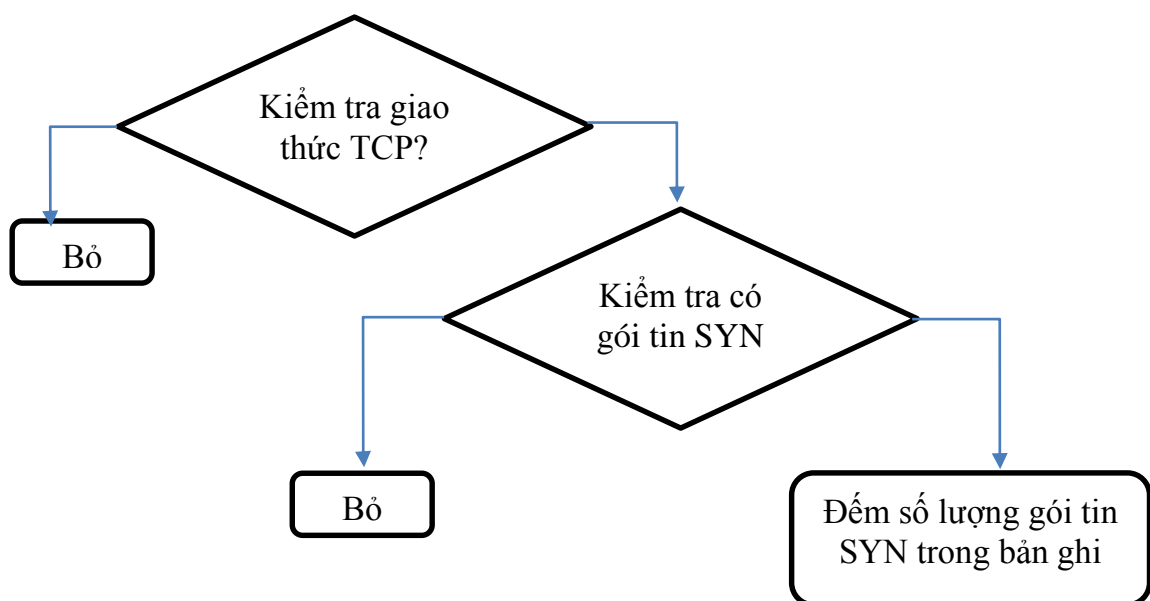
**Bảng 3: Thông tin của flow header format**

Bytes	Contents	Description
0-3	srcaddr	Địa chỉ IP nguồn IP
4-7	dstaddr	Địa chỉ IP đích IP
8-11	nexthop	Địa chỉ IP next-hop của bộ định tuyến
12-13	input	Chỉ số SNMP trong input interface
14-15	output	Chỉ số SNMP trong output interface
16-19	dPkts	Số packets trong 1 flow
20-23	dOctets	Tổng số byte của Layer 3 trong gói dữ liệu của flow
24-27	first	SysUptime khi bắt đầu flow
28-31	last	SysUptime tại thời điểm gói tin cuối cùng của dòng chảy được nhận

32-33	srcport	Tên cổng nguồn TCP/UDP
34-35	dstport	Tên cổng đích TCP/UDP
36	pad1	Byte dự trữ (thường là 0)
37	tcp_flags	Mô tả các flags trong bản ghi
38	prot	Loại giao thức IP (ví dụ TCP = 6; UDP = 17)
39	tos	Loại dịch vụ IP(ToS)
40-41	src_as	Autonomous system number of the source, either origin or peer
42-43	dst_as	Autonomous system number of the destination, either origin or peer
44	src_mask	Các bit mask của địa chỉ nguồn
45	dst_mask	Các bit mask của địa chỉ đích
46-47	pad2	Byte dự trữ (thường là 0)

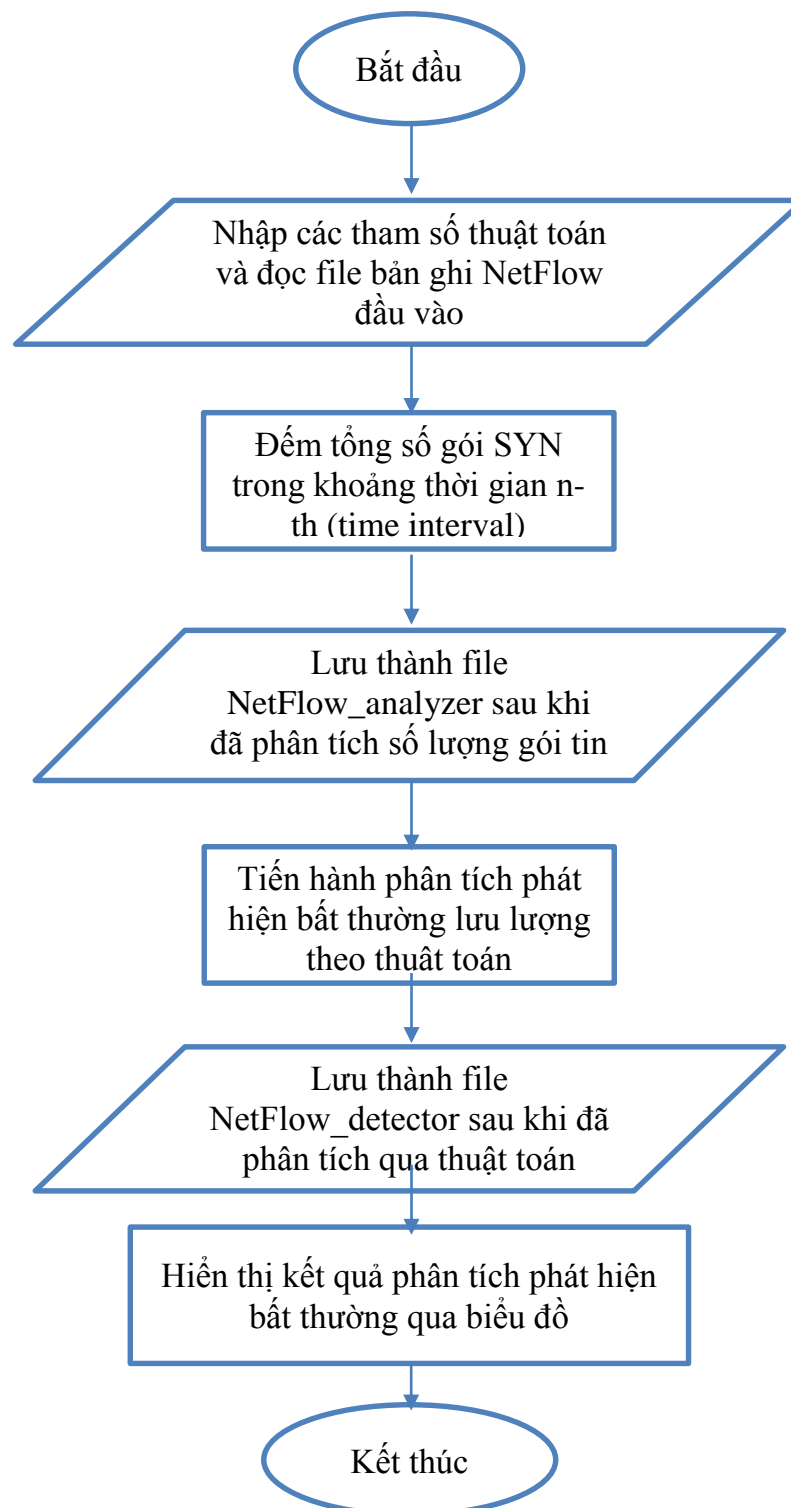
**Bảng 4: Thông tin trong flow record format**

Các bản ghi NetFlow có trường “TCP flags” = {2, 18, 19, 26, 27, 30, 31} tương ứng với bản ghi có chứa cờ SYN.



**Hình 17: Lược đồ phân tích tìm gói tin SYN trong bản ghi NetFlow**

Dưới đây là lược đồ minh họa chương trình đề xuất:

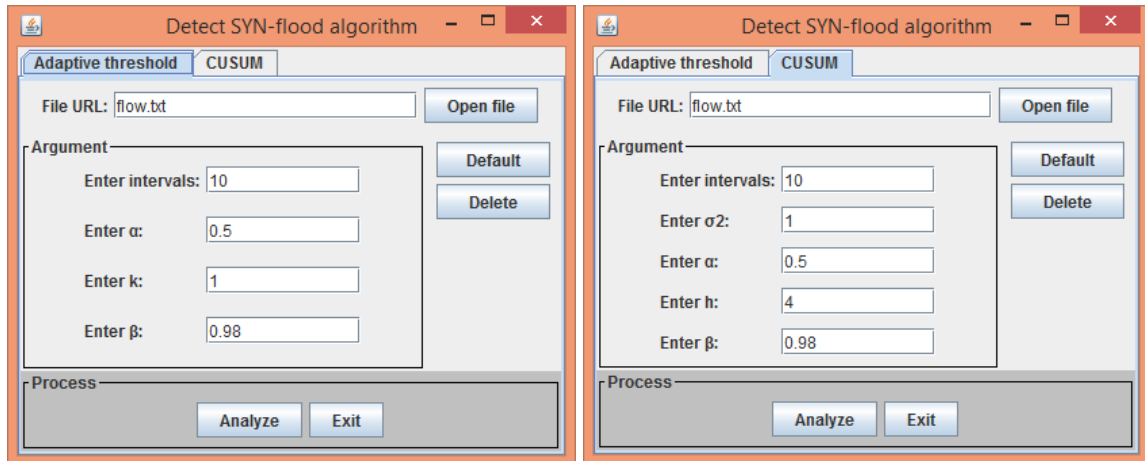


**Hình 18: Lược đồ mô tả chương trình**

### 2.2.3 Tổ chức chương trình

Cấu trúc chương trình:

- Chương trình được viết bằng ngôn ngữ lập trình java.
- Toàn bộ chương trình chính nằm trong folder detect-SYN-flood.
- Các thư viện dùng trong đồ án: bộ thư viện jfreechart.
- Đầu vào là file lưu dữ liệu NetFlow data định dạng text.
- Để xuất ra biểu đồ phân tích phát hiện bất thường mạng cần nhập các tham số của thuật toán phù hợp.



**Hình 19: Giao diện chương trình**

## CHƯƠNG III. ĐÁNH GIÁ THUẬT TOÁN

### 3.1 Bài toán mô phỏng

Thực hiện lấy dữ liệu trong vòng 60 phút, trong đó cách 7 phút tiến hành tấn công SYN-flood trong vòng 3 phút. Dữ liệu nhị phân được thu thập bởi bộ định tuyến Cisco 1921 đóng vai trò là một NetFlow Exporter được đưa đến máy tính cài đặt Flow-tools đóng vai trò như một NetFlow Collector. Sau đó toàn bộ dữ liệu sẽ được chuyển đổi thành file text và đưa máy có cài đặt chương trình thuật toán đóng vai trò như một NetFlow Analyzer.

### 3.2 Đánh giá hiệu quả thuật toán

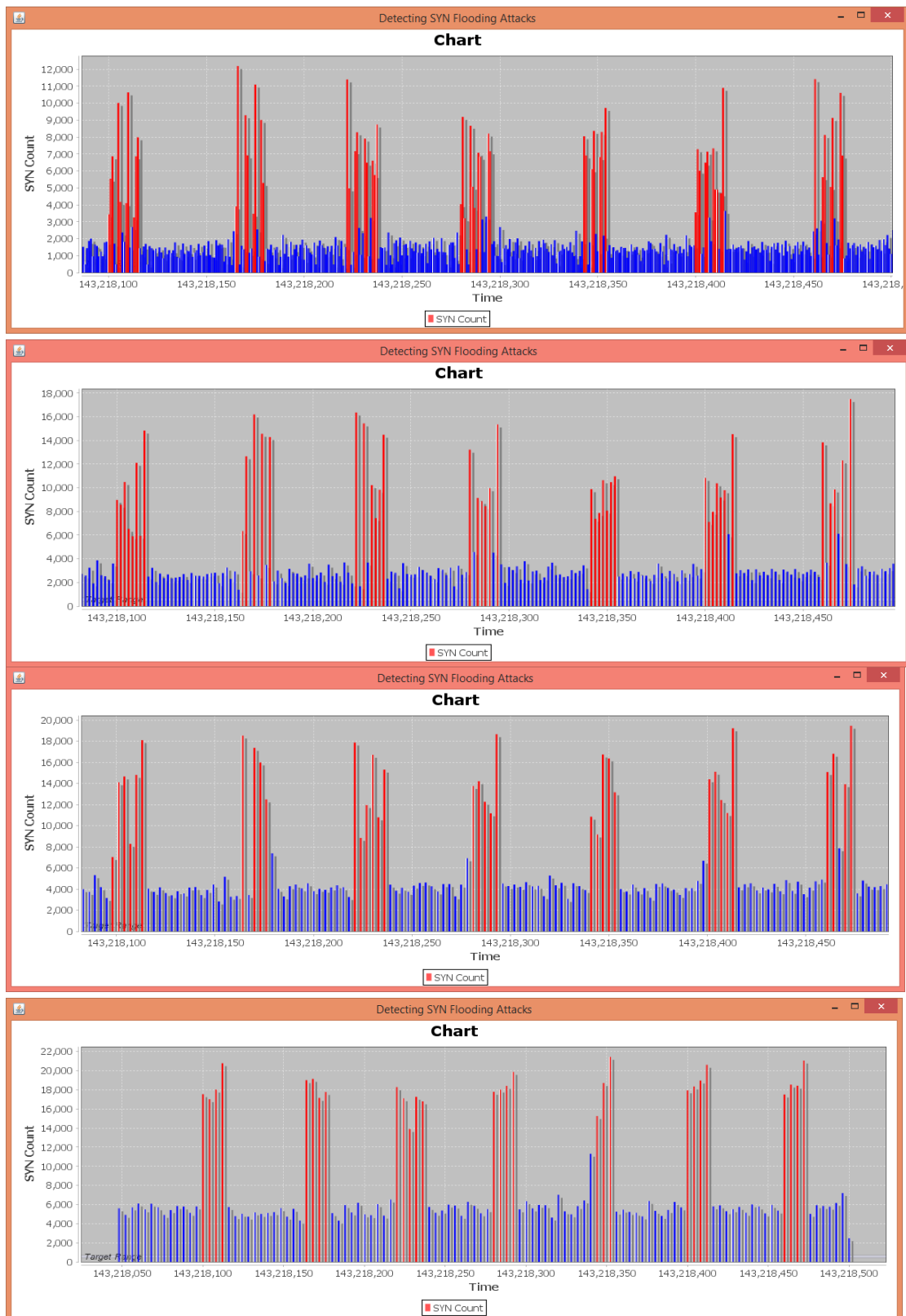
Ta sẽ đánh giá hiệu quả của thuật toán dựa trên hai tỉ lệ sau :

- Tỉ lệ phát hiện (detection probability) là tỉ lệ phần trăm khoảng thời gian phát hiện tấn công SYN-flood so với khoảng thời gian tấn công thực
- Tỉ lệ phát hiện sai (false alarm ratio) là tỉ lệ phần trăm khoảng thời gian phát hiện tấn công SYN-flood nhầm so với khoảng thời gian tấn công thực

#### 3.2.1 Tấn công cường độ cao

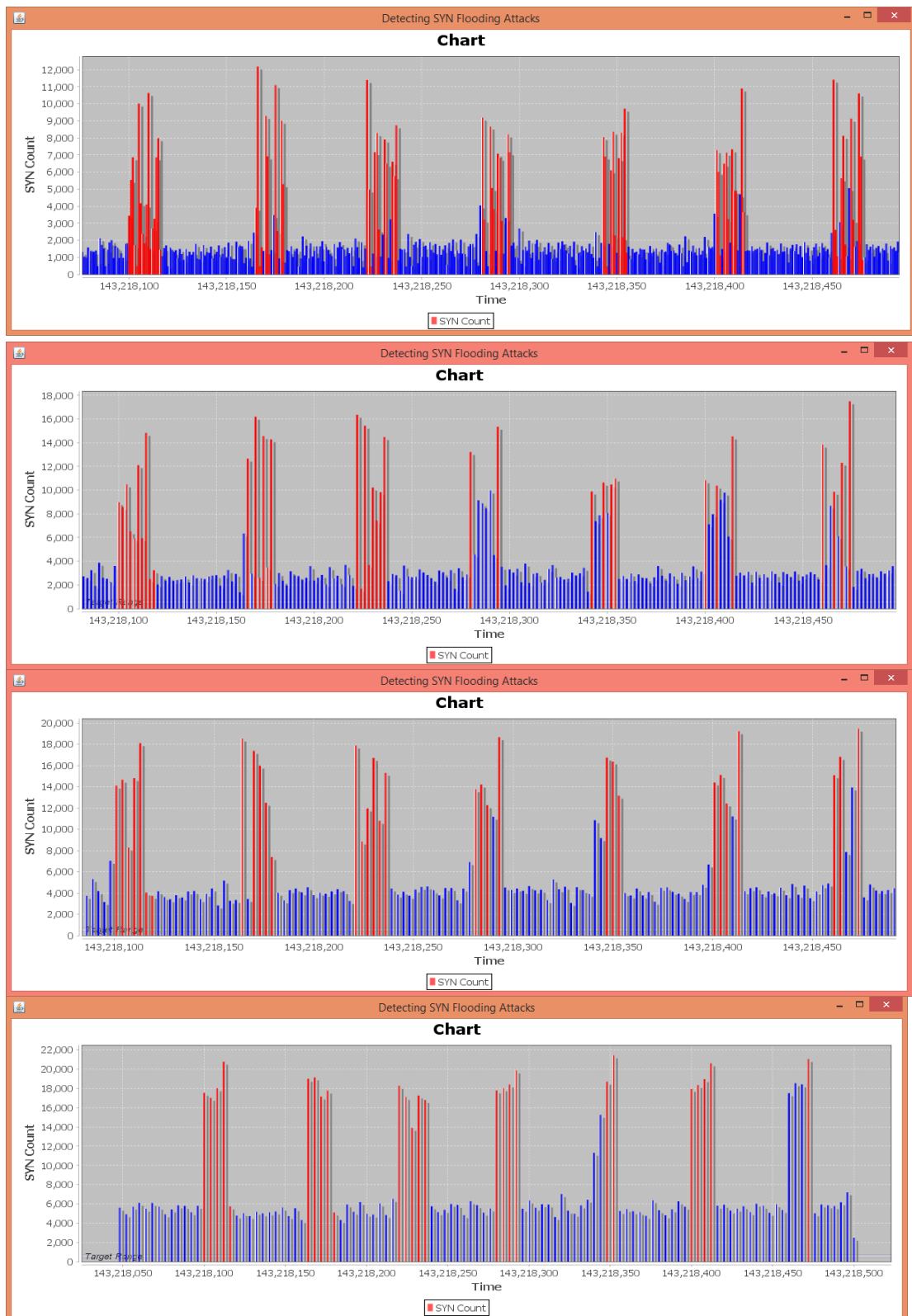
Thử nghiệm tấn công SYN-flood với cường độ 250% so với lưu lượng nhận gói tin SYN trong trường hợp mạng bình thường

Cài đặt thuật toán Adaptive Threshold với khoảng thời gian lấy mẫu là 10 giây, các tham số thử nghiệm lần lượt là  $\alpha = 0.5$ ,  $k = 1$ , và  $\beta = 0.98$



**Hình 20: Biểu đồ phát hiện tấn công SYN-flood sử dụng thuật toán AT với time intervals lần lượt là 10, 20, 30, 40 giây**

Cài đặt thuật toán CUSUM với khoảng thời gian lấy mẫu là 10 giây, các tham số thử nghiệm lần lượt là  $\sigma^2 = 1$ ,  $\alpha = 2.5$ ,  $h = 4$ , và  $\beta = 0.98$



**Hình 21: Biểu đồ phát hiện tấn công SYN-flood sử dụng thuật toán CUSUM với time intervals lần lượt là 10, 20, 30, 40 giây**

Tỉ lệ phát hiện sai FAR của cả hai thuật toán xấp xỉ 0% trong khi đó tỉ lệ phát hiện tấn công được thể hiện dưới bảng sau:



Time intervals (s)	AT	CUSUM
10	48%	65%
20	70%	68%
30	83%	76%
40	90%	83%

**Bảng 5: Tỷ lệ DP của hai thuật toán với tấn công cường độ cao**

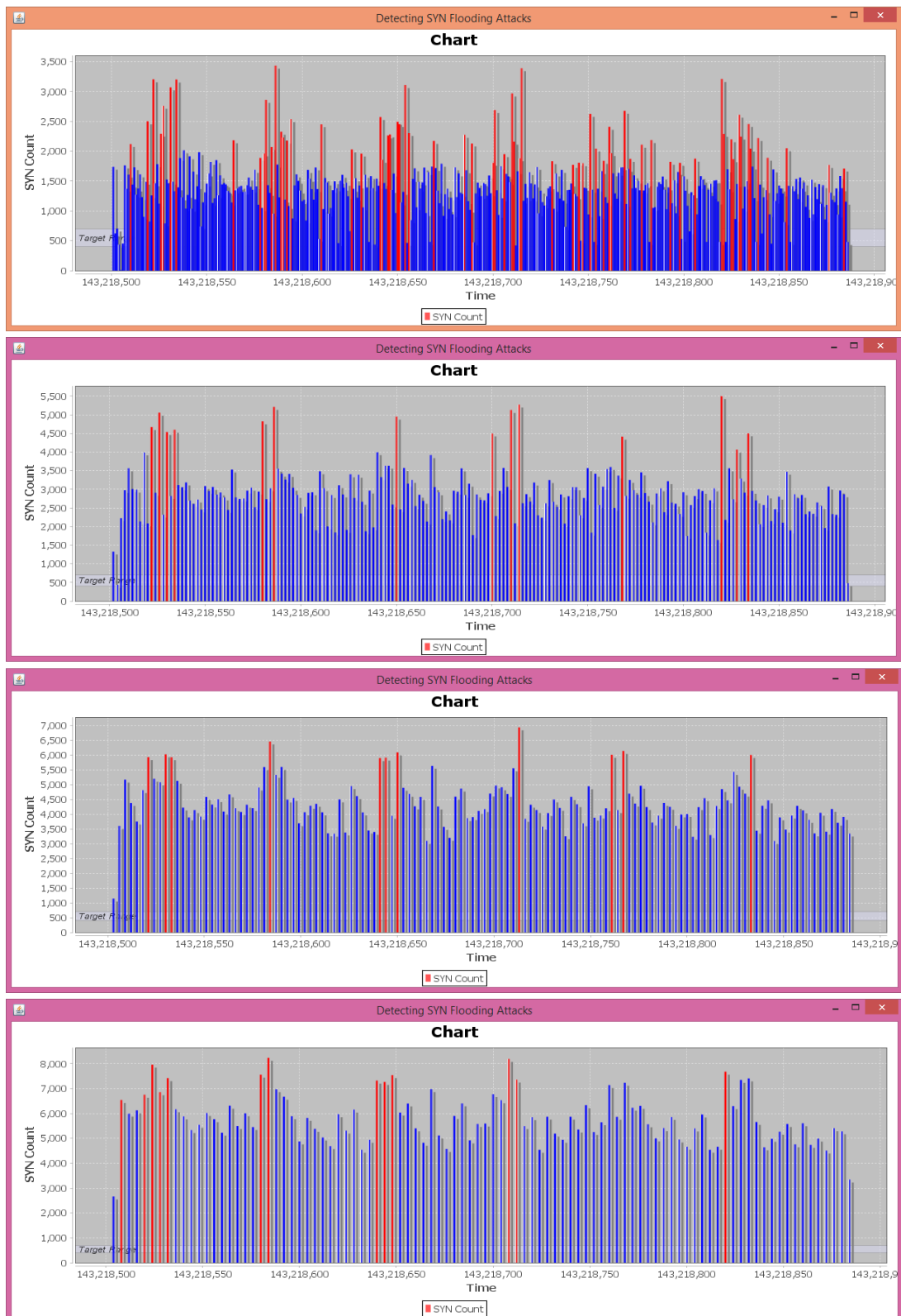
Tỷ lệ phát hiện tấn công trung bình của thuật toán AT là 72.5% trong khi với thuật toán CUSUM là 73%

### 3.2.2 Tấn công với cường độ thấp

Thử nghiệm tấn công SYN-flood với cường độ thấp 50% giá trị trung bình thực tế của lưu lượng các gói tin SYN.

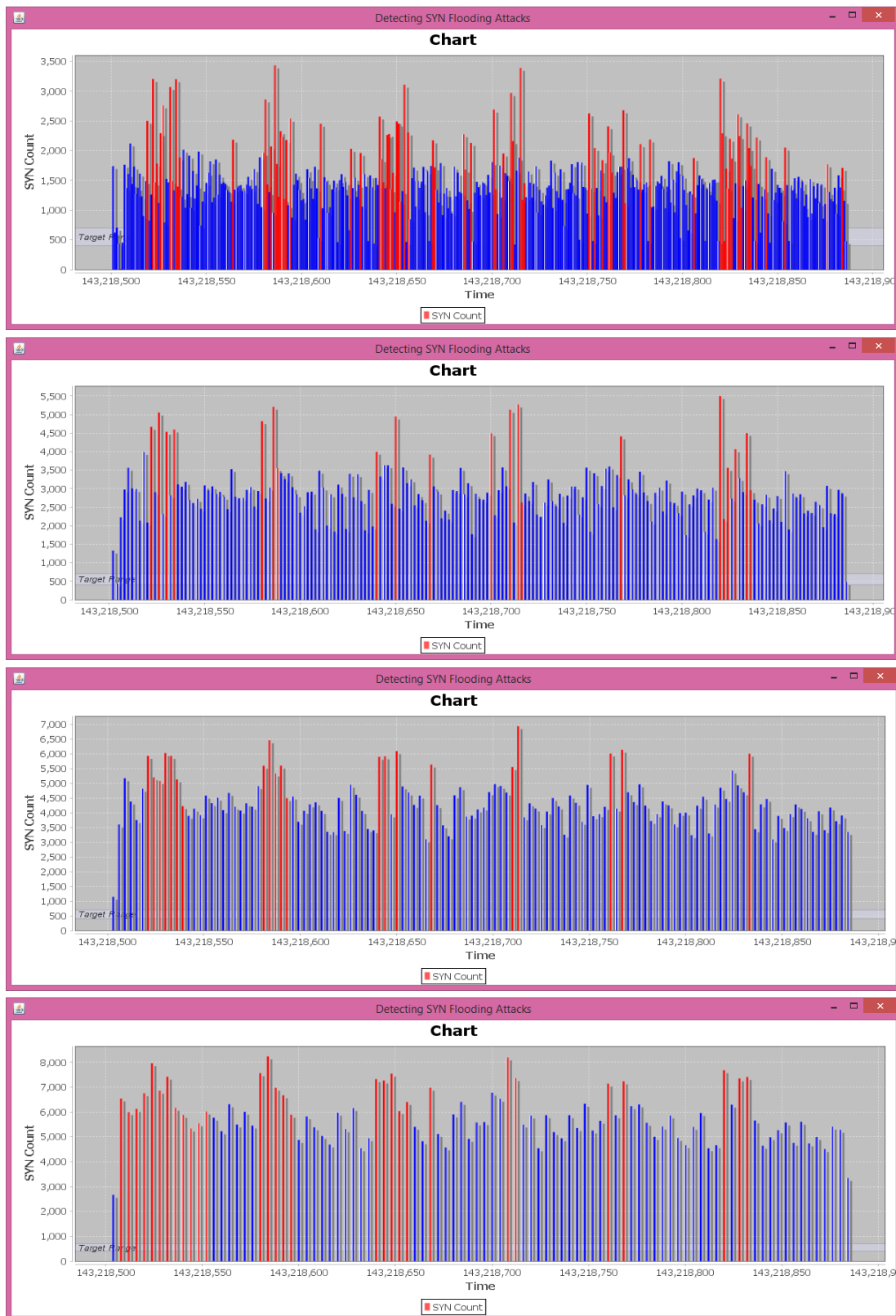
Phát hiện các cuộc tấn công cường độ thấp sẽ cho phép phát hiện sớm các cuộc tấn công có cường độ gia tăng dần, và phát hiện các cuộc tấn công gần với nguồn, hoặc trong các bộ định tuyến hay các trạm giám sát, tạo điều kiện cho việc xác định các máy tính đã bị chiếm quyền và tham gia các cuộc tấn công DoS phân tán.

Cài đặt thuật toán Adaptive Threshold với khoảng thời gian lấy mẫu khác nhau, các tham số thử nghiệm lần lượt là  $\alpha = 0.3$ ,  $k = 1$ , và  $\beta = 0.98$



**Hình 22: Biểu đồ phát hiện tấn công SYN-flood sử dụng thuật toán AT trong tấn công cường độ thấp với time intervals lần lượt là 10, 20, 30, 40 giây**

Cài đặt thuật toán CUSUM với khoảng thời gian lấy mẫu khác nhau, các tham số thử nghiệm lần lượt là  $\sigma^2 = 1$ ,  $\alpha = 0.5$ ,  $h = 4$ , và  $\beta = 0.98$



**Hình 23: Biểu đồ phát hiện tấn công SYN-flood sử dụng thuật toán CUSUM trong tấn công cường độ thấp với time intervals lần lượt là 10, 20, 30, 40 giây**

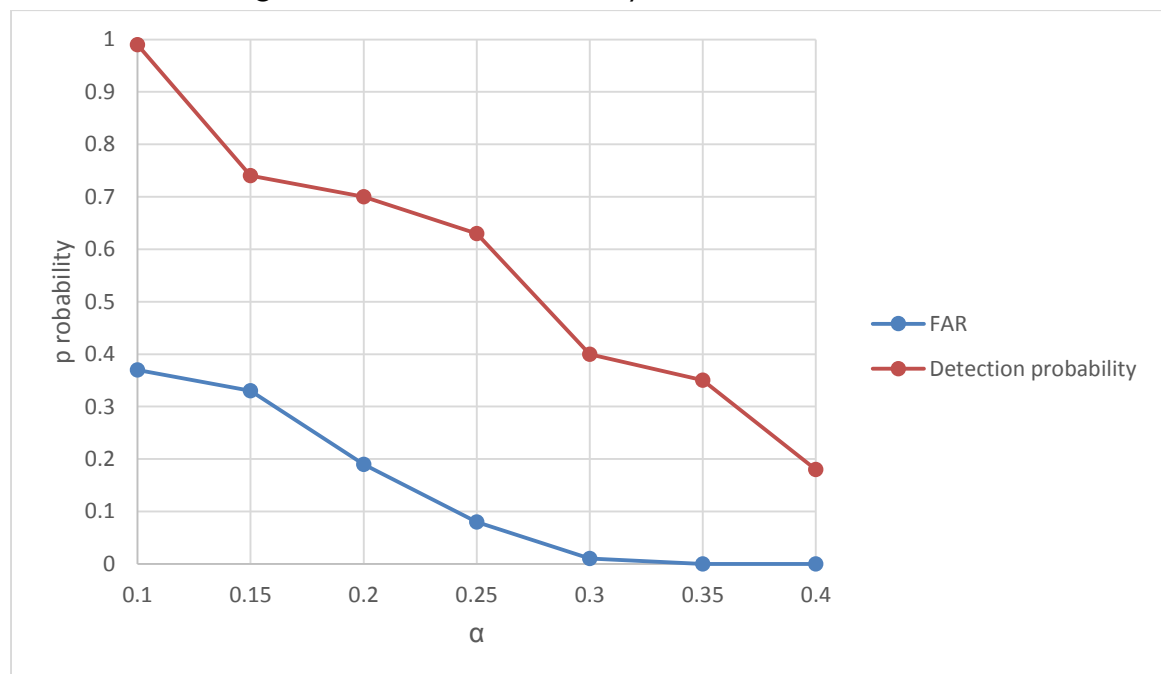
Time intervals (s)	AT	CUSUM
10	35%	61%
20	25%	37%
30	30%	52%
40	44%	77%

**Bảng 6: Tỷ lệ DP của hai thuật toán với tấn công cường độ thấp**

Hình trên cho thấy rằng đối với các cuộc tấn công cường độ thấp thì hiệu quả của cả hai thuật toán đã thấp đi đáng kể, tỷ lệ phát hiện tấn công trung bình của thuật toán AT là 33,5% trong khi với thuật toán CUSUM là 56%

### 3.2.3 Sự cân bằng giữa tỷ lệ DP và FAR trong thuật toán AT

Cài đặt thuật toán Adaptive Threshold với khoảng thời gian lấy mẫu là 40 giây, các tham số thử nghiệm lần lượt là  $k = 1$ , và  $\beta = 0.98$ .

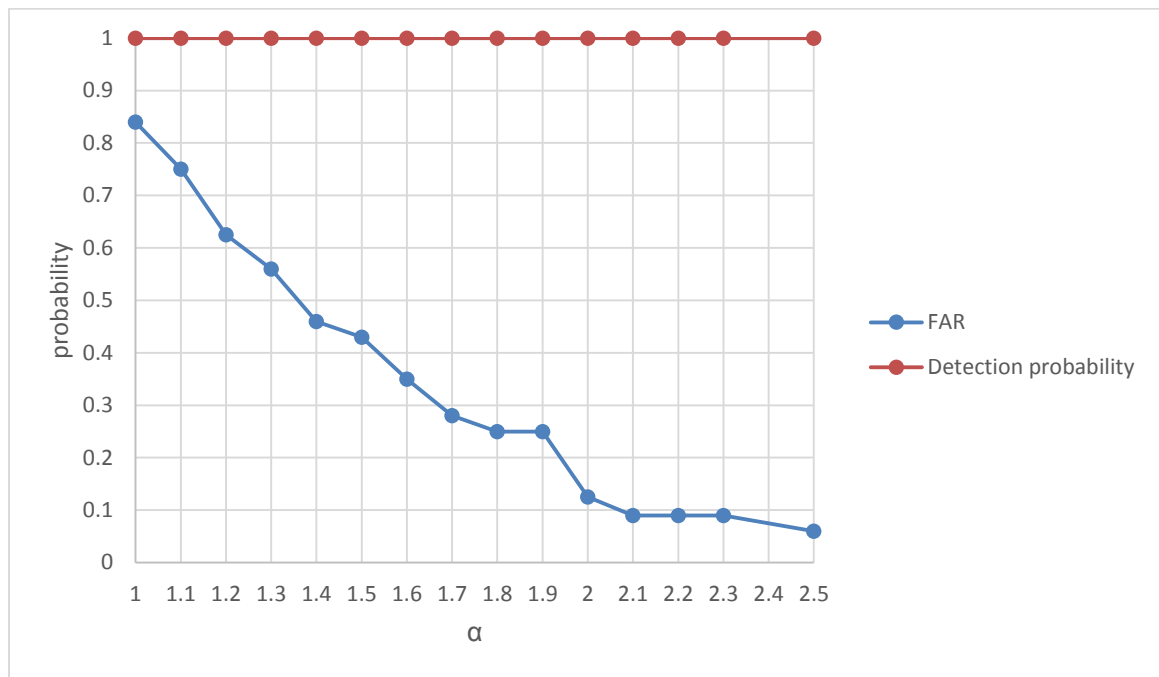


**Hình 24: Biểu đồ đánh giá sự cân bằng giữa xác suất phát hiện và tỷ lệ báo động giả**

Hình trên cho thấy sự biến đổi tỉ lệ thuận giữa xác suất phát hiện và tỷ lệ báo động giả (FAR) cho các giá trị khác nhau của  $\alpha$  trong thuật toán AT.

### 3.2.4 Ảnh hưởng của tham số biên độ $\alpha$

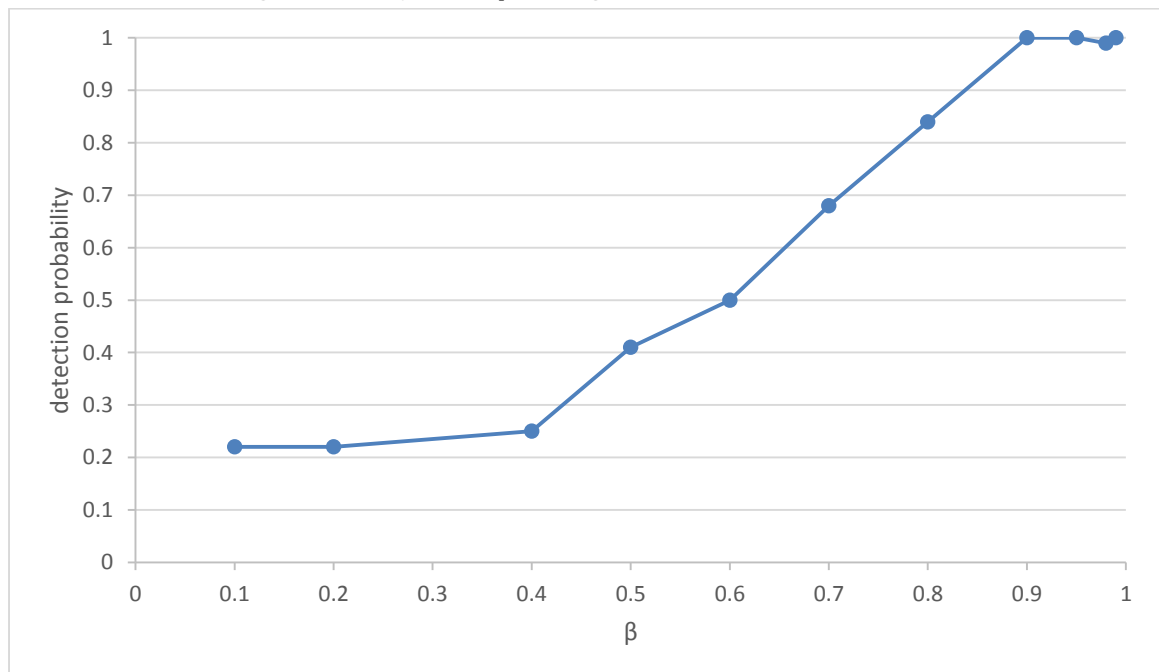
Cài đặt thuật toán CUSUM với khoảng thời gian lấy mẫu là 40 giây, các tham số thử nghiệm lần lượt là  $\sigma^2 = 1$ ,  $h = 4$ , và  $\beta = 0.98$



**Hình 25: Biểu đồ đánh giá ảnh hưởng của biên độ  $\alpha$**

Hình trên cho thấy ảnh hưởng của tham số biên độ  $\alpha$  trong thuật toán CUSUM, khi tham số ngưỡng  $h$  được điều chỉnh để đạt được tỉ lệ phát hiện 100%. Từ biểu đồ cho thấy hiệu quả của thuật toán CUSUM cao nhất khi hệ số  $\alpha$  điều chỉnh trong khoảng  $[2, 2.5]$ .

### 3.2.5 Ảnh hưởng của các yếu tố $\beta$ trong EWMA



**Hình 26: Biểu đồ đánh giá ảnh hưởng của tham số  $\beta$**

Hình trên cho thấy ảnh hưởng tham số  $\beta$  (có trong EWMA) cho thuật toán AT, khi tham số  $\alpha$  được điều chỉnh để đạt được tỉ lệ FAR là 0%. Con số này cho thấy thuật toán CUSUM có hiệu quả cao nhất khi giá trị của  $\beta$  trong khoảng  $[0.9, 0.99]$ .

### 3.3 Nhận xét về giải thuật

Từ kết quả trên cùng với kết quả thực nghiệm khác ta thấy :

Với tấn công cường độ cao là trường hợp tấn công điển hình, ta nhận thấy rằng cả hai thuật toán AT và CUSUM đều cho hiệu quả phát hiện khá cao. Nguyên nhân các thuật toán không phát hiện được 100% các thời điểm tấn công là do cách thức lưu trữ dữ liệu NetFlow.

Với tấn công tấn công với cường độ thấp ta nhận thấy hiệu quả của thuật toán CUSUM cao hơn hẳn so với AT. Nguyên nhân do thuật toán CUSUM duy trì thông tin tốt hơn về số lượng dữ liệu vượt quá lưu lượng dự kiến dựa trên một số tỷ lệ trung bình ước tính.

Kết hợp với các kết quả thực nghiệm khác [1] ta thấy rằng tỉ lệ phát hiện bất thường mạng lớn trong các cuộc tấn công cường độ cao, nhưng đối với các cuộc tấn công cường độ thấp thì tỉ lệ FAR khá lớn.

Đặc điểm của các cuộc tấn công và đặc thù của lưu lượng mạng ảnh hưởng khá lớn đến hiệu suất phát hiện tấn công của các thuật toán.

# KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Phần này trình bày kết luận chung của ĐATN, đánh giá các công việc đã làm được và chưa làm được trong khuôn khổ thực hiện ĐATN. Sau đó là định hướng nhằm phát triển và hoàn thiện các kết quả đã đạt được.

## 1. Kết quả đạt được

Trong quá trình thực hiện đồ án em đã thực hiện được các kết quả sau đây:

- Tìm hiểu công cụ nhúng NetFlow tích hợp trong các thiết bị cisco
- Tìm hiểu cách thức hoạt động của hệ thống giám sát mạng sử dụng NetFlow
- Tìm hiểu cách cấu tạo của bản ghi NetFlow và trích xuất thông tin NetFlow sử dụng để phát hiện bất thường mạng
- Tìm hiểu các phương thức tấn công DoS
- Cài đặt thử nghiệm hệ thống server mô phỏng có tích hợp hệ thống NetFlow để giám sát và phân tích mạng sử dụng thiết bị mạng bộ định tuyến Cisco
- Áp dụng các thuật toán vào phân tích phát hiện bất thường mạng trong tấn công SYN-flood
- Đánh giá hiệu quả của thuật toán đối với dữ liệu chạy trong hệ thống mạng mô phỏng

Tuy nhiên đồ án vẫn còn nhiều thiếu sót và hạn chế, cụ thể như sau:

- Chưa phân tích phát hiện bất thường mạng được trong thời gian thực
- Hệ thống mạng nhỏ nên chưa đánh giá được hiệu quả của thuật toán trong các trường hợp hệ thống mạng phức tạp với lưu lượng dữ liệu lớn
- Chưa đánh giá được hiệu quả của thuật toán trong các trường hợp mạng đặc thù

## 2. Hướng phát triển

Để có thể phát triển trong hệ thống lớn có tính ứng dụng cao, em xin đề xuất các hướng phát triển như sau:

- Giám sát mạng được trong thời gian thực.
- Phân tích phát hiện bất thường được với nhiều trường hợp tấn công vào hệ thống mạng khác
- Cài đặt thêm hệ thống ngăn chặn tấn công hệ thống mạng sau khi phát hiện được bất thường trong mạng
- Tối ưu xử lý dữ liệu NetFlow trong các hệ thống mạng lớn với dữ liệu netflow lên đến hàng GB.

Do thời gian tìm hiểu, nghiên cứu, thực hiện đồ án có hạn, cùng với sự hạn chế về kiến thức chuyên môn lẫn kinh nghiệm thực tiễn nên đồ án mới chỉ dừng lại ở việc tìm hiểu những lí thuyết và xây dựng hệ thống cơ bản, đơn giản, mới phân tích phát hiện được những trường hợp bất thường mạng điện hình, ngoài ra đồ án tốt nghiệp không thể tránh được những sai sót. Tôi rất mong có được những ý kiến đánh giá, góp ý của các thầy cô và các bạn để đồ án thêm hoàn thiện.

Một lần nữa, tôi xin chân thành cảm ơn PGS. TS. Ngô Hồng Sơn đã tạo điều kiện và giúp đỡ tôi trong suốt quá trình làm đồ án này.



## PHỤ LỤC

### 1. Cài đặt minicom trên máy giám sát

Minicom là tool chạy trên Ubuntu cho phép ta giao tiếp với IOS của bộ định tuyến thông qua cổng console.

- Minicom có sẵn trên kho ứng dụng của Ubuntu để tải về

```
sudo apt-get install minicom
```

- Sử dụng lệnh này để tìm tên cổng giao tiếp với console của bộ định tuyến Cisco

```
dmesg | grep tty
```

- Output khi kết nối trực tiếp với các cổng serial

```
[ 22.587279] console [tty0] enabled
[ 24.186230] serial8250: ttyS0 at I/O 0x3f8 (irq = 4) is a 16550A
[ 24.186860] 00:08: ttyS0 at I/O 0x3f8 (irq = 4) is a 16550A
```

```
[ 51.598012] audit(1243322582.732:2): type=1503 operation="inode_permission"
requested_mask="a::" denied_mask="a::" name="/dev/tty" pid=5705
profile="/usr/sbin/cupsd" namespace="default"
```

- Output khi sử dụng cổng USB console

```
[ 0.000000] console [tty0] enabled
[ 5.065029] usb 4-3: pl2303 converter now attached to ttyUSB0
```

Tên của các cổng console có thể là ttyS0 hoặc ttyUSB0 tùy thuộc vào đầu tiếp nối.

- Tiếp theo ta khởi động minicom và cấu hình

```
sudo minicom -s
```

- Cấu hình lại minicom

Minicom Command Summary		
Commands can be called by CTRL-A <key>		
Main Functions		Other Functions
Dialing directory..D	run script (Go)....G	Clear Screen.....C
Send files.....S	Receive files.....R	cOnfigure Minicom..O
comm Parameters....P	Add linefeed.....A	Suspend minicom....J
Capture on/off.....L	Hangup.....H	eXit and reset.....X
send break.....F	initialize Modem...M	Quit with no reset.Q
Terminal settings..T	run Kermit.....K	Cursor key mode....I
lineWrap on/off....W	local Echo on/off..E	Help screen.....Z
Paste file.....Y		scroll Back.....B
Select function or press Enter for none.█		
Written by Miquel van Smoorenburg 1991-1995		
Some additions by Jukka Lahtinen 1997-2000		
i18n by Arnaldo Carvalho de Melo 1998		

**Hình 27: Giao diện chính của minicom**

- Lựa chọn configure minicom -> Serial port setup sau đó sửa lại các thông số cho phù hợp như hình dưới

A -	Serial Device	:	/dev/ttyUSB1
B -	Lockfile Location	:	/var/lock
C -	Callin Program	:	
D -	Callout Program	:	
E -	Bps/Par/Bits	:	115200 8N1
F -	Hardware Flow Control	:	No
G -	Software Flow Control	:	No
Change which setting? █			

**Hình 28: Thông số cài đặt minicom để kết nối với cổng console**

## 2. Cấu hình NetFlow cho bộ định tuyến

Lần lượt cấu hình chi tiết từng thành phần trong bộ định tuyến:

### **Bước 1 : Đặt tên và địa chỉ của các interface**

Router#conf terminal

Router# hostname HCT // đặt tên cho bộ định tuyến

Router(config) #int g0/1

HCT(config-if) #ip address dhcp //sinh dhcp từ mạng internet cấp về

HCT(config-if) #no shutdown

HCT(config-if) #exit

HCT(config) #int g0/0

HCT(config-if) # ip address 192.168.0.1 255.255.255.0 //cấp phát địa  
chỉ cho cổng g0/0

HCT(config-if) #no shutdown

HCT(config-if) #exit

HCT(config) #exit

HCT# show ip int br //hiển thị ip các cổng như hình dưới là ok

```
HCT#show ip int br
Interface                               IP-Address      OK? Method Status        Protocol
Embedded-Service-Engine0/0             unassigned      YES NURAM   administratively down down
GigabitEthernet0/0                      192.168.0.1     YES NURAM   up            up
GigabitEthernet0/1                      192.168.4.253   YES DHCP    up            up
Serial0/1/0                             unassigned      YES NURAM   administratively down down
Serial0/1/1                             unassigned      YES NURAM   administratively down down
GigabitEthernet0/0/0                    unassigned      YES unset   up            up
GigabitEthernet0/0/1                    unassigned      YES unset   down          down
GigabitEthernet0/0/2                    unassigned      YES unset   down          down
GigabitEthernet0/0/3                    unassigned      YES unset   down          down
NUI0                                     unassigned      YES unset   administratively down down
Ulan1                                     unassigned      YES unset   down          down
```

**Hình 29: Hiển thị thông tin và trạng thái các cổng**

HCT# ping 8.8.8.8 // thử ping ra mạng internet

```
HCT#ping 8.8.8.8
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 8.8.8.8, timeout is 2 seconds:
!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 68/111/240 ms
HCT#
```

**Hình 30: Ví dụ kết nối với mạng ngoài**

## **Bước 2: Cấu hình NAT và default route để mạng bên trong truy cập được Internet**

HCT(config) # access-list 1 permit any //NAT toàn bộ mạng trong đi internet

HCT(config) # ip nat inside source list 1 interface f0/1 overload

HCT(config) #int g0/0

HCT(config-if) #ip nat inside

HCT(config-if) #exit

HCT(config) #int g0/1

HCT(config-if) #ip nat outside

HCT(config-if)#exit

Kiểm tra kết quả :

HCT#show ip nat translations

```
HCT#show ip nat translations
Pro Inside global      Inside local      Outside local      Outside global
tcp 192.168.4.253:49478 192.168.0.2:49478 37.187.172.34:5938 37.187.172.34:5938
tcp 192.168.4.253:49724 192.168.0.2:49724 198.199.14.40:443 198.199.14.40:443
tcp 192.168.4.253:51566 192.168.0.2:51566 31.13.79.246:443 31.13.79.246:443
udp 192.168.4.253:51760 192.168.0.2:51760 192.168.41.236:161 192.168.41.236:161
udp 192.168.4.253:51762 192.168.0.2:51762 192.168.4.200:161 192.168.4.200:161
tcp 192.168.4.253:51927 192.168.0.2:51927 1.9.56.80:443 1.9.56.80:443
tcp 192.168.4.253:51934 192.168.0.2:51934 31.13.79.246:443 31.13.79.246:443
tcp 192.168.4.253:51936 192.168.0.2:51936 58.27.86.34:443 58.27.86.34:443
udp 192.168.4.253:60837 192.168.0.2:60837 192.168.4.200:161 192.168.4.200:161
tcp 192.168.4.253:49178 192.168.1.101:49178 77.234.43.63:80 77.234.43.63:80
tcp 192.168.4.253:49243 192.168.1.101:49243 111.221.72.33:443 111.221.72.33:443
tcp 192.168.4.253:49262 192.168.1.101:49262 64.233.187.188:5228 64.233.187.188:5228
tcp 192.168.4.253:49318 192.168.1.101:49318 216.58.221.110:443 216.58.221.110:443
tcp 192.168.4.253:49319 192.168.1.101:49319 216.58.221.67:80 216.58.221.67:80
tcp 192.168.4.253:49320 192.168.1.101:49320 204.79.197.200:443 204.79.197.200:443
tcp 192.168.4.253:49321 192.168.1.101:49321 204.79.197.200:443 204.79.197.200:443
udp 192.168.4.253:51840 192.168.1.101:51840 216.58.221.110:443 216.58.221.110:443
```

**Hình 31: Hiển thị thông tin bảng NAT**

### **Bước 3: Cấu hình DHCP cho các cổng**

```
HCT(config)#ip dhcp pool TEST
```

```
HCT(dhcp-config)#network 192.168.0.0 /24
```

```
HCT(dhcp-config)#default-router 192.168.0.1
```

```
HCT(dhcp-config)#dns-server 8.8.8.8 8.8.4.4
```

```
HCT(dhcp-config)#exit
```

### **Bước 4: Kích hoạt và cấu hình các cổng switch tại layer 3 gắn trong bộ định tuyến**

- Mở mạng Vlan ảo để tạo switch tích hợp trên bộ định tuyến

```
HCT(config)#int vlan 20
```

```
HCT(config-if)#ip address 192.168.1.1 255.255.255.0
```

```
HCT(config-if)#ip nat inside
```

```
HCT#show ip int br
Interface                               IP-Address      OK? Method Status        Protocol
Embedded-Service-Engine0/0             unassigned      YES NURAM        administratively down down
GigabitEthernet0/0                      192.168.0.1     YES NURAM        up            up
GigabitEthernet0/1                      192.168.4.253   YES DHCP        up            up
Serial0/1/0                             unassigned      YES NURAM        administratively down down
Serial0/1/1                             unassigned      YES NURAM        administratively down down
GigabitEthernet0/0/0                    unassigned      YES unset       up            up
GigabitEthernet0/0/1                    unassigned      YES unset       down          down
GigabitEthernet0/0/2                    unassigned      YES unset       down          down
GigabitEthernet0/0/3                    unassigned      YES unset       down          down
NUI0                                     unassigned      YES unset       administratively down down
Ulan1                                    unassigned      YES unset       down          down
Ulan20                                   192.168.1.1     YES NURAM        up            up
HCT#
```

**Hình 32: Hiển thị trạng thái mạng VLAN**

- Lần lượt mở và cấu hình các cổng trên switch để kết nối vào hệ thống mạng

```
HCT(config)#int gigabitEthernet 0/0/0
```

```
HCT(config-if)#switchport access vlan 20
HCT(config-if)#no shutdown
HCT(config-if)#exit
```

```
HCT(config)#int gigabitEthernet 0/0/1
HCT(config-if)#switchport access vlan 20
HCT(config-if)#no shutdown
HCT(config-if)#exit
```

```
HCT(config)#int gigabitEthernet 0/0/2
HCT(config-if)#switchport access vlan 20
HCT(config-if)#no shutdown
HCT(config-if)#exit
```

```
HCT(config)#int gigabitEthernet 0/0/3
HCT(config-if)#switchport access vlan 20
HCT(config-if)#no shutdown
HCT(config-if)#exit
```

- Tiến hành định tuyến tĩnh cho bộ định tuyến

```
HCT(config)#int gigabitEthernet 0/1
HCT(config-if)#ip route 0.0.0.0 0.0.0.0 192.168.4.1
```

- Liên kết các vùng mạng trong bộ định tuyến (Routing Information Protocol)

```
HCT(config)#router rip
HCT(config-router)#network 192.168.0.0
HCT(config-router)#network 192.168.4.0
HCT(config-router)#network 192.168.1.0
```

- Mở DHCP trên vùng mạng mới

```
HCT(config)#ip dhcp pool TEST
HCT(dhcp-config)#network 192.168.1.0 /24
HCT(dhcp-config)#default-router 192.168.1.1
HCT(dhcp-config)#dns-server 8.8.8.8 8.8.4.4
```

### **Bước 5 : Cấu hình NetFlow trên HCT**

- Xác định thiết bị sẽ nhận flow cache (lưu ý phải bao gồm giá trị port):

```
HCT(config) #ip flow-export destination 192.168.0.3 5000
```

- Xác định version của NetFlow:

```
HCT(config) # ip flow-export version 5
```

- Xác định cổng và hướng được theo dõi:

```
HCT(config)#interface g0/0
```

```
HCT(config-if)#ip flow egress
```

```
HCT(config)#interface g0/1
```

```
HCT(config-if)#ip flow ingress
```

### Các câu lệnh cần thiết khác:

Có thể sử dụng lệnh sau để theo dõi theo cả hai hướng

```
HCT(config-if)#ip route-cache flow
```

- Có thể cấu hình snmp-server để cho phép đa dạng thông tin được theo dõi :

```
HCT(config)# snmp-server community cisco
```

Lệnh xem dữ liệu NetFlow trực tiếp: show ip cache flow

last clearing of statistics never							
Protocol	Total	Flows	Packets	Bytes	Packets	Active(Sec)	Idle(Sec)
-----	Flows	/Sec	/Flow	/Pkt	/Sec	/Flow	/Flow
TCP-WWW	94290	3.0	28	551	87.4	0.9	3.2
TCP-other	7069	0.2	12	662	2.9	3.1	11.2
UDP-DNS	707	0.0	1	65	0.0	0.0	15.5
UDP-NTP	8	0.0	4	76	0.0	6.6	15.4
UDP-other	6497	0.2	8	130	1.7	6.0	15.4
ICMP	168	0.0	120	75	0.6	42.1	15.2
IP-other	1	0.0	1	72	0.0	0.0	15.1
Total:	108740	3.5	26	543	92.7	1.4	4.5

SrcIf	SrcIPAddress	DstIf	DstIPAddress	Pr	SrcP	DstP	Pkts
U120	192.168.1.101	Gi0/0x	192.168.0.2	06	C13D	0050	13
U120	192.168.1.101	Gi0/0x	192.168.0.2	06	C13C	0050	17
U120	192.168.1.101	Gi0/0x	192.168.0.2	06	C139	0050	20
U120	192.168.1.101	Gi0/0x	192.168.0.2	06	C138	0050	80
U120	192.168.1.101	Gi0/0x	192.168.0.2	06	C13B	0050	67
U120	192.168.1.101	Gi0/0x	192.168.0.2	06	C13A	0050	129

**Hình 33: Ví dụ hiển thị thông tin NetFlow**

Lệnh xem thông tin cấu hình NetFlow: show ip flow export



```
HCT#show ip flow export
Flow export v5 is enabled for main cache
Export source and destination details :
  URF ID : Default
    Destination(1) 192.168.0.3 (5000)
Version 5 flow records
108731 flows exported in 4310 udp datagrams
0 flows failed due to lack of export packet
0 export packets were sent up to process level
0 export packets were dropped due to no fib
2 export packets were dropped due to adjacency issues
0 export packets were dropped due to fragmentation failures
0 export packets were dropped due to encapsulation fixup failures
```

**Hình 34: Ví dụ hiển thị trạng thái Netflow**

lưu cấu hình: copy running-config startup-config

xóa cấu hình triệt để: erase startup-config/reload

## TÀI LIỆU THAM KHẢO

- [1] H. Wang, D. Zhang, and K. G. Shin, “Detecting SYN flooding attacks”, in Proceedings of Annual Joint Conference of the IEEE Computer and Communications Societies(INFOCOM), 23-27, 2002.
- [2] Youngseok Lee, Wonchul Kang, Hyeongu Son, “An Internet Traffic Analysis Method with MapReduce,”in: 2010 IEEE/IFIP Network Operations and Management Symposium Workshops.
- [3] Vasilios A. Siris and Fotini Papagalou, “Application of Anomaly Detection Algorithm For Detecting SYN Flooding Attacks,” in: 2004 IEEE
- [4] Mitko Bogdanoski, “Analysis of the SYN Flood DoS Attack,” in: I.J. Computer Network and Information Security, 2013, 8, 1-11.
- [5] Karthik Pai, B.H. Nagesh H.R., Ph.D. Abhijit Bhat, “Detection and Performance Evaluation of DoS/DDoS Attacks using SYN Flooding Attacks”, in: International Conference on Information and Communication Technologies (ICICT- 2014).
- [6] Rui Zhong and Guangxue Yue, “DDoS Detection System Based on Data Mining,” Proceedings of the Second International Symposium on Networking and Network Security, Jinggangshan, P. R. China, 2-4, April. 2010, pp. 062-065.
- [7] Vasilios A. Siris, “Denial of Service and Anomaly Detection,” in: SCAMPI BoF, Zagreb, May-21-2002
- [8] Dipti J. Suryawanshi, U. A. Mande, “Parallel Processing of Internet Traffic Measurement and Analysis Using Hadoop,” Volume 3, Issue 5, September 2014
- [9] Tongguang Zhang, “Cumulative Sum Algorithm for Detecting SYN Flooding Attacks,”
- [10] F. Palmieri and U. Fiore, "Network anomaly detection through nonlinear analysis," Computers & Security, vol. 29, pp. 737-755, 2010.
- [11] Yeonhee Lee, Wonchul Kang, and Youngseok Lee, “A Hadoop-based Packet Trace Processing Tool”
- [12] “Cisco IOS NetFlow Configuration Guide”, August 6, 2008
- [13] Tongguang Zhang, “Cumulative Sum Algorithm for Detecting SYN Flooding Attacks”