

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



TRẦN THANH HÒA

**NGHIÊN CỨU MỘT SỐ KỸ THUẬT PHÁT HIỆN
XÂM NHẬP MẠNG BẰNG PHƯƠNG PHÁP SO KHỚP**

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

THÁI NGUYÊN - 2016

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



TRẦN THANH HÒA

**NGHIÊN CỨU MỘT SỐ KỸ THUẬT PHÁT HIỆN XÂM NHẬP MẠNG BẰNG
PHƯƠNG PHÁP SO KHỚP**

Chuyên ngành: Khoa học máy tính

Mã số: 60 48 0101

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

***Người hướng dẫn khoa học:* TS. Nguyễn Ngọc Cương**

THÁI NGUYÊN - 2016

LỜI CAM ĐOAN

Học viên xin cam đoan luận văn “**Nghiên cứu một số kỹ thuật phát hiện xâm nhập mạng bằng phương pháp so khớp**” là công trình nghiên cứu của cá nhân học viên tìm hiểu, nghiên cứu dưới sự hướng dẫn của TS. Nguyễn Ngọc Cương. Các kết quả là hoàn toàn trung thực, toàn bộ nội dung nghiên cứu của luận văn, các vấn đề được trình bày đều là những tìm hiểu và nghiên cứu của chính cá nhân học viên hoặc là được trích dẫn từ các nguồn tài liệu được trích dẫn và chú thích đầy đủ.

TÁC GIẢ LUẬN VĂN

Trần Thanh Hòa

LỜI CẢM ƠN

Học viên xin bày tỏ lời cảm ơn chân thành tới tập thể các thầy cô giáo Viện công nghệ thông tin, các thầy cô giáo Trường Đại học Công nghệ thông tin và truyền thông - Đại học Thái Nguyên đã mang lại cho học viên kiến thức vô cùng quý giá và bổ ích trong suốt quá trình học tập chương trình cao học tại trường. Đặc biệt học viên xin bày tỏ lòng biết ơn sâu sắc tới thầy giáo TS.Nguyễn Ngọc Cương - Học viện an ninh đã định hướng khoa học và đưa ra những góp ý, gợi ý, chỉnh sửa quý báu, quan tâm, tạo điều kiện thuận lợi trong quá trình nghiên cứu hoàn thành luận văn này.

Cuối cùng, học viên xin chân thành cảm ơn các bạn bè đồng nghiệp, gia đình và người thân đã quan tâm, giúp đỡ và chia sẻ với học viên trong suốt quá trình học tập.

Do thời gian và kiến thức có hạn nên luận văn không tránh khỏi những thiếu sót nhất định. Học viên rất mong nhận được những sự góp ý quý báu của thầy cô và các bạn.

Thái Nguyên, ngày 10 tháng 5 năm 2016
HỌC VIÊN

Trần Thanh Hòa

DANH MỤC CÁC HÌNH VẼ

| | |
|---|----|
| Hình 1.1. Hệ thống chống xâm nhập IDS | 8 |
| Hình 1.2. Mô hình Network based IDS – NIDS | 9 |
| Hình 1.3. Mô hình Host based IDS – HIDS | 10 |
| Hình 1.4. Các thành phần của hệ thống phát hiện xâm nhập..... | 14 |
| Hình 2.1. Xây dựng hàm Goto | 41 |
| Hình 2.2. Xây dựng hàm Failure..... | 42 |
| Hình 2.3. Xây dựng hàm Output..... | 42 |
| Hình 3.1. Mô hình hệ thống phát hiện xâm nhập mạng..... | 45 |
| Hình 3.2. Các thành phần của Snort..... | 48 |
| Hình 3.3. Lệnh Snort –W xem thông số card mạng..... | 55 |
| Hình 3.4. Mô hình mạng ở trung tâm GDTX Định Hóa..... | 56 |
| Hình 3.5. Mô hình giải pháp kết hợp IDS..... | 57 |
| Hình 3.6. Gói tin được mã hóa..... | 59 |

MỤC LỤC

| | |
|--|-----|
| LỜI CAM ĐOAN | i |
| LỜI CẢM ƠN | ii |
| DANH MỤC CÁC HÌNH VẼ | iii |
| MỞ ĐẦU | 3 |
| 1. Lý do chọn đề tài..... | 3 |
| 2. Hướng nghiên cứu của luận văn | 4 |
| 3. Đối tượng và phạm vi nghiên cứu..... | 5 |
| 4. Những nội dung nghiên cứu chính..... | 5 |
| 5. Phương pháp nghiên cứu..... | 5 |
| 1.1.1. Định nghĩa..... | 6 |
| 1.1.1.1. Phát hiện xâm nhập | 6 |
| 1.1.1.2. Hệ thống phát hiện xâm nhập | 7 |
| 1.1.2. Tính năng của IDS | 8 |
| 1.1.3. Phân loại IDS | 8 |
| 1.1.3.1. Network based IDS – NIDS | 9 |
| 1.1.3.2. Host based IDS – HIDS | 10 |
| 1.1.4. Cơ chế hoạt động của hệ thống IDS..... | 11 |
| 1.1.4.1. Phát hiện sự lạm dụng | 12 |
| 1.1.4.2. Phát hiện sự bất thường..... | 13 |
| 1.1.5. Ưu điểm và hạn chế của IDS | 13 |
| 1.1.5.1. Ưu điểm..... | 13 |
| 1.1.5.2. Hạn chế..... | 13 |
| 1.2. Các thành phần của hệ thống phát hiện xâm nhập..... | 13 |

| | |
|--|----|
| 1.3. Phân biệt những hệ thống không phải là IDS | 14 |
| 1.4. Một số kỹ thuật phát hiện xâm nhập mạng | 15 |
| 1.4.1. Phương pháp tiếp cận dựa trên xác suất thống kê..... | 15 |
| 1.4.2. Phương pháp tiếp cận dựa trên trạng thái | 15 |
| 1.4.3. Phương pháp tiếp cận dựa trên hệ chuyên gia | 16 |
| 1.4.4. Phương pháp tiếp cận dựa trên khai phá dữ liệu..... | 17 |
| 1.4.5. Hệ thống phát hiện xâm nhập dựa trên mạng nơ-ron | 17 |
| 1.4.6. Tiếp cận dựa trên so khớp mẫu | 19 |
| 1.5. Kết chương..... | 19 |
| 2.1. Bài toán so khớp mẫu..... | 20 |
| 2.1.1. So khớp mẫu là gì?..... | 20 |
| 2.1.2. Lịch sử phát triển..... | 22 |
| 2.1.3. Các bước xử lý | 22 |
| 2.1.4. Các cách tiếp cận..... | 22 |
| 2.1.5. Độ phức tạp tính toán..... | 23 |
| 2.1.6. Ứng dụng của so khớp mẫu | 24 |
| 2.2. Các thuật toán so khớp trong phát hiện xâm nhập mạng..... | 24 |
| 2.2.1. Thuật toán Brute Force..... | 24 |
| 2.2.2. Các thuật toán so khớp đơn mẫu trong phát hiện xâm nhập mạng..... | 28 |
| 2.2.2.1 Thuật toán Knuth-Morris-Pratt (KMP)..... | 28 |
| 2.2.2.2 Thuật toán Boyer-Moore Horspool (BMH)..... | 32 |
| 2.2.2.3 Thuật toán Karp-Rabin..... | 34 |
| 2.2.3. Đánh giá ưu, nhược điểm các thuật toán so khớp đơn mẫu..... | 38 |

| | |
|---|----|
| 2.2.3.1. Ưu điểm..... | 39 |
| 2.2.3.2 Nhược điểm..... | 40 |
| 2.2.4. Các thuật toán so khớp đa mẫu trong phát hiện xâm nhập mạng..... | 40 |
| 2.2.4.1. Thuật toán Aho-Corasick (AC) [8]..... | 40 |
| 2.2.4.2. Thuật toán Commentz-Walter (CW) | 43 |
| 2.3. Kết chương | 44 |
| 3.1. Mô hình phát hiện xâm nhập mạng dựa trên so khớp mẫu..... | 45 |
| 3.1.1. Thành phần thu nhập gói tin | 45 |
| 3.1.2. Thành phần phân tích gói tin..... | 46 |
| 3.1.3. Thành phần phản hồi..... | 46 |
| 3.2. Hệ thống phát hiện xâm nhập mạng Snort..... | 47 |
| 3.2.1. Giới thiệu..... | 47 |
| 3.2.2. Các thành phần của Snort..... | 48 |
| 3.3. Hệ thống luật | 52 |
| 3.3.1. Định nghĩa..... | 52 |
| 3.3.2. Các thành phần của tập luật | 52 |
| 3.3.2.1. Rule header..... | 52 |
| 3.3.3.2. Rule options | 52 |
| 3.4. Cài đặt và cấu hình Snort | 53 |
| 3.5. Mô hình triển khai..... | 55 |
| 3.5.1. Mô hình bài toán: | 55 |
| 3.5.1.1. Đặt ra giải pháp | 56 |
| 3.5.1.2. Yêu cầu..... | 57 |

| | |
|---|-----|
| 3.6. Thực hiện..... | 57 |
| 3.6.1. Đầu vào của bài toán | 57 |
| 3.6.2. Đầu ra của bài toán..... | 58 |
| 3.6.3. Thử nghiệm và đánh giá kết quả..... | 58 |
| 3.6.3.1. Tập dữ liệu thử nghiệm | 58 |
| 3.6.3.2. Tiền xử lý dữ liệu | 59 |
| 3.6.3.3. Dữ liệu lựa chọn..... | 59 |
| 3.6.3.4. Thiết kế thử nghiệm | 60 |
| 3.6.3.5. Kết quả thử nghiệm..... | 60 |
| 3.6.3.6. Nhận xét kết quả..... | 60 |
| 3.6.4. Sử dụng chương trình so khớp đa mẫu Aho-Corasick vào Snort | 61 |
| 3.6.4.1. Cài đặt chương trình so khớp đa mẫu Aho – Corasick..... | 61 |
| 3.6.4.2. Cài đặt thuật toán so khớp đa mẫu Aho Corasick vào Snort | 61 |
| 3.7. Kết chương | 63 |
| KẾT LUẬN | 634 |
| TÀI LIỆU THAM KHẢO..... | 635 |
| PHỤ LỤC | 637 |

MỞ ĐẦU

1. Lý do chọn đề tài

Mạng Internet từ cuối thế kỷ XX đã mở ra một làn sóng mới về xu hướng phát triển của xã hội - thời đại của công nghệ thông tin và truyền thông, trong đó các dịch vụ trực tuyến được phát triển mạnh mẽ như thương mại điện tử, thanh toán trực tuyến, kinh doanh, tài chính, công nghiệp, an ninh, y tế,...

Người sử dụng nhờ mạng Internet có thể truy cập, khai thác và chia sẻ thông tin mọi lúc mọi nơi, tuy nhiên Internet cũng là không gian rộng mở cho kẻ xấu lợi dụng thực hiện những vụ tấn công, truy cập trái phép vào các hệ thống máy tính và mạng của người dùng.

Phát hiện xâm nhập mạng là một trong các thành phần quan trọng trong hệ thống các giải pháp đảm bảo an ninh cho các mạng hiện đại

Hệ thống phát hiện xâm nhập mạng IDS (Intrusion Detection System) có nhiệm vụ phân tích các thông tin, theo dõi, phát hiện và ngăn chặn sự xâm nhập trái phép tài nguyên làm tổn hại đến tính bảo mật, tính toàn vẹn và tính sẵn sàng của hệ thống.

Hiện nay, trong việc phát triển hệ thống IDS có nhiều phương pháp để phát hiện xâm nhập vào hệ thống như:

- *Sử dụng phương pháp Bayes* với ý tưởng tính toán xác suất Bayes để có thể dịch chuyển ngược thời gian và tìm ra nguyên nhân của sự kiện, phù hợp cho việc tìm kiếm lý do cho một sự bất thường đặc biệt trong hành vi mạng.

- *Sử dụng mạng nơ ron nhân tạo* trên mô hình Kohonen's Self Organizing features Map (SOM) và sử dụng máy học vectơ. Mục tiêu chính của việc sử dụng mạng nơ ron nhân tạo là cung cấp một phương pháp phân