TRANG THÔNG TIN LUẬN ÁN TIẾN SĨ

Tên đề tài luận án tiến sĩ: **Nghiên cứu các kỹ thuật phát hiện DGA Botnet**

Chuyên ngành: Hệ thống thông tin

Mã số: 9.48.01.04

Họ và tên NCS: **Vũ Xuân Hạnh**

Người hướng dẫn khoa học:

1. PGS.TS. Hoàng Xuân Dậu

2. TS. Ngô Quốc Dũng

Cơ sở đào tạo: Học viện Công nghệ Bưu chính Viễn thông

NHỮNG KẾT QUẢ MỚI CỦA LUẬN ÁN:

Luận án tập trung nghiên cứu, đề xuất một số mô hình phát hiện DGA botnet dựa trên các kỹ thuật học máy. Cụ thể, luận án tập trung vào các mục tiêu sau: (i) Nghiên cứu, đánh giá các phương pháp, kỹ thuật, giải pháp, công cụ phát hiện botnet hiện có; (ii) Nghiên cứu, đề xuất các mô hình phát hiện botnet dựa trên học máy có giám sát và học kết hợp sử dụng các tập đặc trưng phân loại tên miền mới nhằm nâng cao độ chính xác, giảm cảnh báo sai, đồng thời cho phép phát hiện nhiều dạng DGA botnet; (ii) - Cài đặt, thử nghiệm và đánh giá các mô hình phát hiện botnet đã đề xuất sử dụng các tệp dữ liệu thực tế. Đóng góp mới của quá trình nghiên cứu thể hiện trong luận án như sau:

Một là: đề xuất mô hình phát hiện DGA botnet dựa trên học máy sử dụng các đặc trưng ký tự và các đặc trưng từ. Mô hình sử dụng các đặc trưng ký tự có khả năng phát hiện hiệu quả các character-based DGA botnet - là các botnet tự sinh tên miền sử dụng thuật toán ghép ngẫu nhiên các ký tự. Mô hình sử dụng các đặc trưng từ có khả năng phát hiện hiệu quả các word-based DGA botnet - là các botnet tự sinh tên miền sử dụng thuật toán ghép các từ theo từ điển..

Hai là: đề xuất mô hình phát hiện DGA botnet dựa trên học kết hợp (*ensemble learning*). Mô hình này cho phép phát hiện hiệu quả cả character-based và word-based DGA botnet sử dụng thuật toán học kết hợp.

Ba là: Các đề xuất phát hiện DGA botnet dựa trên tên miền thực thi hiệu quả hơn so với các phương pháp dựa trên lưu lượng mạng bởi giảm thiểu các đặc trưng, xử lý dữ liệu luồng và gói tin, do đó sẽ nhanh hơn, chi phí đỡ tốn kém hơn. Các mô hình khi đưa vào ứng sẽ được cài đặt tại DNS server nhằm ngăn chặn các bot có thể liên lạc được với CnC server hoặc trước firewall trong các hệ thống đơn lẻ nhằm phát hiện máy tính nào là bot.

CÁC ỨNG DỤNG, KHẢ NĂNG ỨNG DỤNG TRONG THỰC TIỄN HOẶC NHỮNG VẤN ĐỀ CÒN BỎ NGỎ CẦN TIẾP TỤC NGHIÊN CỨU:

Trên thực tế đã và đang trở thành một trong các nguy cơ gây mất an toàn thông tin hàng đầu do chúng không ngừng phát triển về cả quy mô và mức độ tinh vi trong các kỹ thuật chỉ huy và kiểm soát. Nhiều dạng botnet sử dụng kỹ thuật DGA để sinh và đăng ký nhiều tên miền ngẫu nhiên khác nhau cho máy chủ CnC của chúng nhằm chống lại việc bị kiểm soát và vô hiệu hóa. Vì vậy, liên quan đến những đề xuất mới của luận án, có thể liệt kê những vấn đề cần nghiên cứu trong các công trình tiếp theo như sau:

Một là: Các hạn chế của mô hình kết hợp bao gồm: (i) thời gian huấn luyện và phát hiện dài hơn so với mô hình thành phần và (ii) mô hình kết hợp không có khả năng phát hiện một số DGA botnet thuộc họ mixed DGA, như Banjori.

Hai là: việc phát triển hệ thống phát hiện DGA botnet dựa trên các mô hình phát hiện đề xuất chưa được ứng dụng trong thực tế mà mới ở mức độ thực nghiệm và đánh giá.

Ba là: Trong vấn đề đề xuất mô hình phát hiện word-based DGA botnet, phạm vi nghiên cứu của luận án sử dụng các từ điển, danh mục từ thuần tiếng Anh. Xét đến các tên miền của Việt Nam, sử dụng các từ tiếng Việt trong dấu chưa được đề cập trong luận án.

Xác nhận của đại diện tập thể	**Nghiên cứu sinh**
Người hướng dẫn khoa học	
PGS.TS. Hoàng Xuân Dậu	**Vũ Xuân Hạnh**