THÔNG TIN LUẬN ÁN TIẾN SĨ

Tên luận án: **KẾT HỢP PHÂN TÍCH TĨNH VÀ KIỂM TRA ĐỘNG TRONG VIỆC XÂY DỰNG ĐỒ THỊ LUỒNG ĐIỀU KHIỂN PHỤC VỤ PHÂN TÍCH MÃ NHỊ PHÂN**

Chuyên ngành: **KHOA HỌC MÁY TÍNH**

Mã số chuyên ngành: **62.48.01.01**

Họ và tên NCS: **NGUYỄN MINH HẢI**

Người hướng dẫn khoa học: **PGS. TS. QUẢN THÀNH THƠ**

Cơ sở đào tạo: **TRƯỜNG ĐẠI HỌC BÁCH KHOA – ĐHQG TP.HCM**

Tóm tắt nội dung luận án:

Phân tích chương trình là một lĩnh vực đang thu hút rất nhiều sự chú ý của các nhà nghiên cứu. Tuy nhiên, một trong những vấn đề lớn của bài toán này nằm ở việc xây dựng *đồ thị luồng điều khiển* của chương trình nhị phân. Đây là một khó khăn lớn do đồ thị luồng điều khiển thường được tạo ra và thay đổi một cách ngẫu nhiên do những kỹ thuật như *mã tự thay đổi*. Tuy nhiên, khó khăn điển hình trong lĩnh vực xây dựng đồ thị luồng điều khiển là quá trình xử lí những câu *lệnh nhảy không trực tiếp*. Trong luận án này, chúng tôi đề xuất hướng tiếp cận *kiểm thử thực thi ký hiệu động* kết hợp quá trình *phân tích tĩnh* và *kiểm tra động* để xây dựng đồ thị luồng điều khiển từ mã nhị phân của chương trình. Mục tiêu của hướng tiếp cận này là để giảm sự không chính xác trong xử lý lệnh nhảy động của quá trình xây dựng đồ thị luồng điều khiển.

Các đóng góp chính của luận án:

Các đóng góp chính của luận án được tóm tắt như sau:

i. Luận án đề xuất một khung thức tổng quát cho xây dựng đồ thị luồng điều khiển từ mã nhị phân của chương trình một cách tự động.

ii. Luận án đưa ra giải pháp để tăng tốc quá trình thực thi của chương trình bằng cách áp dụng giải thuật song song hóa với tính toán đa luồng để tăng tốc độ xử lý các trạng thái.

iii. Luận án đề xuất cách khai thác tri thức dựa trên đồ thị luồng điều khiển của mã nhị phân. Luận án tập trung vào vấn đề nhận diện chương trình đóng gói trên mã độc với hai hướng tiếp cận: (i) sử dụng kiểm thử Chi bình phương; (ii) sử dụng mô hình Markov ẩn.

iv. Luận án trình bày giải pháp nhận diện mã độc dựa trên phương pháp học sâu.

v. Cuối cùng, luận án cũng đã xây dựng một công cụ hoàn chỉnh có tên là BE-PUM cho việc xây dựng đồ thị luồng điều khiển từ mã nhị phân.

Khả năng ứng dụng thực tiễn:

Nghiên cứu trong luận án này có tầm quan trọng rất cao trong lĩnh vực an toàn thông tin nói chung và trong lĩnh vực phân tích mã độc nói riêng. Phân tích chương trình, đặc biệt ở dạng mã nhị phân là bài toán quan trọng trong lĩnh vực an toàn thông tin, xuất phát từ nhu cầu thực tế để phân tích các ứng dụng chứa mã độc. Vì vậy, việc xây dựng đồ thị luồng điều khiển để phục vụ phân tích mã nhị phân là bài toán có ý nghĩa khoa học và thực tiễn cao.

Hướng phát triển tiếp theo của luận án:

Đối với công việc tương lai, nghiên cứu này còn nhiều vấn đề để xem xét. Chi tiết các hương mở rộng này được trình bày trong những phần dưới đây.

i. Mở rộng hỗ trợ tập lệnh x86 và Windows API.

ii. Mở rộng quá trình nhận diện các chương trình đóng gói.

iii. Nhận diện mã độc với các phương pháp học máy khác.

iv. Xây dựng mô hình weighted pushdown trên BE-PUM nhằm mục tiêu giảm số trạng thái cần xử lý trong bài toán kiểm tra mô hình.

Xác nhận của cán bộ hướng dẫn khoa học **Nghiên cứu sinh**

PGS. TS. Quản Thành Thơ Nguyễn Minh Hải