INFORMATION OF THE DOCTORAL THESIS

Research title: **COMBINE STATIC ANALYSIS AND DYNAMIC TESTING IN GENERATING CONTROL FLOW GRAPH FOR ANALYZING MALWARE**
Major: **COMPUTER SCIENCE**
Major code: **62.48.01.01**
PhD student: **NGUYEN MINH HAI**
Scientific advisor: **Assoc. Prof. Dr. QUAN THANH THO**
University: **HCMC University of Technology – Vietnam National University Ho Chi Minh City**

The thesis summary:

Program analysis has attracted much attention. The difficulty lies in constructing a *Control Flow Graph* (CFG), which is dynamically generated and modified by obfuscation techniques such as self-modifying code. Typical examples are handling dynamic jump instructions, in which destinations may be directly modified by rewriting loaded instructions on memory. In this thesis, we propose a hybrid approach which combines static analysis and dynamic testing to construct CFG from binary code. Our aim is to minimize false targets produced when processing indirect jumps during the CFG construction.

The main contributions:

The main contributions of the thesis are summarized as follows.

i. The thesis proposes a general framework for automatically generating the control flow graph of binary code

ii. The thesis proposes an approach for reducing the processing time of exploring states by applying the computational power of multiple threads.

iii. The thesis proposes an approach of exploiting the knowledge discovery based on control flow graph of binary code. It focuses on packer identification of malware using (i) Chi-square test and (ii) Hidden Markov Model.

iv. The thesis proposed an approach of malware classification using deep learning.

Finally, the thesis has built a complete tool called BE-PUM for generating the control flow graph of binary code.

Applications of the thesis:

The research in this thesis is important in the field of information security in general and malware analysis in particular. Program analysis especially for binary code is essenstial for analyzing malicious applications. Therefore, building flow control graphs for binary code could be applied to detect malware in practice with accurate results.

Further research of the thesis:

The thesis are very potential to further research, such as the followings.

i. Support more x86 instructions and Windows APIs.
ii. Support more packers with many obfuscation techniques.
iii. Identify malware using other machine learning methods.
iv. Apply weighted pushdown model in BE-PUM for reducing processing states.

Scientific advisor	**PhD student**
Assoc. Prof. Dr. Quan Thanh Tho	Nguyen Minh Hai