INFORMATION OF THE DOCTORAL THESIS

Thesis title: *" Research on DGA Botnet detection techniques "*
Speciality: Information System
Code: 9.48.01.04
PhD. Candidate: Vu Xuan Hanh
Scientific supervisors:
1. **Assoc. Prof. Hoang Xuan Dau, PhD**
2. **PhD. Ngo Quoc Dung**
Training institution: Posts and Telecommunications Institute of Technology

NEW FINDINGS OF THE THESIS

The thesis focuses on researching and proposing some DGA botnet detection models based on machine learning techniques. Specifically, the thesis focuses on the following objectives: (i) Research and evaluate existing methods, techniques, solutions and tools to detect botnets; (ii) Research and propose botnet detection models based on supervised machine learning and combined learning using new domain classifier feature sets to improve accuracy, reduce false alarms, and at the same time allow detection of many types of DGA botnets; (ii) Install, test and evaluate proposed botnet detection models using actual data files. The new contributions of the research process shown in the thesis are as follows:

1. Propose a machine learning-based DGA botnet detection model using character features and word features. The model using character features has the ability to effectively detect character-based DGA botnets, which are domain-generating botnets using a random character matching algorithm. The model using word features has the ability to effectively detect word-based DGA botnets – which generate domain names using a dictionary matching algorithm…
2. Propose a DGA botnet detection model based on associative learning (*ensemble learning*). This model enables efficient detection of both character-based and word-based DGA botnets using associative learning algorithms.
3. Domain-based DGA botnet detection recommendations perform better than network traffic-based methods by minimizing features, processing stream and packet data, and therefore faster and less costly. The application models will be installed at the DNS server to prevent bots from communicating with the CnC server or in front of the firewall in individual systems to detect which computer is the bot.

APPLICATIONS, PRACTICAL APPLICABILITY AND MATTERS THAT NEED FURTHER STUDIES

In fact, they have become one of the leading information security threats as they are constantly growing in both size and sophistication in command and control techniques. Many types of botnets use DGA to generate and register various random domains for their CnC servers to prevent censorship and disabling. Therefore,

regarding the new proposals of the thesis, it is possible to list the issues that need to be studied in the next works as follows:

1. The limitations of the association model include: (i) the training and detection time is longer than that of the component model and (ii) the combined model's inability to detect some botnet DGAs belongs to the mixed DGA family, like Banjori.
2. The development of a botnet DGA detection system based on the proposed detection models has not been applied in practice but is at the experiment and evaluation level.
3. In the problem of proposing a word-based DGA botnet detection model, the research scope of the thesis uses pure English dictionaries and word lists. Considering the domain names of Vietnam, using Vietnamese words in accents has not been mentioned in the thesis.

Confirmation of representative Scientific supervisor	**PhD. Candidate**
Assoc. Prof. Hoang Xuan Dau, PhD	**Vu Xuan Hanh**