ČESKÉ
VYSOKÉ
UČENÍ
TECHNICKÉ
V PRAZE

REVIEWER‘S OPINION OF FINAL THESIS

I. IDENTIFICATION DATA

Thesis name:	**Identification of network users by profiling their behavior**
Author’s name:	**David Kubeša**
Type of thesis :	master
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Computer science
Thesis reviewer:	Carlos Catania
Reviewer’s department:	Department of Computer Science, National University of Cuyo. Mendoza, Argentina.

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment

challenging

Evaluation of thesis difficulty of assignment.

I personally consider this work attempts to provide a solution to a complex and significant problem using state of the art algorithms and techniques.

Satisfaction of assignment

fulfilled

Assess that handed thesis meets assignment. Present points of assignment that fell short or were extended. Try to assess importance, impact or cause of each shortcoming.

The thesis met the assignment. The central aspects of a user behavior detection method have been developed. I believe the approach followed in the profile to profile classifier was better analyzed and evaluated than the user classifier. However, given the difficult associated to the user detection problem, I consider the approach proposed in the thesis was an adequate initial strategy for solving the problem.

Method of conception

correct

Assess that student has chosen correct approach or solution methods.

The student has followed the methodology used in Machine Learning. He has analyzed the state of the art, and proposed a new set of features that could deal with the problem. Then, the student has validated his hypothesis on real traffic captures following machine learning standard procedures.

Technical level

A - excellen

Assess level of thesis specialty, use of knowledge gained by study and by expert literature, use of sources and data gained by experience.

The student has proven himself capable of dealing with a new problem and provided a valid solution using a different set of tools. He has showed expertise in several areas such as software development, machine learning and network security.

Formal and language level, scope of thesis

A - excellen

Assess correctness of usage of formal notation. Assess typographical and language arrangement of thesis.

In general, thesis was well written. The student expressed in a clear language the different aspects involved in the process of building a profile classifier using formal notation when required.

Selection of sources, citation correctness

A - excellen

Present your opinion to student’s activity when obtaining and using study materials for thesis creation. Characterize selection of sources. Assess that student used all relevant sources. Verify that

all used elements are correctly distinguished from own results and thoughts. Assess that citation ethics has not been breached and that all bibliographic citations are complete and in accordance with citation convention and standards.
The student has always made reference to third-party articles and software applications used for meeting the thesis assignment. All references used in the work followed the proper quality standards

Additional commentary and evaluation *Present your opinion to achieved primary goals of thesis, e.g. level of theoretical results, level and functionality of technical or software conception, publication performance, experimental dexterity etc.*
Please insert your commentary (voluntary evaluation).

III. OVERALL EVALUATION, QUESTIONS FOR DEFENSE, CLASSIFICATION SUGGESTION

Summarize thesis aspects that swayed your final evaluation. Please present apt questions which student should answer during defense.

In the present thesis the student has proposed a new method based on features considering the statistical distribution of the network flows together with the application of a machine learning approach for performing user detection. A profile classifier conformed the basis of a preliminary user detection classifier based on a simple but straightforward strategy, which consists of averaging the results of profile classifier and then training a threshold. Two pieces of software were developed from scratch for extracting features from network data and visually comparing the resulting profiles. In addition, a machine learning approach was proposed for performing profile detection. The proposed approach was validated using three datasets created from real traffic captures considering 19 identified users following a standard machine learning methodology. The student has proved himself competent in several areas such as software development, machine learning, experimental design and network security.

Apt questions:
1) Several sets of features were described in section 3.2, however only Port and Connection based features were used. Has the student considered the use of some sort of automatic feature selection algorithm to select the optimal subset?
2) Has the student considered using other algorithms ? One class classifiers, for example, have proved been successful in profile detection
3) Can the student bring some insights about the computational time of the application?
4) Has the student considered the application of re sampling techniques such as cross validation for selecting the threshold for the user classifier?

I evaluate handed thesis with classification grade A - excellent ▾

Date: 06/05/18

Signature: