Multimed Tools Appl
DOI 10.1007/s11042-017-4586-0

Bio-inspired computational paradigm for feature investigation and malware detection: interactive analytics

Ahmad Firdaus[1,2] · **Nor Badrul Anuar**[1] · **Mohd Faizal Ab Razak**[1,2] · **Arun Kumar Sangaiah**[3]

Received: 13 December 2016 / Revised: 27 February 2017 / Accepted: 6 March 2017

Abstract Recently, people rely on mobile devices to conduct their daily fundamental activities. Simultaneously, most of the people prefer devices with Android operating system. As the demand expands, deceitful authors develop malware to compromise Android for private and money purposes. Consequently, security analysts have to conduct static and dynamic analyses to counter malware violation. In this paper, we adopt static analysis which only requests minimal resource consumption and rapid processing. However, finding a minimum set of features in the static analysis are vital because it removes irrelevant data, reduces the runtime of machine learning detection and reduces the dimensionality of datasets. Therefore, in this paper, we investigate three categories of features, which are permissions, directory path, and telephony. This investigation considers the features frequency as well as repeatedly used in each application. Subsequently, this study evaluates the proposed features in three bio-inspired machine learning classifiers in artificial neural network (ANN) category to signify the usefulness of ANN type in uncovering unknown malware. The classifiers are multilayer perceptron (MLP), voted perceptron (VP) and radial basis function network (RBFN). Among all these

✉ Ahmad Firdaus
ahmadfirdaus@um.edu.my

✉ Nor Badrul Anuar
badrul@um.edu.my

Mohd Faizal Ab Razak
faizalabrazak@siswa.um.edu.my

Arun Kumar Sangaiah
arunkumarsangaiah@gmail.com

[1] Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

[2] Faculty of Computer Systems & Software Engineering, University Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Kuantan, Pahang, Malaysia

[3] School of Computing Science and Engineering, VIT University, Vellore, Tamil Nadu 632014, India

Published online: 18 March 2017

three classifiers, the outstanding outcomes acquire is the MLP, which achieves 90% in accuracy and 87% in true positive rate (TPR), as well as 97% accuracy in our Bio Analyzer prediction system.

Keywords Static analysis · Malware · Feature selection · Android · Machine learning · Neural network