Android Malware Classification: Updating Features Through Incremental Learning Approach(UFILA)

1st Zakaria SAWADOGO
Laboratory LANI
Gaston Berger University
Saint Louis, Senegal
zakaria.sawado@gmail.com

2nd Gervais MENDY
Laboratory LITA
University Cheikh Anta Diop of Dakar
Dakar,Senegal
gervais.mendy@ucad.edu.sn

3rd Jean Marie DEMBELLE
Laboratory LANI
Gaston Berger University
Saint Louis, Senegal
jean-marie.dembele@ugb.edu.sn

4th Samuel OUYA
Laboratory LITA
University Cheikh Anta Diop of Dakar
Dakar,Senegal
samuel.ouya@gmail.com

***Abstract*—exponential growth of the use of connected objects, in particular of smartphones, is the consequence of the digitization of services. All types of applications, from the least critical to the most critical are available on mobile devices through mobile applications. The daily penetration of mobile applications in widely used devices brings certain threats. We find in the software repositories malwares and good application at the same time, which is a major cybersecurity problem. To resolve this problem, machine learning approaches have been proposed in the literature for the detection of malware in general and Android malicious applications in particular. Obfuscation techniques are used by developers to hide malicious applications that implies the need to update Android malware detection models. But many approaches in the literature are more focus on data than features. Hence our contribution is an incremental learning approach capable of detecting Android malware. We propose through UFILA approach an updating of features for the detection and classification of android malware by adding new features. We evaluated 13 classification algorithms and chose four most efficient algorithms to implement our approach. The results obtained by our approach surpass several malware detection approaches in the literature. The values of the metrics obtained respectively by the Accuracy, the precision, the recall and an F1-Score are 99%, 99% ,98.6 %98%.**

***Index Terms*—Android malware classification, classification algorithm, incremental learning, Android malware detection, machine learning, cyber-security.**

Zakaria SAWADOGO received his Master II degree in software engineering and information systems from Joseph Ki-Zerbo University in Burkina Faso. He is currently a researcher in cybersecurity including artificial intelligence and is affiliated with the Laboratoire d'Informatique, de Télécommunications et Applications (LITA) of the University Cheikh Anta Diop of Dakar and the Laboratoire d'Analyse Numérique et d'Informatique (LANI) of the University Gaston Berger of Saint Louis. His research interests include mobile security, system security and artificial intelligence. He is a member of the IEEE.

Pr. Gervais MENDY is a researcher-scientist at the ESP polytechnic school at UCAD university where he was head of the IT department from 2012 to 2016.Holder of a PhD in Computer Science from Paris-Sud XI University, his research interests are in Computer Combinatory, Social Network Analysis, Internet of Things (IoT), Artificial Intelligence and IT Security.He is a member of the LITA Laboratory of UCAD.

Pr. Jean-Marie DEMEBELE is Assimilated Professor in Computer Science at the UGB, ex Director of the UFR SAT and member of the UMI UMMISCO Senegal. He has to his credit, several scientific publications in international peer-reviewed journals. His research interests include artificial intelligence, agent-based modeling, dynamical systems, evolutionary algorithms, genetic regulatory networks, machine learning

Pr. Samuel Ouya is currently the director of the LITA laboratory at the UCAD. He was from 2013 to May 2017 the first Director of Infrastructure and Information System of the first virtual university of Senegal (UVS). Holder of a thesis in Applied Mathematics from the Gaston Berger University of Saint-Louis Senegal and a Telecommunications Thesis from the UCAD university in Dakar-Senegal, he is interested in he is interested in Applications of innovative telecom services to virtual organizations.