

Research Article

C SVM Classification and KNN Techniques for Cyber Crime Detection

K. Veena,¹ K. Meena,² Yuvaraja Teekaraman ,³ Ramya Kuppusamy ,⁴ and Arun Radhakrishnan ⁵

¹Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, 600119, Chennai, India

²Department of Computer Science and Engineering, Institute of Aeronautical Engineering, Hyderabad, India 500043

³Department of Electronic and Electrical Engineering, The University of Sheffield, Sheffield S1 3JD, UK

⁴Department of Electrical and Electronics Engineering, Sri Sairam College of Engineering, 562106, Bangalore City, India

⁵Faculty of Electrical & Computer Engineering, Jimma Institute of Technology, Jimma University, Ethiopia

Correspondence should be addressed to Yuvaraja Teekaraman; yuvarajastr@ieee.org and Arun Radhakrishnan; arun.radhakrishnan@ju.edu.et

DWw[hW \$' EVbFW_ TVd\$ " \$ #- DW[eW \$ & @ahW_ TVd\$ " \$ #- 3UWbFW %6Ww_ TVd\$ " \$ #- BgT' [eZW #) <S gSk \$ " \$ \$

Academic Editor: Deepak Kumar Jain

Copyright © 2022 K. Veena et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the digital age, cybercrime is spreading its root widely. Internet evolution has turned out to a boon as well as curse for those confronting the issues of privacy, national security, social decency, IP rights, child protection, ghting, detecting, and prosecuting cybercrime. Hence, there arises a need to detect the cybercriminal. Cybercrime identi cation utilizes dataset that is taken from CBS open dataset. For identifying the cybercriminal, support vector machine (SVM) in the C SVM classi cation and K-nearest neighbor (KNN) models is utilized for determining the cybercrime information. The evaluation of the performance is done taking the following metrics into consideration: true positive, false positive, true negative and false negative, false alarm rate, detection rate, accuracy, recall, precision, speci city, sensitivity, classi cation rate, and Fowlkes-Mallows Scores. Expectation maximization (EM) calculation is utilized for evaluating the presentation of the Gaussian mixture model. The performance of classi er's presentation is also done. Accuracy is accomplished in the event of grouping by means of SVM classi er as 89% in the supervised method.

1. Introduction

Cybercrime involves attempting a criminal offense via computer and a network wherein the computer can act as a tool, goal, or both of them. Many unauthorized computer enabled activities take place via global electronic networks. The research categorizes cybercrime into the following:

- (i) A criminal activity involving computer for the execution of the crime
- (ii) A criminal activity involving computer for crime related information storage and not necessarily for the execution of the crime

1.1. Necessity of Detection of Cybercrime. Cybercrime has spread its roots far on a global scale and portrays a great danger towards occurrence of either a criminal or a terrorist activity. These threats can affect both internal and external security (military) without reacting to single authority policing methodologies. There is loss of both personal and nancial data if the cybercrime goes undetected. There have already been attacks on information infrastructure and Internet services. Online fraud and hacker attacks are only two instances of computer-related crimes that occur on a daily basis. Cybercrime is said to have caused huge nancial damage. Malicious software costs up to USD 17 billion in damages in 2003 alone. According to some estimates,

