Abstract

Over the years, the Internet has overwhelmed humankind and increased the dependence of human civilization on it. With the rising number of users, the Internet has become the primary source of security threats to computing systems. These computing systems, viz., computers, mobiles, and now IoT devices, have proliferated in the past decade. Amongst these, mobiles have become nearly ubiquitous with human existence. With the number of people browsing the Internet increasing exponentially, web-based attacks have become the most preferred means of attack. Security firms are now trying to battle these web-based attacks. However, these attacks are evolving rapidly, making it difficult for previous generations of security solutions to keep pace with them. In the present cybersecurity scenario, Artificial Intelligence (AI) enthuses hope to overcome these rapidly evolving web threats. This hope emanates from the fact that AI has made rapid advances in the past decade and is now influencing the growth of every other field. With such benefits being accrued from AI in every arena, the area of web security cannot afford to stay behind from reaping the benefits of AI. The work done in this thesis is a step in this direction. The research covered in this thesis attempts to solve critical web security problems with AI.

The thesis has taken on web security challenges on both computer and mobile platforms. The following are the contributions of the thesis in the area of web security:

- A focused web crawler named 'MalCrawler', which seeks and crawls websites. This crawler facilitates the collection of webpages, especially malicious webpages. Compared to a generic crawler, it enables a higher than normal collection of malicious webpages. Moreover, it is designed

to overcome the evasion techniques employed by malicious websites. Thus, the crawler stands as a significant contribution in the field of web security, as it enables the collection of webpages, especially malicious, for preparing datasets that can be used for ML based analysis and solutions.

- Presents a detailed analysis of attributes that can be used for the classification of malicious webpages. It evaluates various attributes using information gain, resource utilization in pre-processing, and prediction performance with different Conventional ML algorithms. Furthermore, using these attributes, it proposes multiple Conventional ML based models to predict malicious webpages with improved accuracy.

- Developed suitable deep learning models to predict malicious webpages using both structured and raw data as input. It has used both supervised and unsupervised approaches for producing these deep learning models. Using state-of-the-art techniques, these Deep learning models have surpassed previous performance metrics.

- Carried out a detailed ML-based security analysis of hybrid apps on the Android mobile platform. Based on the analysis, it detected vulnerable hybrid apps on the Google Play store. Further, two Android apps were developed as part of the research. The first app, named 'WebView Tool', helps to understand the Android WebView component and the hybrid apps that are made using this component. The second app, called 'WebView Monitor', monitors the security of the hybrid apps in the background.

- Provided a Federated Learning (FL) based web security solution for the mobile platform. This novel solution is proposed as a futuristic and privacy-preserving mobile security solution, which can handle newer ever-evolving threats. Also, the solution dovetailed the latest research in differential privacy, secure aggregation, and Hierarchical FL (HFL) to improve the overall effectiveness of the approach. Seeing the results

achieved, the work has set the pace for the next generation of mobile security solutions using state-of-the-art FL technology.

All the work done in this thesis has been implemented, experimented with, and benchmarked with existing solutions, if any. Overall, the research carried out as part of this thesis has made a significant contribution in the field of web security using the latest advancements in AI!