

ABSTRACT

Malicious Web sites largely promote the growth of Internet criminal activities and constrain the development of Web services. As a result, there has been strong motivation to develop systemic solution to stopping the user from visiting such Web sites. We propose a learning-based approach to classifying Web sites into 3 classes: Benign, Spam and Malicious. Our mechanism only analyzes the Uniform Resource Locator (URL) itself without accessing the content of Web sites. Thus, it eliminates the run time latency and the possibility of exposing users to the browser based vulnerabilities. By employing learning algorithms, our scheme achieves better performance on generality and coverage compared with blacklisting service. The main objective of this project is to detect the malicious malware