A Dynamic Method of Detecting Malicious Scripts Using Classifiers

Article *in* Journal of Computational and Theoretical Nanoscience · June 2017

DOI: 10.1166/asl.2017.7374

CITATIONS
6

READS
73

3 authors, including:

Nayeem Khan
Albaha University
17 PUBLICATIONS **112** CITATIONS
SEE PROFILE

Shahid Khan
Hämeen ammattikorkeakoulu University of Applied Sciences
136 PUBLICATIONS **1,101** CITATIONS
SEE PROFILE

Advanced Science Letters

A Dynamic Method of Detecting Malicious Scripts Using Classifiers

Nayeem Khan, Johari Abdullah, Adnan Shahid Khan
Department of Computer Systems & Communication Technologies
Faculty of Computer Science & Information Technology
University Malaysia Sarawak
94300 Kota Samarahan, Sarawak, Malaysia
Email: 15010049@siswa.unimas.my, {ajohari, skadnan}@fit.unimas.my

Due to the increasing importance of Internet in every aspect of our life, the World Wide Web which is accessed by end users through web browsers is becoming the next platform for criminal or individual with the malicious intent to conduct malicious activities either for personal or economic gains. Malicious scripts work as a primary source of infection for malicious software or also known as malware. This paper proposes an efficient method of detecting malicious scripts by employing an interceptor on the client side by using a set of supervised and unsupervised classifiers. The proposed method will be implemented to achieve high detection rate with low false alarms and minimal performance overheads.

Keywords: Web Security, XSS, Detection, Interceptor, Machine Learning, Supervised Classifiers, Unsupervised Classifiers

1. INTRODUCTION

Due to the increasing importance of Internet in every aspect of our life, the World Wide Web which is accessed by end users through web browsers is becoming the next platform for criminal or individual with the malicious intent to conduct malicious activities either for personal or economic gains. The top most security challenges are that malicious code is being sent to individuals and organisations by cyber criminals to attack for their interests. As per authors, [1, 2, 3] malware or malicious software detection refers to the process of detecting weather a given program *P* is malicious or benign based on prior knowledge. The purpose of malware is to gain sensitive information from the users without his/her knowledge, disrupt normal computer operations, display unwanted advertisements, sending spam emails and to extort money etc.

As per OWASP cross site scripting (XSS) is the 2nd among the top ten security vulnerabilities. Among all the reported vulnerabilities XSS accounted 43% as per Web Incident Database. XSS is vulnerability in a web application that provides a way for the cyber criminals to implant or inject malicious scripts into a source code of a web page to perform an attack. Malicious code which is sent from server to client could take several forms including JavaScript, VBScript, Flash, Active, HTML and any other sort of code which a browser is able to execute. This exploitation could materialize only when the code send is not properly sanitized. XSS which is considered as vulnerability on sever side used to target users on the client side via web browsers. A number of solutions for protection against attacks have been proposed but the fail in one way or the other. Many authors [4, 5, 6] have strongly advocated that there is no such solution on the server side which will completely be able to protect client side by getting attacked.

In order to prevent any threat, common users on the client side are dependent on anti-viruses to protect their systems. The architecture of most of the antiviruses is based on signature and pattern matching for detection of attacks. The signature based mechanism in anti-viruses couldn't provide a full security protection with the reason that signature based approaches need regular updating of malicious patterns in order to match. Signature based detection approaches can only detect previously known threats only. Another problem associated with signature based approach is that it has high false alarm rate with an