Journal of Network and Computer Applications 60 (2016) 95–112

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

Review

Web application protection techniques: A taxonomy

Victor Prokhorenko, Kim-Kwang Raymond Choo, Helen Ashman

University of South Australia, Australia

ARTICLE INFO

Article history:
Received 16 July 2015
Received in revised form
28 October 2015
Accepted 19 November 2015
Available online 11 December 2015

Keywords:
Web IDS
Web protection
Web application
Web security

ABSTRACT

The growing popularity of web applications makes them an attractive target for malicious users. Large amounts of private data commonly processed and stored by web applications are a valuable asset for attackers, resulting in more sophisticated web-oriented attacks. Therefore, multiple web application protections have been proposed. Such protections range from narrow, vector-specific solutions used to prevent some attacks only, to generic development practices aiming to build secure software from the ground up. However, due to the diversity of the proposed protection methods, choosing one to protect an existing or a planned application becomes an issue of its own.

This paper surveys the web application protection techniques, aiming to systematise the existing approaches into a holistic big picture. First, a general background is presented to highlight the issues specific to web applications. Then, a novel classification of the protections is provided. A variety of existing protections is overviewed and systematised next, followed by a discussion of current issues and limitation inherent to the existing protection methods. Finally, the overall picture is summarised and future potentially beneficial research lines are discussed.

Contents

E-mail addresses: Victor.Prokhorenko@mymail.unisa.edu.au (V. Prokhorenko), Raymond.Choo@unisa.edu.au (K.-K. Choo), Helen.Ashman@unisa.edu.au (H. Ashman).

http://dx.doi.org/10.1016/j.jnca.2015.11.017
1084-8045/

1. Introduction

This review examines recent web-based protection techniques from the perspective of a detailed classification. Perhaps, due to a large variety of web technologies used in modern web application development, the majority of existing reviews commonly cover only a specific type of attack or summarise a class of protection techniques (e.g. Asghar Sandu et al., 2011; Hossein Manshaei et al., 2013; Scholte et al., 2012; Du et al., 2011; Mitchell and Chen, 2014). Comprehensive reviews aiming to systematise currently existing techniques have been conducted as well (Li and Xue, 2014). However, given the multitude of incompatible views, this review aims to provide an alternative and more holistic classification of the existing web application protection techniques. A brief overview of security software evolution is first provided to explain the current landscape of various software protection techniques.

The history of software development shows that software applications are generally designed to provide features that appeal to consumers rather than with security in mind (the latter is also known as *security by design*). Retrofitting security in such software is usually not economically viable, due to the costs and resources required (e.g. integrating security in an application with a large code base without introducing compatibility issues). A common approach is to implement an additional layer of protection on top of the existing application, and these protection layers are known as *envelope protections*.

In the 1980s, the main security threats were viruses, thus antivirus software came into existence. With the development and penetration of networks in computing, firewalls emerged as a separate class of security software. Firewalls were designed to protect against malicious network-based activities (including virus propagation over networks). However, strict permission models employed by early firewalls were not flexible enough in some situations, as the control of the network connections happens at a lower level. For example, a traditional firewall could completely block a TCP port, but could not selectively pass only the packets with some specific data. This type of limitation caused Intrusion Detection Systems (IDSs) to be developed.

The purpose of an IDS is to serve as an intermediate layer between the application being protected and the users. The layer would analyse user actions and detect possibly malicious activity. Once a suspicious activity is detected, such a protection layer would raise an alarm to warn the system administrator. An obvious improvement is to not only detect, but also prevent (if possible) the attack, as implemented by Intrusion Prevention Systems (IPSs).

Another firewall improvement emerged due to the growth of computing power available. As computational speeds increased, a direct analysis of each byte in all the network packets became feasible for firewalls. Such technology was named Deep Packet Inspection (DPI) and is now commonly used on the market. In contrast to traditional firewalls which made decisions on whether to pass or reject a packet based on the packet headers, a DPI firewall can analyse the packet data payload as well.

In some ways, a network-level IDS can be considered to be the same thing as a DPI firewall. Both approaches require analysing all of the network packets passing through, paying attention to the packet data. The difference mainly lies in the decision phase, whereas the data capturing and analysis are technically the same.

Web technologies were born out of the necessity to communicate with remote clients. With further growth of the networks, the popularity of web technologies increased. Initially, web sites were used to serve static content, however with growing user demands, various techniques were introduced to serve dynamic content as well. Such techniques include both client- and server-side scripting. While the code executed at the client side does not directly affect the web application, introducing the notion of server-side programming brought complex web applications into existence. The increasing complexity of the server-side code caused bugs to appear in web applications making it possible to perform successful attacks.

Due to conceptual differences introduced by web technologies, traditional protection mechanisms such as antiviruses, firewalls or network level IDSs are not directly applicable to protecting web sites. For example, non-DPI firewalls would only either provide or block access to the web server as a whole. Network level IDSs would face the same kind of limitation without having additional information about the web application structure. Traditional host-side antivirus protections solve a somewhat different problem of detecting and preventing malicious code execution. Modern antiviruses tend to monitor multiple virus intrusion vectors such as network or USB disks and the strict difference between antiviruses