php 程序中比较经常用到的remove_xss 通防绕过

笔记本： 学习记录
创建时间： 2017/1/24 14:44 **更新时间：** 2017/2/7 15:31
作者： 546325574@qq.com
URL： http://www.freebuf.com/articles/web/55505.html

```
<?php

function remove_xss($string) {
    $string = preg_replace('/[\x00-\x08\x0B\x0C\x0E-\x1F\x7F]+/S', '', $string);

    $parm1 = Array('javascript', 'vbscript', 'expression', 'applet', 'meta', 'xml', 'blink', 'link',
'script', 'embed', 'object', 'iframe', 'frame', 'frameset', 'ilayer', 'layer', 'bgsound', 'title',
'base');

    $parm2 = Array('onabort', 'onactivate', 'onafterprint', 'onafterupdate',
'onbeforeactivate', 'onbeforecopy', 'onbeforecut', 'onbeforedeactivate',
'onbeforeeditfocus', 'onbeforepaste', 'onbeforeprint', 'onbeforeunload', 'onbeforeupdate',
'onblur', 'onbounce', 'oncellchange', 'onchange', 'onclick', 'oncontextmenu',
'oncontrolselect', 'oncopy', 'oncut', 'ondataavailable', 'ondatasetchanged',
'ondatasetcomplete', 'ondblclick', 'ondeactivate', 'ondrag', 'ondragend', 'ondragenter',
'ondragleave', 'ondragover', 'ondragstart', 'ondrop', 'onerror', 'onerrorupdate',
'onfilterchange', 'onfinish', 'onfocus', 'onfocusin', 'onfocusout', 'onhelp', 'onkeydown',
'onkeypress', 'onkeyup', 'onlayoutcomplete', 'onload', 'onlosecapture', 'onmousedown',
'onmouseenter', 'onmouseleave', 'onmousemove', 'onmouseout', 'onmouseover',
'onmouseup', 'onmousewheel', 'onmove', 'onmoveend', 'onmovestart', 'onpaste',
'onpropertychange', 'onreadystatechange', 'onreset', 'onresize', 'onresizeend',
'onresizestart', 'onrowenter', 'onrowexit', 'onrowsdelete', 'onrowsinserted', 'onscroll',
'onselect', 'onselectionchange', 'onselectstart', 'onstart', 'onstop', 'onsubmit', 'onunload');

    $parm = array_merge($parm1, $parm2);

    for ($i = 0; $i < sizeof($parm); $i++) {
        $pattern = '/';
        for ($j = 0; $j < strlen($parm[$i]); $j++) {
            if ($j > 0) {
                $pattern .= '(';
                $pattern .= '(&#[x|X]0([9][a][b]);?)?';
                $pattern .= '|(&#0([9][10][13]);?)?';
                $pattern .= ')?';
            }
            $pattern .= $parm[$i][$j];
        }
```

```
        $pattern .= '/i';
        $string = preg_replace($pattern, ' ', $string);
    }
    return $string;
}

$str='<a href="javascrip&#116&#58alert(/xss/);">asd </a>';
echo remove_xss($str);

?>
```

++++++++++++++++++++++++++++++++++++++

经常有白名单限制

但是允许你输入a标签

所以 payload 是 asd

漏洞原理，这里过滤了Javacript base xml 所以一些伪协议是不能执行的。过滤了像link，object等一些常用的冷门标签。

不考虑html5或者啥最新出来 的新标签，新功能的情况。可以尝试编码绕过。

参考http://www.freebuf.com/articles/web/55505.html

http://bobao.360.cn/learning/detail/292.html

HTML编码的存在就是让他在代码中和显示中分开， 避免错误。他的命名实体：构造是&加上希腊字母，字符编码：构造是&#加十进制、十六进制ASCII码或unicode字符编码，而且浏览器解析的时候会先把html编码解析再进行渲染。但是有个前提就是必须要在“值”里，比如属性src里，但却不能对src进行html编码。不然浏览器无法正常的渲染。

这里 src ， href 是标签属性，支持html实体编码。所以直接编码即可绕过 函数检查