

人工智能中心-数据抓取组



js混淆的利与弊



常见Js混淆与反混淆方法

TITLE CONTENTS

目 录



高级Js混淆与解析方法



项目分析

一、js混淆的利与弊

Js混淆的利与弊

站在网站开发者的角度

- 1、是为了保护我们的前端代码逻辑
- 2、精简代码、加快传输

站在爬虫者的角度

- 1、增加了获取数据的难度
- 2、增加了获取数据的难度

二、常见JS混淆方法

常见JS混淆方法(一)

JS压缩

特征:

原理:

削减是一个从源代码中删除不必要的字符的技术使它看起来简单而整洁。这种技术也被称为代码压缩和最小化

常见JS混淆方法(一)

常用压缩工具:

https://javascript-minifier.com/

破解方法-代码格式化:

http://tool.oschina.net/codeformat/js/

常见JS混淆方法(二)

Base62混淆

特征:

最明显的特征是生成的代码以eval (function (p, a, c, k, e, r)) 开头

常见JS混淆方法(二)

原理:

这类混淆的关键思想在于将需要执行的代码进行一次编码,在执行的时候还原出浏览器可执行的合法的脚本

常用混淆工具:

http://dean.edwards.name/packer/

破解方法-浏览器

- 1. 打开 谷歌 或者 火狐 浏览器
- 2. 按 F12 打开控制台
- 3. 把代码复制进去
- 4. 删除开头 eval 这4个字母
- 5. 按回车键

常见JS混淆方法(三)

公钥加密关键字段

特征:

```
var publicKey = "MIGfMA0GCSqGSIb3DQEBAQUAA4GNADCBiQKBgQCXC4vf5+K50WpDTSq/wxwjHGZm"
+ "suvmeSkVQzXMk+yfABciLK0EyEfVyf4zH19yoSigS4k6shS2Esx4aZqPxcJjFgvt"
+ "PxNAEARsvvFvb7hPqFWwcAOWWozTxT7yjM4x0P/bmbClvMfeWGzE/OQZFRQOQxfg"
+ "ReeSKR9IQphr//mAdQIDAQAB";
var encrypt = new JSEncrypt();
    encrypt.setPublicKey(publicKey);
    //将data数组赋给ajax对象
    var PassWords = encrypt.encrypt(password);
    var rst = {username:userName,password:PassWords};
    dmsCommon.ajaxRestRequest({
        url: dmsCommon.getDmsPath()["sysAuth"] + "/login",
        data: JSON.stringify(rst),
        type: 'POST',
        sucessCallBack: function (response) {
```

常见JS混淆方法(三)

原理:

- 1. 用户访问客户端,客户端向服务器请求获取一个RSA公钥以及键值key,存储在本地
- 2. 用户在本地公钥失效前发起登录请求,则使用已有公钥对用户密码进行加密; 若已过期则执行1后再加密
- 3. 客户端将密文与key一起传回后台
- 4. 后台通过key找到缓存里面的私钥,对密文进行解密

常用混淆工具:

http://travistidwell.com/jsencrypt/

破解方法:

from Crypto.Cipher import PKCS1_v1_5 as Cipher_pkcs1v1_5 from Crypto.PublicKey import RSA

三、高级JS混淆方法

高级JS混淆方法(一)

变量名混淆

特征:

```
var _0x1b14=["\x6C\x65\x6E\x67\x74\x68","","\x66\x72\x6F\x6D\x43\x68\x61\x72\x43\x6F\x64\x65","\x63\x68\x61\x72\x43\x6F\x64\x65\x41\x74","\x63\x6F\x6E\x63
   \x61\x74","\x30\x31\x32\x33\x34\x35\x36\x37\x38\x39\x61\x62\x63\x64\x65\x66","\x63\x68\x61\x72\x41\x74","\x41\x42\x43\x44\x45\x46\x47\x48\x49\x4A\x4B
   \x4C\x4D\x4E\x4F\x50\x51\x52\x53\x54\x55\x56\x57\x58\x59\x5A\x61\x62\x63\x64\x65\x66\x67\x68\x6B\x6C\x6D\x6E\x6F\x70\x71\x72\x73\x74\x75\x76
   \x77\x78\x79\x74","\x72\x61\x6E\x64\x6F\x6D","\x66\x6C\x6F\x6F\x72","\x72\x65\x70\x6C\x61\x63\x65","\x73\x70\x6C\x69\x74","\x64\x6F\x69\x6E","
   \x67\x65\x74\x54\x69\x6D\x65","\x73\x65\x74\x54\x69\x6D\x65","\x78\x20\x65\x78\x70\x69\x72\x65\x73\x3D","\x74\x6F\x55\x54\x43\x53\x74\x72\x69\x6E\x67"
   "\x63\x6F\x6F\x6B\x69\x65","\x3D","\x3B\x20\x70\x61\x74\x68\x3D\x2F","\x69\x6E\x64\x65\x78\x4F\x66","\x63\x6F\x6F\x6F\x6F\x65\x45\x6E\x61\x62\x6C\x65
   \x64","\u8BF7\u4FEE\u6539\u6D4F\u89C8\u5668\u8BBE\u7F6E\x2C\u5141\u8BB8\x63\x6F\x6B\x69\x65\u7F13\u5B58","\x68\x72\x65\x66","\x70\x72\x6F\x74\x6F
   \x63\x6F\x6C","\x68\x74\x70\x73\x3A","\x6C\x6F\x61\x74\x69\x6F\x6E","\x73\x75\x62\x73\x74\x72\x69\x6E","\x58\x59\x32\x6C\x35\x53\x53\x4E
   \x4D\x4F\x6C\x7A\x43\x75\x4F\x57\x50\x72\x56\x41\x51\x36\x4F\x38\x62\x6D\x42\x31\x30\x52\x39\x30\x35\x32\x57\x72\x75\x53\x44\x36\x2B\x58\x77\x3D","\x31
   \x39\x33\x31\x33\x38\x37\x37\x35\x39\x37\x32\x31\x33","\x61\x6E\x74\x69\x70\x61\x73","\x68\x74\x70\x73\x38\x2F\x2F"]; function safeAdd( @xcab3x2,
   <u>_0xcab3x3</u>){var _0xcab3x4=(_0xcab3x2& 0xFFFF)+ (_0xcab3x3& 0xFFFF);var _0xcab3x5=(_0xcab3x2>> 16)+ (_0xcab3x3>> 16)+ (_0xcab3x4>> 16);return (_0xcab3x5
   << 16)| (_0xcab3x4& 0xFFFF)}function bitRotateLeft(_0xcab3x7,_0xcab3x8){return (_0xcab3x7<< _0xcab3x8)| (_0xcab3x7>>> (32- _0xcab3x8))}function cmn(
   _@xcab3xa,_@xcab3xb,_@xcab3xc,_@xcab3x2,_@xcab3xd,_@xcab3xe){return safeAdd(bitRotateLeft(safeAdd(safeAdd(_@xcab3xb,_@xcab3xa),safeAdd(_@xcab3x2,
   Oxcab3xe)), Oxcab3xd), Oxcab3xc)}function ff(_0xcab3xb,_0xcab3xc,_0xcab3x10,_0xcab3x11,_0xcab3x2,_0xcab3xd,_0xcab3xe){return cmn((_0xcab3xc&
   _0xcab3x10)| ((~_0xcab3xc) & _0xcab3x11),_0xcab3xb,_0xcab3xc,_0xcab3x2,_0xcab3xd,_0xcab3xe)}function gg(_0xcab3xb,_0xcab3xc,_0xcab3x10,_0xcab3x11,
    0xcab3x2, 0xcab3xd, 0xcab3xe){return cmn(( 0xcab3xc& 0xcab3x11)| ( 0xcab3x10& (~ 0xcab3x11)), 0xcab3xb, 0xcab3xc, 0xcab3x2, 0xcab3xd, 0xcab3xd, 0xcab3xd)
```

高级JS混淆方法(一)

原理:

字符串字面量混淆: 首先提取全部的字符串, 在全局作用域创建一个字符串数组, 同时转义字符增大阅读难度, 然后将字符串出现的地方替换成为数组元素的引用

变量名混淆:不同于压缩器的缩短命名,此处使用了下划线加数字的格式,变量之间区分度很低,相比单个字母更难以阅读

成员运算符混淆:将点运算符替换为字符串下标形式,然后对字符串进行混淆删除多余的空白字符:减小文件体积,这是所有压缩器都会做的事

常用混淆工具:

http://javascriptobfuscator.com/Javascript-Obfuscator.aspx

http://js.51tools.info/

破解方法-IDE、解密工具、浏览器:

http://jsnice.org/

http://js.51tools.info/

高级JS混淆方法(二)

使用特定符号编写js脚本

特征:

高级JS混淆方法(二)

原理:

jsfuck源于一门编程语言brainfuck,其主要的思想就是只使用8种特定的符号来编写代码。而jsfuck也是沿用了这个思想,它仅仅使用6种符号来编写代码。它们分别是(、)、+、[、]、!。

常用混淆工具:

http://www.jsfuck.com/

破解方法:

无

高级JS混淆方法(三)

引入加密密钥串

特征:

高级JS混淆方法(三)

原理:

加密核心函数,保护关键代码

常用混淆工具:

无

破解方法:

浏览器

高级JS混淆方法(四)

对浏览器自动化检测

https://defyun.oss-cn-shanghai.aliyuncs.com/areyouhuman.html

,	Are you human?
Im huma	n
Submit	

高级JS混淆方法(四)

原理:

通过js脚本检测环境中是否存在selenium调用浏览器参数

破解方法:

通过Linux shell命令启动浏览器

```
runBotDetection = function () {
    var documentDetectionKeys = [
        " webdriver evaluate",
          selenium evaluate",
          webdriver script function",
          webdriver script func",
          webdriver script fn",
          fxdriver evaluate",
          driver unwrapped",
          webdriver unwrapped",
           driver evaluate",
          selenium unwrapped",
         fxdriver_unwrapped",
    var windowDetectionKeys = [
        " phantom",
        " nightmare",
        " selenium",
        "callPhantom",
        "callSelenium",
        " Selenium IDE Recorder",
    ];
```

四、项目分析

瓜子二手车js混淆分析 路虎登录界面js混淆分析 蝉大师js混淆分析 抖音视频js混淆分析