美容与野兽：

转向现代网络浏览器来构建独特的浏览器指纹

1. 浏览器指纹的介绍
2. 论文内容的介绍
3. 总结

1.1 什么是浏览器指纹

当你使用浏览器访问某个网站的时候，浏览器【必定会暴露】某些信息给这个网站。为什么强调“必定”呢？因为这些信息中，有些是跟 HTTP 协议相关的。只要你基于 HTTP 协议访问网站，浏览器就【必定】会传输这些信息给网站的服务器。只要你通过浏览器访问 Web，必定是基于 HTTP 协议的。因此，Web 网站的服务器必定可以获取到跟你的浏览器相关的某些信息。

1.2 与cookie的区别

两者的原理类似——都是利用某些特殊的信息来定位你的身份。两者的本质差异在于：

cookie 需要把信息保存在浏览器端，所以会被用户发现，也会被用户清除。

“浏览器指纹”无需在客户端保存任何信息，不会被用户发觉，用户也无法清除（换句话说：你甚至无法判断你访问的网站到底有没有收集浏览器指纹）。

1.3 如何暴露隐私

假如网站没有采用“指纹追踪”的技术，那么你可以在该网站上注册若干个帐号（马甲）。当你需要切换身份的时候，只需要先注销用户，清空浏览器的 cookie，然后用另一个帐号登录。网站是看不出来的。一旦网站采用“指纹追踪”的技术，即使你用上述方式伪造马甲，但因为你用的是同一个浏览器，浏览器指纹相同。网站的服务器软件可以猜测出，这两个帐号其实是同一个网民注册的。

2.1 传统网络环境

传统网络上泄露的浏览器指纹信息：

User Agent

其实简单的说User-Agent就是客户端浏览器等应用程序使用的一种特殊的网络协议，在每次浏览器（邮件客户端/搜索引擎蜘蛛）进行 HTTP 请求时发送到服务器，服务器就知道了用户是使用什么浏览器（邮件客户端/搜索引擎蜘蛛）来访问的。既然是人为规定的协议，那么就是说不管什么浏览器，默认的UA都是可以更改的。

屏幕分辨率这一项不仅包括屏幕的尺寸，还包括颜色深度（比如你的屏幕是16位色、24位色、还是32位色）。

时区

浏览器的插件信息

浏览器的字体信息

和浏览器相关的一些字体信息。 如果你的浏览器安装了 Flash 或 Java 插件，有可能会暴露某些字体信息。

2.2 背景

更互动的网络（例如，JavaScript库的繁荣，HTML5的每周创新）

更可用的网络（例如，移动设备的爆炸）

更安全的网络（例如，Flash正在消失，NPAPI插件正在被弃用）

更私人的网络（例如，增加立法反对cookies，扩展的巨大成功，如Ghostery和AdBlock）。

结论：

现代的浏览器技术，提供了美丽和力量的网络，也提供了一个黑暗的一面，一个丰富的可利用的数据，可用于构建独特的浏览器指纹的生态系统。

2.3 工作

Attribute	Source	Distinct values	Unique values	Example
User agent	HTTP header	11,237	6,559	Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/41.0.2272.118 Safari/537.36
Accept	HTTP header	131	62	text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,*/*;q=0.8
Content encoding	HTTP header	42	11	gzip, deflate, sdch
Content language	HTTP header	4,694	2,887	en-us,en;q=0.5
List of plugins	JavaScript	47,057	39,797	Plugin 1: Chrome PDF Viewer. Plugin 2: Chrome Remote Desktop Viewer. Plugin 3: Native Client. Plugin 4: Shockwave Flash...
Cookies enabled	JavaScript	2	0	yes
Use of local/session storage	JavaScript	2	0	yes
Timezone	JavaScript	55	6	-60 (UTC+1)
Screen resolution and color depth	JavaScript	2,689	1,666	1920x1200x24
List of fonts	Flash plugin	36,202	31,007	Abyssinica SIL,Aharoni CLM,AR PL UMing CN,AR PL UMing HK,AR PL UMing TW...
List of HTTP headers	HTTP headers	1,182	525	Referer X-Forwarded-For Connection Accept Cookie Accept-Language Accept-Encoding User-Agent Host
Platform	JavaScript	187	99	Linux x86_64
Do Not Track	JavaScript	7	0	yes
Canvas	JavaScript	8,375	5,533	Cwm fjordbank glyphs vext quiz, 😃 Cwm fjordbank glyphs vext quiz, 😃
WebGL Vendor	JavaScript	26	2	NVIDIA Corporation
WebGL Renderer	JavaScript	1,732	649	GeForce GTX 650 Ti/PCIe/SSE2
Use of an ad blocker	JavaScript	2	0	no

提供了

一个使用现代网络技术的

17属性指纹识别脚本

2.3 属性

•HTTP头列表：当连接到服务器时，浏览器发送用户代理，网页所需的语言，浏览器支持的编码类型以及其他头部。一些软件和浏览器扩展修改或添加标题，提供有关设备配置的额外详细信息。在HTTP协议中定义，这些头部总是可以由服务器获取，并且不依赖于JavaScript。

•平台：“navigator.platform”属性中的值提供有关用户操作系统的信息。虽然此信息已经在用户代理中，但我们收集“平台”值以检测修改或不一致的指纹，例如，返回的值与user-agent中的值不同

2.3 属性

•不跟踪/使用广告拦截器：这两个属性与隐私直接相关，这些值可以帮助我们区分具有隐私权的用户与其他人。

•WebGL供应商和渲染器：这两个属性与HTML WebGL API一起添加，以提供关于设备的底层GPU的信息。

•帆布：HTML5 Canvas元素使我们能够通过请求浏览器在一组固定的指令之后渲染图片，从而对硬件和操作系统执行测试。

3. 总结

工作重点是现代网络技术对通过浏览器指纹识别设备的能力的影响。了解现代网络技术提供了大大改善的用户体验，虽然有损于隐私。

提供了关于最新浏览器API的影响的新见解，包括HTML5画布上的第一个大规模分析，以及最近趋势的影响，例如Flash和其他插件的存在减少网络。


谢谢观看