



讯实网络商业白皮书

ComRatings Whitepaper

网络 **Cookie** 删除 对广告服务统计及网站流量统计造成巨大偏差

2011 年 6 月更新

June 2011 Updated

一、Cookie 概述	2
什么是 Cookie?	2
Cookie 浏览器应用管理	2
Cookie 是一个事件处理对象	2
Cookie 应用.....	3
Flash Cookie	3
Flash Cookies 和 Cookies 有什么区别?	3
Cookie 删除.....	4
第一方 Cookie vs. 第三方 Cookie	4
二、讯实网络的研究方法论.....	5
研究概述	5
研究假设	5
三、Cookie 删除行为分析	6
第一方 Cookie 删除.....	6
第三方 Cookie 删除.....	6
安全保护程序对 Cookie 删除的影响	7
四、Cookie 删除问卷调查	9
用户对 Cookie 使用的看法.....	9
用户对 Cookie 删除的看法.....	9
五、结论.....	11
讯实网络的研究揭示了如下关键发现:	11
Cookie 删除导致了下述网站用户度量的不精确:	11
Cookie 删除导致了下述广告服务度量的不精确:	11
附录.....	12
关于讯实网络 About Comratings	12
讯实网络的 <u>Hybrid</u> 复合网络广告分析和优化技术.....	12

一、Cookie 概述

为了更好地理解如何评估 Cookie 管理，有必要解释一下 Cookie 的工作原理。

什么是 Cookie?

简单来说，Cookie 就是服务器暂存在用户电脑里的资料（采用.txt 格式的文本文件），这样可以方便服务器识别用户电脑的唯一性。

当用户浏览网页时，网站 Web 服务器会在用户电脑的硬盘上植入一个非常小的 Cookie 文件。由于单个网站的所有信息，通常就被存储在一个特定的 Cookie 文件中（它可以记录用户的 ID、密码、浏览过的网页、停留的时间等信息），因此，当用户再次访问同一网站时，Web 服务器就会马上检查用户上次保留的 Cookie 资料，并依据 Cookie 里的信息识别出使用该浏览器的用户是否是重复用户。有些用户或许会注意到，当使用同一台电脑同一个浏览器再次登陆某网站时，会发现不必输入用户名和密码就已经登录了，这其实就是 cookie 的功用了。

Cookie 浏览器应用管理

Cookie 是被浏览器使用的文件，它与操作系统、ISP 提供商和互联网没有直接的关联。Cookie 一般保存在 IE 浏览器和 Firefox 浏览器的信息存储文件夹中。

值得一提的是，每个浏览器只有一个 Cookie 存储文件夹。而 IE 和 Firefox 采用不同路径的 Cookie 存放文件夹。因此，某一访问同一网站但采用不同类型浏览器的用户，会在电脑中存放两个不同路径的 Cookie。此外，每次 Windows 登录的 Cookie 也不同。例如，共享同一台 XP 电脑的两个独立登录的用户，会得到不同的 Cookie 存放文件夹。

Cookie 是一个事件处理对象

Cookie 是一个动态的事件处理对象，每当用户与网站发生互操作的时候，它都会被网站实时的读取并修改。一旦出现用户的页面需求，网站服务器都会要求浏览器上传相应的 Cookie 信息，这通常叫做“Cookie 获取”事件。当网站发起“Cookie 获取”事件时，一般会有三种浏览器响应方式：

- 第一种方式：如果一个 Cookie 处在被开启的 Cookie 文件夹内且没有过期，那么，浏览器就会返回该 Cookie 文件存储的所有信息（Cookie 通常会关联某个特定的域名，但也有被设置成关联一个特定的页面或者网址，而这种情况并不常见）。一般来讲，这个动作会保留通过浏览器上传给网站服务器的唯一识别符。
- 第二种方式：如果一个 Cookie 不在被开启的 Cookie 文件夹内或者已经过期了，那么 Cookie 存储的信息就不会从浏览器传递到网站服务器。于是，网站服务器就会要求浏览器设置一个全新的 Cookie（“Cookie 设置”事件），而新的 Cookie 也会要求网站建立一个与之对应的新的唯一识别符。
- 第三种方式：如果浏览器被设置为拒绝接收 Cookie 状态，那么就不会有 Cookie 信

息被网站服务器发现。这时,尽管网站可能会再一次发出要求浏览器设置新的 Cookie 的指令,但浏览器将不予理会。

Cookie 应用

Cookie 是构建在 HTTP 协议之上的一个应用,但目前没有出台具体针对 Cookie 设置、跟踪等问题的标准,这就导致 Cookie 应用的多样性和随意性。目前,最为典型的 Cookies 应用是,网站保留用户的初次登录信息,并在用户每次返回时,能自动“记住”用户,免去用户再次输入 ID、密码的麻烦。另一个重要应用是,网站保留用户的浏览行为,如浏览过的网页、停留的时间等信息,这样,当用户再次来到该网站时,网站通过读取 Cookies,得知用户的相关信息,就可以做出相应的动作。

Cookie 还可以跨网站应用。比如,用户登录某一网站浏览信息,而在该网站上同时还有许多来自各广告网站推送的广告信息。这些广告信息由第三方网站提供,当你点击某广告时,与该广告相关联的第三方网站就会通知你的浏览器设置与该网站相联系的 Cookie,第三方网站通过 Cookie 跟踪用户行为,进而可以找到最佳的广告投放站点。

Flash Cookie

Flash Cookie 是由 FlashPlayer 控制的客户端共享存储技术,它具备以下特点:

- 类似 HTTP Cookie, Flash Cookie 利用 SharedObject 类实现本地存储信息, SharedObject 类用于在用户计算机上读取和存储有限的的数据量,共享对象提供永久贮存在用户计算机上的对象之间的实时数据共享;
- 本地共享对象是作为一些单独的文件来存储的,它们的文件扩展名为.SOL;
- 本地共享对象并不是基于浏览器的,所以普通的用户不容易删除它们。如果要删掉它们的话,首先要知道这些文件所在的具体位置。这使得本地共享对象能够长时间的保留在本地系统上。

Flash Cookies 和 Cookies 有什么区别?

- 存储大小不同 cookies 仅允许存储 4KB,而 flash cookies 则存储 100KB—这只是默认的,还可以调整大小。
- 存储时长不同 一般来说,cookies 是有消亡期的,它会在一段时间后自动消失;而 flash cookies 并不,如果你没有删除它,它就永远保留在你的电脑上。
- 存储位置不同 普通 cookies 的位置人们并不需要知道,因为他们可以通过许多软件进行删除,甚至浏览器本身都内置了这一功能。而 flash cookies 则是存储在 C:\Documents and Settings\用户名\Application Data\Macromedia\Flesh Player 文件夹下。其中#sharedobjects 文件夹用于存储 flash cookies, macromedia.com 存储 flash cookies 的全局设置。

Cookie 删除

Cookie 删除，是指从用户电脑中删除 Cookie。Cookie 删除方式包括：

- 用户手工删除 Cookie（从该用户的 Cookie 文件夹中）；
- 使用诸如 IE 浏览器的“Internet 选项”来删除 Cookie；
- 使用安全保护程序来清除 Cookie。

用户可以通过设置浏览器让其处于“拒绝接收 Cookie”状态。互联网广告管理局（IAB）发布的相关研究显示，全球有 12% 的用户拒绝接受 Cookie。由于我们讨论的是这个动态的服务在服务器日志中如何夸大了独立用户，所以浏览器设置为拒绝接收 Cookie 的用户不包含在我们的研究范畴之内。

以基于服务器的统计方法来看，Cookie 删除的重要影响是，单个用户可能被误认为是多个访问者。下面的例子是说一个月内访问了某网站四次，但这期间重设了两次 Cookie 的用户在服务器中可能被计数了三次，故网站将这个独立用户当作了三个独立用户来看待。

第一方 Cookie vs. 第三方 Cookie

与某一网站直接传送页面请求相联系的 Cookies 被认为是第一方 Cookie，这样的 Cookie 通常被用来改善某网站的用户体验。有时，用户明知存在第一方 Cookie，但他仍然会浏览该网站。

第三方 Cookie 通常在不引人注意的地方起作用，它常与另一网站 Web 页面内的某个目标相联系，并不是通过用户直接请求。这些 Cookie 可能与广告，嵌入式内容等者丰富的媒体应用相联系。

第三方 Cookie 可能在网页浏览的中间活动过程中被设置，比如在浏览由第三方提供的呈现于某页面的广告的过程中就很有可能设置了一个第三方 Cookie。第三方 Cookie 的很多应用为在广泛的网络中进行用户追踪提供了手段。很多第三方 Cookie 被认为是“跟踪 Cookie”，这个词带有一点侵犯隐私的意味。这因为如此很多第三方 Cookie 应用程序会自动标记或者清除起跟踪作用的 Cookie。

二、讯实网络的研究方法论

研究概述

本研究旨在确定：用 Cookie 的唯一识别符来测量某个网站的用户数量，会存在多大的统计偏差？本研究的目的是为了得到重设 Cookie 用户的总体比例，研究对象是独立用户，而不是某个网站或 Cookie。

一般来讲，有两种类型的 Cookies 识别方式：

- Cookie 登录识别：用户须登录验证，Cookies 才可识别出用户的唯一识别符
- Cookie 被动识别：用户无须登陆验证，Cookies 即可识别出用户的唯一识别符

以上两种方式的主要区别在于，一旦发现没有 Cookies 信息对应用户的页面需求，网站服务器会做出如何反应？

- 对于登陆 Cookie 识别来讲，网站会发送一个通用的表格，或者提示用户重新登录，之后网站会重设用户的 Cookies 唯一识别符。
- 对于被动 Cookie 识别来讲，由于缺少有效的登录验证来重建用户的唯一识别符，因此，假如网站未发现对应用户网页需求的识别符，那么网站将会重设一个新的唯一识别符。

本研究通过评估被动 Cookie 识别的一致性来估计 Cookie 重设所造成的统计偏差。需要强调的是，那些设置成拒绝接收 Cookies 的电脑不在本次研究对象范围内。

研究假设

为了评估每个 Cookie 的“唯一性”，讯实网络对被测对象进行了长时间的跟踪观测。本次研究基于的基本假设是，被测对象最初的 Cookie 识别符一直存在，观测的第一个值应该等于观测的最后一个值。但如果一个新的观测出现在这段时间内，并且在接下来的观测期间一直出现，则认为最初的 cookie 没有被保存下来。

据此，将调查对象分为两组：一组是 Cookie 标识被保存的那些调查对象，另一组是没有被保存的调查对象。

三、Cookie 删除行为分析

第一方 Cookie 删除

以讯实网络的样本库为基础，发现每台电脑平均有 2.5 个不同的 Cookie。此数据表明，根据站服务器日志所统计的独立用户数可能是实际的 2.5 倍，即夸大了 150%。实际的高估程度依赖于访问该网站的频率。访问该网站越频繁，夸大的程度就越高。

第一方 Cookies			
	重设 Cookies 的个人电脑数量占比	平均每台个人电脑的 Cookies 数量	Cookies 数量占比
所有被测个人电脑	100.0%	2.5	100.0%
未作 Cookies 重设	69.3%	1.5	41.8%
做过 1 次以上重设	30.7%	4.7	58.2%
1 次重设	16.1%	2.0	12.8%
2 次重设	5.1%	3.0	6.1%
3 次重设	2.5%	4.0	4.0%
4 次以上重设	7.1%	12.5	35.3%

数据来源：© 2009 讯实网络 ComRatings

由上表可以看出，大约有 31% 的网民重设过第一方 Cookie。在这部分用户中，平均每个站点有 4.7 个不同的 Cookie。在那些重设 Cookie 的电脑中，重设一次的用户最多（占总数的 16%）。然而，重设次数超过 4 次的用户尽管只占 7%，但其重设 Cookies 数量的占比却高达 35.3%。也就是说，这一小部分用户的行为是夸大网站独立用户数的主要原因。

第三方 Cookie 删除

通过对第三方 Cookie 的分析发现，平均每台电脑有 2.6 个不同的 Cookies，这个结果和第一方 Cookie 结果差不多。这一发现也彻底颠覆了先前的观念。过去，人们多以为第三方 Cookie 的删除率应该比第一方 Cookie 的删除率高。这是因为第三方 Cookie 更具隐私侵犯性，因此，用户更倾向于开启安全保护程序删除第三方 Cookie。

第三方 Cookies			
	重设 Cookies 的个人电脑数量 占比	平均每台个人电脑的 Cookies 数量	Cookies 数量占比
所有被测个人电脑	100.0%	2.6	100.0%
未作 Cookies 重设	73.0%	1.5	42.6%
做过 1 次以上重设	27.0%	5.5	57.4%
1 次重设	13.6%	2.0	10.6%
2 次重设	4.2%	3.0	4.9%
3 次重设	2.2%	4.0	3.5%
4 次以上重设	7.0%	14.2	38.4%

数据来源：© 2009 讯实网络 ComRatings

为进一步分析此问题，讯实网络比较了开启安全保护程序的电脑的第一方 Cookie 和第三方 Cookie 的删除率。

安全保护程序对 Cookie 删除的影响

在所监测的样本 PC 中，尽管第一方 Cookie 和第三方 Cookie 的删除率相近，讯实网络对开启安全保护程序(SPP)电脑的 Cookie 的解释揭示了第一方 Cookie 和第三方 Cookie 的区别。

开启 SPP 后，平均每台电脑有 2.5 个第一方 Cookie，重设次数超过 4 次的用户电脑仅占 7.0%，却包含了近 36%的第一方 Cookie。

开启 SPP：第一方 Cookies			
	重设 Cookies 的个人电脑数量 占比	平均每台个人电脑的 Cookies 数量	Cookies 数量占 比
所有被测个人电脑	100.0%	2.5	100.0%
未作 Cookies 重设	68.4%	1.5	40.5%
做过 1 次以上重设	31.6%	4.7	59.5%
1 次重设	16.7%	2.0	13.4%
2 次重设	5.2%	3.0	6.3%
3 次重设	2.7%	4.0	4.3%
4 次以上重设	7.0%	12.6	35.6%

数据来源：© 2009 讯实网络 ComRatings

然而，从开启 SPP 的电脑中又可以看出，第三方 Cookie 的删除率明显高得多，平均每台电脑大约有 3 个 Cookie。如此高的删除率，原因之一是由于用户的重度删除（4 次以上重设 Cookie）行为所致。

开启 SPP: 第三方 Cookies			
	重设 Cookies 的电脑数量 占比	平均每台电脑的 Cookies 数量	Cookies 数 量占比
所有被测个人电脑	100.0%	3.0	100.0%
未作 Cookies 重设	62.6%	1.6	34.0%
做过 1 次以上重设	37.4%	5.3	66.0%
1 次重设	17.2%	2.0	11.4%
2 次重设	6.4%	3.0	6.4%
3 次重设	3.9%	4.0	5.2%
4 次以上重设	9.9%	13.1	43.1%

数据来源: © 2009 讯实网络 ComRatings

四、Cookie 删除问卷调查

为了获取 Cookie 删除的进一步认识，讯实网络发起了一个针对 500 个调查对象的调查，针对他们对 Cookie 的知识、态度和行为进行问卷调查。从调查结果中，我们发现了很多有趣的东西，这些结果进一步验证了讯实网络的行为分析。

用户对 Cookie 使用的看法

在互联网发展初期，Cookie 获取信息时候常常会造成电脑混乱甚至牵涉到隐私侵犯。作为调查的一部分，针对上述问题，我们对被调查者设置了相应的问题。比如，存在于电脑中的 Cookie 对你的上网使用感受有何影响？



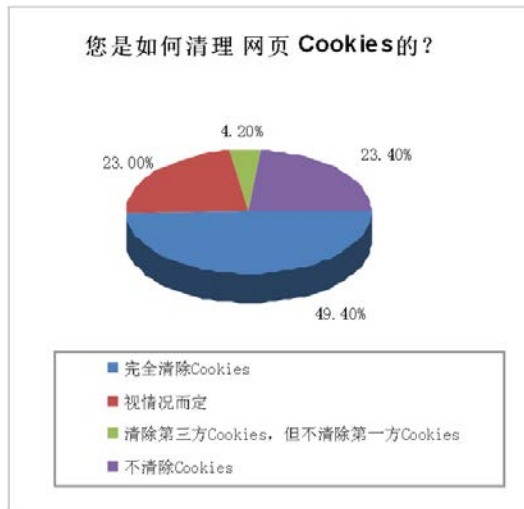
数据来源：© 2009 讯实网络 ComRatings

尽管 Cookie 有一定的负面影响，讯实网络的研究发现，对 Cookie 存在的普遍态度是中立的。特别地，将近一半（47.6%）的用户认为 Cookie 对用户上网既有正面影响也有负面影响。那些有明确立场的用户中，13% 的人认为 Cookie 能够改善上网体验，而有 15.4% 的人持有明显的反对态度；另外大约 24% 的用户对此问题没有给出确定的回应。

用户对 Cookie 删除的看法

正如前面所讨论的，讯实网络的行为研究表明第一方 Cookie 删除和第三方 Cookie 删除有惊人的相似，这与传统的观念有些违背。为了进一步核实这些发现，讯实网络专门对用户的使用习惯作了调查。

开始调查时，我们首先询问了第一方 Cookie 和第三方 Cookie 的区别。结果显示仅有 29.8% 的用户知道这一差异，其余 70.2% 的调查者要么不知道、要么不能确定。



数据来源：© 2009 讯实网络 ComRatings

通过对问卷的分析，我们发现，当调查者被问及如何清除 Cookie 时，仅有 4.2% 的用户清楚第三方 Cookie、而不清楚第一方 Cookie。这一结果和讯实网络行为研究的结果是相吻合的，它同时也证实了第一方 Cookie 删除和第三方 Cookie 删除存在的差异比较小。调查显示，最常见的删除行为是“全部删除”，有将近一半（49.4%）的被调查者都是这么做的。

五、结论

讯实网络的研究揭示了如下关键发现：

- ❖ 约有 30% 左右的用户在一个月内重设过他们的第一方 Cookie。
- ❖ 第一方 cookie 和第三方 cookie 删除率非常相似，一个月内平均每台电脑中有 2.5 个第一方 Cookie，而第三方 Cookie 有 2.6 个。
- ❖ 在安装了安全保护程序的电脑中，第三方 Cookie 删除率要比第一方 Cookie 删除率高一些。
- ❖ 用户的重度 Cookie 删除行为（4 次以上重设 Cookie）对夸大服务器日志的用户数统计有重大影响。
- ❖ 由于较高的 Cookie 删除率，如果用由以网站服务器日志端测量一个网站的独立用户，其统计结果将是实际的 2.5 倍。类似的，用于跟踪用户到达率和频次的在线营销广告服务系统统计出的结果可能是实际的 2.6 倍。

Cookie 删除的影响是深远的，它对网站用户分析及广告服务效果分析都有影响，这些影响最终导致了仅仅依赖从服务器端采集数据而产生的统计偏差。

Cookie 删除导致了下述网站用户度量的不精确：

- ❖ 夸大了独立用户数
- ❖ 低估了重复访问人数
- ❖ 低估了转换率

Cookie 删除导致了下述广告服务度量的不精确：

- ❖ 高估了到达率
- ❖ 低估了访问频次

附录.

关于讯实网络 **About Comratings**

ComRatings 是中国领先的互联网数字分析和营销优化(DAO)技术平台,为互联网营销客户和广告公司提供不同的软件应用工具,帮助这些客户分析和优化 互联网营销的工作流,深入了解数字消费者洞察,并以目标用户为核心,有效评估和提升网络广告和营销的效果。

ComRatings is the leading digital analytics & optimization (DAO) technology platform in Asia/Pacific. We provide differing software applications to the digital marketers and agencies to help them analyze and optimize the digital work flow: obtain deep digital consumer insights, measure and improve the effectiveness of online marketing programs, and directly improve their bottom line.

讯实网络的 **Hybrid 复合网络广告分析和优化技术**

讯实网络是国内唯一的将样本用户测量和网站 Cookie 测量相结合,提供网络广告数据分析和优化的技术公司。通过 Hybrid 复合网络测量和分析技术,我们可以确保网络广告和网站流量的分析能够更加精确,从而真正帮助提升网络广告和营销效果

如果您需要更多关于讯实网络方法论的信息,请访问 www.comratings.com。