루멘소프트 보안기술연구팀

Google Analytics Cookies and the Forensic Implications

Google Analytics Cookies 를 통해 포렌식적 증거 추출

Deok9

2012-02-27

Google Analytics Cookies and the Forensic Implications

일반적인 쿠키는 사용자 계정이 웹 사이트에 접속을 했는지에 대해서만 사용되었기 때문에 구조를 가지지 않지만, Google Analytics Cookie(이하 GA Cookie)는 정형화된 구조를 가지고 있다.

이는 포렌식 조사관들에게 수사 시 매우 유용한 정보를 획득 가능하게 해준다.

GA Cookie는 Google Analytics 코드를 포함한 사이트에 한해서 생성되며, 본래는 마케팅 관련 통계 용도로 쓰였다.

```
<script type="text/javascript">

 var _gaq = _gaq || [];
 _gaq.push(['_setAccount', 'UA-29715216-1']);
 _gaq.push(['_trackPageview']);

 (function() {
  var ga = document.createElement('script'); ga.type = 'text/javascript'; ga.async = true;
  ga.src = ('https:' == document.location.protocol ? 'https://ssl' : 'http://www') + '.google-analytics.com/ga.js';
  var s = document.getElementsByTagName('script')[0]; s.parentNode.insertBefore(ga, s);
 })();

</script>
```

[Google Analytics Code]

아래는 GA Cookie에 대해 간단히 설명한 표이다.

Name	Description	Expiration
__utma	This cookie is typically written to the browser upon the first visit to your site *from that Web browser*. If the cookie has been deleted by the browser operator, and the browser subsequently visits your site, a new __utma cookie is written with a different unique ID. This cookie is used to determine unique visitors to your site and is updated with each page view. Additionally, this cookie is provided with a unique ID that Google Analytics uses to ensure the validity and accessibility of the cookie as an extra security measure.	2 years from set/update.
__utmb	This cookie is used to establish and continue a user session with your site. When a user views a page on your site, the Google Analytics code attempts to update this cookie. If it does not find the cookie, a new one is written and a new session is established. This cookie expires when a user pauses on a page on your site for longer than 30 minutes.	30 minutes from set/update.
__utmc	This cookie operates in conjunction with the __utmb cookie to determine whether to establish a new session for the user. The absence of the __utmc cookie indicates that a new session needs to be established, despite that the __utmb cookie has not yet expired.	Not set.
__utmz	This cookie stores the type of referral used by the visitor to reach your site, whether via a direct method, a referring link, a Web site search, or a campaign such as an ad or an e-mail link.	6 months from set/update.
__utmv	This cookie is not normally present in a default configuration of the tracking code.	2 years from set/update.
__utmx	This cookie is not normally present in a default configuration of the tracking code.	2 years from set/update.

[Google Analytics Cookie]

가장 우측의 만료 값은 사이트 운영자에 의해 변경 될 수 있으며, 위의 쿠키들을 통하여 조금 더 나은 정보를 제공 받을 수 있다.

__utma Cookie 는 아래와 같은 구조를 가지고 있다.(로컬 시스템 기반의 UNIX 타임스탬프)

<domain hash>.<visitor ID>.<first visit>.<previous>.<last>.<# of sessions>

domain hash	하위 경로에만 관계없이 똑같은 값을 가지고 있다. -> lumensoft.co.kr/index.html과 lumensoft.co.kr은 같은 Hash intro.lumensoft.co.kr과 lumensoft.com은 다른 Hash
visitor ID	해당 사이트의 새로운 방문자에게 주어지는 유일한 ID이다. -> A가 lumensoft.co.kr에 들어간 후 B가 A의 로그인 세션이 유지되어 있는 브라우저를 통해 lumensoft.co.kr에 들어간다 하더라도 visitor ID는 같은 값일 것이다. 물론 쿠키가 삭제 된다면 새로운 visitor ID를 할당 받게 된다.
first visit	쿠키가 처음으로 시스템에 생성된 시간이며, 사이트를 처음 방문한 시간을 절대적으로 표현하지는 않는다.
previous	현재 세션 이전에 사이트를 방문 했을 때의 시간을 가진다.
last	가장 최근 방문 시간을 가진다.
# of session	HTTP Cookies와 GA Cookies의 구별을 위해 쓰일 수 있으며, 이 숫자는 사이트를 재 방문했다고 증가하지는 않는다.

[__utma Cookie]

※ __utma는 해당 쿠키가 persistent cookie 인지 확인 하는데도 유용하다.

ex) 쿠키의 마지막 업데이트 시간 2년 후 또는 관리자가 지정한 기간 후에도 해당 쿠키가 존재한다면 자동 로그인 설정과 같은 persistent cookie 이다.

__utmb 쿠키는 아래와 같은 구조를 가지고 있다.

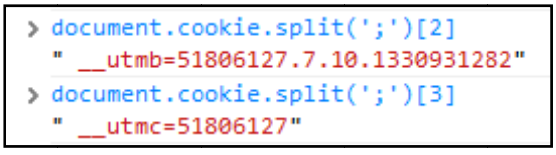
<domain hash>.<pages viewed>.10.<last time>

domain hash	__utma의 domain hash와 유사하며, __utmb가 같은 도메인이라면 해당 값은 동일하다.
page viewed	사용자가 해당 도메인에서 본 페이지 수를 나타낸다.
last time	페이지를 보거나 갱신했을 때의 값을 나타낸다.

[__utmb Cookie]

__utmc 쿠키는 domain hash만 가지고 있으며, 세션이 파기될 때 자동으로 삭제 된다.
그러므로, 이 쿠키가 존재한다면 세션이 활성 상태 인 것을 의미하며, 오직 램에만 저장되기 때문에 웹 브라우저 플러그인이나 물리 메모리 덤프를 통해서만 확인 가능하다.

아래는 크롬의 자바스크립트 콘솔을 이용하여 확인한 __utmb와 __utmc 쿠키이다.



[__utmb와 __utmc]

__utmz 쿠키는 조사에 있어 중요한 데이터들을 가장 많이 제공해 주며, 구조는 아래와 같다.

<domain hash>.<last time>.<sessions>.<sources>.<variables>

여기서 가장 중요하게 보아야 할 것은 variables 부분 이며, 아래와 같은 내용을 가지고 있다.

Name	Description
utmcsr	Last source/site used to access the target site.
utmccn	Ad campaign information commonly used with Google Adwords™. This value is usually the same as utmcmd.
utmcmd	Last type of access.
utmctr	Keyword(s) from search that found the target site (described below).
utmcct	The path to the page on the site of the referring link.

[__utmz cookie variables]

조금 더 이해를 돕기 위해 구글에서 wargame.kr을 검색한 후 해당 쿠키를 확인해 보았다.

```
> document.cookie.split(';')[4]
  " __utmz=51806127.1330932633.5.2.utmcsr=google|utmccn=(organic)|utmcmd=organic|utmctr=wargame.kr"
```

[__utmz cookie variables sample]

organic 이란, 검색엔진을 통해 해당 사이트에 접속한 것을 뜻하며 여기서는 utmcsr을 통해 구글인 것을 알 수 있다. 이 외에 사이트 배너와 같은 링크를 통해 접속했을 때에는 referral, url 주소를 직접 입력하여 접속하거나 bookmark를 통해 접속 했을 때에는 direct로 나오게 된다.

또한, utmctr값을 통해 구글에서 wargame.kr을 입력한 것도 알 수 있다.

크롬의 경우에는 쿠키를 sqlite 형태로 저장하기 때문에 sqlite browser를 통해 조금더 쉽게 확인 가능하다.(파이어 폭스도 sqlite 형태로 저장)

<user_profile>₩AppData₩Local₩Google₩Chrome₩User Data₩Default₩Cookies

creation_	host_key	name	value	path	expires_t	secure	httponly	last_acce	has_expi	persistent
1297422...	.demon,...	__utmz	9452954,...	/	1298999,...	0	0	1297422,...	1	1
1297422...	.downlo,...	__utmz	2138389,...	/	1298999,...	0	0	1297422,...	1	1
1297422...	.macup,...	__utmz	1209426,...	/	1298999,...	0	0	1297422,...	1	1
1297422...	.forensi,...	__utmz	2481173,...	/	1298999,...	0	0	1297422,...	1	1
1297422...	.clublan,...	__utmz	1115545,...	/	1298999,...	0	0	1297422,...	1	1
1297422...	.ahnlab,...	__utmz	1204702,...	/	1298999,...	0	0	1297436,...	1	1
1297422...	.framel,...	__utmz	1649665,...	/	1298999,...	0	0	1297422,...	1	1
1297422...	.forensi,...	__utmz	1161106,...	/	1298999,...	0	0	1297422,...	1	1
1297422...	.wotsit,...	__utmz	1230671,...	/	1298999,...	0	0	1297422,...	1	1
1297422...	.compu,...	__utmz	5873222,...	/	1298999,...	0	0	1297435,...	1	1
1297422...	.sans.org	__utmz	1571781,...	/	1298999,...	0	0	1297435,...	1	1
1297422...	.dumpa,...	__utmz	1288615,...	/	1298999,...	0	0	1297422,...	1	1
1297422...	.exploit,...	__utmz	1714632,...	/	1298999,...	0	0	1297422,...	1	1
1297422...	.openrc,...	__utmz	1772620,...	/	1298999,...	0	0	1297422,...	1	1
1297422...	.offensi,...	__utmz	3014652,...	/	1298999,...	0	0	1297422,...	1	1

[select * from cookies where cookies.name like "%utmz%" limit 0,15]

database 결과에서 나오는 시간의 경우 1601년 1월 1일 00:00:00을 기준으로 한 microsecond 이며 이를 보기 편하게 변환하기 위한 쿼리는 아래와 같다.

datetime(cookies.last_access_utc/100	host_key	value
2012-02-20 16:16:26	.demonoid.me	94529545,1329754587,1,1,utmcsr=dem...
2012-02-20 16:16:30	.download.cnet.com	213838977,1329754590,1,1,utmcsr=(dir...
2012-02-20 16:16:39	.macupdate.com	120942612,1329754600,1,1,utmcsr=(dir...
2012-02-20 16:16:58	.forensiccontrol.com	248117392,1329754618,1,1,utmcsr=(dir...
2012-02-20 16:17:08	.clublandmp3.com	111554518,1329754629,1,1,utmcsr=(dir...
2012-02-20 16:17:59	.frameloss.org	164966504,1329754680,1,1,utmcsr=(dir...
2012-02-20 16:18:37	.forensic.korea.ac.kr	116110609,1329754718,1,1,utmcsr=(dir...
2012-02-20 16:18:58	.wotsit.org	123067149,1329754738,1,1,utmcsr=(dir...
2012-02-20 16:19:35	.dumpanalysis.org	128861593,1329754776,1,1,utmcsr=(dir...
2012-02-20 16:19:59	.exploit-db.com	171463284,1329754800,1,1,utmcsr=(dir...
2012-02-20 16:20:03	.openrce.org	177262075,1329754803,1,1,utmccn=(dir...
2012-02-20 16:20:11	.offensivecomputing.net	30146529,1329754812,1,1,utmcsr=(dire...
2012-02-20 16:20:19	.tuts4you.com	75405463,1329754819,1,1,utmcsr=(dire...
2012-02-20 16:20:20	.winapi.co.kr	196138410,1329754821,1,1,utmcsr=(dir...
2012-02-20 16:20:51	.malc0de.com	125106710,1329754851,1,1,utmcsr=(dir...

[select datetime(cookies.last_access_utc/1000000 - 11644473600,'unixepoch'), host_key, value from cookies where cookies.name like '%utmz%' order by last_access_utc limit 0,15]

이처럼 GA Cookie는 단순히 통계 및 분석 용도 뿐만 아니라, 포렌식적 증거 추출에 도움되는 여러 가지 정보들을 제공해 주고 있기 때문에, 포렌식 조사관이라면 이에 대한 지식을 꼭 가지고 있어야 할 것이다.