Brian Pennington
is the marketing director and co-founder of Brilaw International. He provides consultancy on IT security and often speaks on the issues and problems of IT security and the Internet.

Cookies — Are they a tool for Web marketers or a breach of privacy?

Brian Pennington

Received (in revised form): 27 September 2000

Keywords: security, privacy, data protection, active code

Abstract
The foundation of any marketing campaign is the accuracy of the target data. But with the advent of the Internet, companies began to lose control of who they targeted, as anyone could visit a site from anywhere in the world and look at any product or aspect of the business. This is why programmers developed 'cookies'. This paper looks at what cookies are, what they are used for, whether they are a security risk or an infringement of privacy, and how they work.

Introduction

The foundation of any marketing campaign is the accuracy of the target data. With the advent of the Internet, companies began to lose control of who they targeted, as anyone could visit a site from anywhere in the world and look at any product or aspect of the business.

This is why programmers developed 'cookies'.

So what are cookies?

A 'cookie' is a small piece of information sent by a Web server to store on a Web browser so it can later be read back from that browser.

They are invisible to the actual website visitor and are often sent without their knowledge.

What are they used for?

An example is when a browser stores your passwords and user IDs. They are also used to store preferences of start pages. Both Microsoft and Netscape use cookies to create personal start pages.

There are a number of common uses for cookies.

Targeted marketing

This is probably one of the main uses of cookies. They can be used to build up a profile of where you go and what adverts you click on. This information is then used to target adverts at you. Companies also use cookies to store which adverts have been displayed so the same advert does not get displayed twice.

There are several companies providing a managed cookie/advertising service, including New York-based Doubleclick. Doubleclick specialises in providing Web-based advertising services to high-profile firms,

Brian Pennington, Marketing Director,
Brilaw International Ltd,
Brilaw House,
Parkgate Close,
Bredbury Industrial Park,
Stockport SK6 2SH,
UK
Tel: 44 (0) 161 455 7502
Fax: 44 (0) 161 455 7585
E-mail: bpennington@brilaw.com

including IBM, Microsoft and AT&T. Cookies are central to Doubleclick's ability to target advertising, regulate frequency, and provide the detailed analysis of advertising to reach the market demands.

So how does Doubleclick do it? While the site you are visiting is sending you its home page, it is also connecting to Doubleclick. Doubleclick checks its database to see if you are a part of it. If not, it passes you one of its cookies. Once this cookie is shared, Doubleclick begins to log information on what ads you are seeing as you surf the site. Each time you load a new page, the invisible 'handoff' to Doubleclick occurs; with each handoff your cookie is updated with information on the new page or new ad you are checking out. If you have already been to the site before, Doubleclick will recognise your cookie and send an ad to the page you are on that is tailored just for you, based on information previously saved in your cookie. It all takes place in less than 20 milliseconds.

At the moment, the primary service Doubleclick is selling to its customers is the ability to prevent a particular cookie holder from seeing the same ad too many times. Internet advertising suffers more than any other form from viewer boredom and annoyance at seeing the same ad banner again and again and again. By tracking how many times you have seen a particular banner, Doubleclick can regulate frequency so that you are shown an ad a specified number of times, thereby preventing banner burnout.

Unique services

All this allows some unique services. Advertisers can specify they will not pay for showing the same ad to any given user more than *x* times — perhaps three, five or ten. By comparing a carefully constructed database where domain names have been resolved into SIC codes, advertisers can target ads to users from any industry group.

Dave Merriman, senior engineer with Doubleclick, says:

> 'A year ago advertising on the Web was fairly new and people were willing to experiment. Now advertisers are demanding intelligent systems that give them more control and detailed reports. Advertisers want to target ads by many criteria, including geographic location, operating system, domain name, time of day, and frequency of exposure.'

Detailed reports generated from collected data also give Doubleclick a valuable carrot for advertisers. They can demonstrate a very predictable response rate (they call it a click rate) based on frequency of exposure. The more times a user sees an ad the less likely they are to respond. These data can be matched up against advertising rates, and companies can balance exposure with cost-effective response.

Online ordering systems

Since cookies are designed to collect information and pass it back and forth between you and a server, they can be used for lots of helpful reasons. For example, a site that lets you order goods or services could store your shopping list in your cookie file as you make your selections. If your shopping session was interrupted for some reason, you could return later and pick up where you left off.

Some sites let you personalise what they have to offer, based on your interests. Your preferences are then stored, like the shopping list above, in a cookie file. A useful example of this is www.search.com which offers links to dozens of searchable websites and search engines. Search.com lets you choose your 20 favourite search places; the next time you visit the site, you will see a new link called 'Your Page', which is just that: a personalised page that displays what you chose as your top 20 picks.

Site personalisation

Target is the key word here: the idea is that if I am interested in buying a car and you are a photographer, we can go to the same site at exactly the same time, and I will see a Rover ad and you will see a Nikon ad. Of course, Doubleclick charges advertisers for providing this profile-rich ad-viewing scheme, but companies can create their own less-sophisticated cookies.

Website tracking

All website owners want to know which parts of their sites are visited the most. Site tracking can show you what pages cause users to leave your site and visit another, as well as what pages are the most popular. This information enables changes to the site to be made constructively and provides more accurate data about the traffic the site is actually receiving. For example, you could differentiate between 50 unique people visiting your site and one person hitting the refresh button 50 times.

Information enables constructive changes

How do they work?

Cookies are small data files which are delivered to the surfer's machine by the website being visited. The website may deliver one or more cookies to the client. The surfer stores the cookie data in one or more data files on the surfer's local hard drive. In certain cases (determined by the data within the cookie itself), the client returns the cookie to the server that originally delivered it.

Cookie data are generally stored unencrypted on the user's local hard drive (although during actual communications it is stored in RAM memory). The file name for cookies is different for each operating system platform. For example, on Windows machines, the cookie information is stored in a folder called 'cookies' located under 'windows'.

The need for Web browsers and servers to exchange cookies is due to the stateless nature of HTTP. (A stateless server — for example, a World Wide Web server — considers each page request independently. A request specifies the entire document, without requiring any context or memory of previous requests. Unless a cookie is set, no information is carried across requests.) Unless something special is done, Web servers are only aware of users when a transaction sending or receiving information is in process. The moment the transaction is complete the server forgets about the user and makes no attempt to correlate subsequent transactions with previous exchanges.

There are basically two types of cookies: temporary and persistent. A temporary cookie maintains state information for shoppers. This enables a site to spread products and information over multiple pages, or to put the order entry forms on a separate page. As the user selects products they want to buy, the server indexes these selections to the session key carried as a cookie by the user's browser. Persistent cookies are those stored on the user's file system, and provide a convenient location to store user preferences that are likely to be used each time the user visits a website.

Cookies are designed to be read only by the site that originally issues them, not by other sites, and may have an expiry date. If no expiry date is specified, the cookie is deleted based on your individual browser settings.

Are they a security risk?

Internet active content is a global term for anything that makes the Internet interactive.

Cookies, Java applets, ActiveX controls, JavaScript and other forms of mobile code, which download and run immediately from Web pages on visiting users' machines, represent a massive change in the way people should view and use the Internet.

New active-content technologies can as easily be configured to execute malicious activities like erasing data stored on hard disks or surreptitiously copying and transmitting data to eavesdropping third parties as they can to help marketers' target users.

It is the prospect that a website might contain malicious code which worries Internet users the most, but simply disabling all mobile codes at the browser will serious reduce the benefits of using the Internet. For example, many e-newsletter subscription systems require active code to record the subscription request.

This is why several software companies (eg www.finjan.com) have developed active-content filters that filter all access to the Web's many and varied interactive benefits, and can spot potentially malicious code and stop the users being exposed to the code.

Without these specific security measures, programs running on your computer generally have access to all your files. With free rein, a vandal program could send e-mail in your name, erase your files, or even install a virus.

Are they an infringement of a visitor's privacy?

In the USA consumers are particularly concerned over privacy issues. 1995 research by Harris said that 82 per cent of Americans are concerned about threats to their own privacy and 78 per cent believe that consumers have lost control over how businesses use personal information. Research in the UK has shown similar results.

Improved interaction

Cookies are not so much a threat to privacy as a way of improving your interaction with the Internet. However, other examples of active code can create damage or store the e-mail address.

A consequence of someone obtaining your e-mail address is that you could become part of a direct marketing mailing list, which means your electronic mailbox could pile up with junk e-mail. Junk e-mail is more commonly referred to as spam because a lot of people do not like it.

And once someone has your e-mail address, it is easy to discover more about you. The most direct way is by querying one of the many Internet white-page databases. These freely available Web databases store millions of names, addresses, phone numbers and e-mail addresses. They mainly operate in the USA, but in the near future they will be operating in the UK. Try a quick search, using yourself as a guinea pig, on www.switchboard.com, www.iaf.net or http://people.yahoo.com.