You cannot hide when you are online

16 Jun 08

Sherman Tan

Your PC is Exposed

Many of us think that surfing the web and searching the internet for information is a solitary experience, but in fact, every time you visit a new website, that website gathers information about you. While it is mandatory for us to lock our doors at night (except for some fortunate folks who live in crime-free estate) to keep out intruders, whenever we turn on our computer and connect it to the internet, we are in fact, putting up a big sign across the internet saying “Welcome to my PC, take what you want from it!”.

Is the situation that serious or have I been exaggerating?

A couple of years ago, AOL together with the National Cyber Security Alliance found that as many as 80% of all computers surveyed contained spyware and 90% of these PC owners were not even aware that their machines were infected. As high as 20% of these PCs had active viruses running. A recent study showed that nearly 80% of our emails are spammed mails.

Some of us may feel that there is always a trade-off; getting information instantly from all over the world in exchange for some annoying spam emails, pop-up advertisements or the occasional viruses. The bad news is that these occurrences are more than just annoyances and they can do more harm to your PC and you than you may have imagined. For instance, some websites track every page you visit and the amount of time you spend on each page. It might examine your IP address and find out your geographic location and potentially your place of work. With these and other supplemented information from other databases, a surprisingly sophisticated and complete profile of who you are and your interests could be assembled.

Websites that are put up by hackers contain malicious software that exploit certain flaws in some popular browsers and attack these browsers that visit these sites. Many users are often attracted to web sites that provide free software downloads, code breaker software, video and picture downloads of popular artistes. Unknown to them, these websites download malicious codes to the user’s PC whenever a link, a pop-up box or a download button is clicked. Malicious codes can come in many forms from key logger and adware to Trojan horses that can control your entire PC and convert it to a zombie PC that forms part of an entire zombie network that launch attacks on other PCs and servers across the world.

Web security is very wide topic on its own and there are many books written on this subject matter covering the various types of threats such as adware, spyware, home-page hijackers, viruses, Trojans, bots, worms, spam, phishing and provide in-depth information on how to prevent and combat the various cyber-threats. This article however, focuses on how websites monitor your visits and how your privacy can be compromised if you do not take some precaution when surfing online.

How Websites Invade Your Privacy

Basically, three technologies are involved in tracking your website activities, namely cookies, web tracking and web bugs.

Cookies are bits of data put on a hard disk when someone visits certain websites. There are legitimate uses for cookies – for example, they make it easier for people to use websites that require a username and password. The cookie on the hard disk has

the username and password on it, so that you don't have to log in every page that requires that information. Instead, the cookie sends the information to the server on your behalf so that you can visit these web pages freely including registration and filling goods in electronic shopping cart for online purchases.

Cookies also store the name of the website that placed the cookie. Only that website can read the cookie information, so information from one website cannot be shared with information from another website. Cookie information is put into a special file on a hard disk. The location and files vary according to the type of computer and the browser. On PCs using Internet Explorer, for example, cookies are often found in the folder Documents and Settings\Owner\Cookies, with each cookie in its own text file.

Whenever you visit a website, your browser examines the URL you're visiting and looks into your cookie file. If it finds a cookie associated with that URL, it sends that cookie information to the server. The server can now use that cookie information. Cookies can contain many kinds of information, such as the last time you visited the website, your favourite pages on the website and other similar information. Cookies can be used to track you as you go through a website and to help gain statistics about which types of pages you like to visit, how long you spend on the page and related searches that you made.

While cookies can be used to track how people use a website, there are other methods that are deployed for **web tracking**, for instance; many web masters examined the web server logs in detail to identify the most popular pages on the site, the sites people have just visited, how many pages people read in a typical visit and similar information. Another method requires the use of software sniffers that examine every packet coming into or going out of a website.

Typically, a sniffer sits on the internet and analyses the traffic to the website and this is done by analyzing all the TCP/IP packets that enters and leaves the website. There are various ways for the sniffer to identify who is entering the website that it is assigned to monitor. If the website uses cookies, the sniffer will uses the cookie to identify the person. Besides, cookies, a sniffer can also make use of the Open Profiling Standard (OPS) information that is stored on the visitor's web browser. This information usually contains the browser type and version, operating system, default settings such as language, country, etc. If no cookies or OPS information is present, the sniffer will use the visitor's IP address.

The sniffer examines every packet that enters/leaves the website and record any time an action is taken, for example when a visitor requests a web page, or whenever an action is completed. All the information collected is sent to the designated database and many types of reports can be created including the average amount of time a visitor spent on a web site, the top ten web pages visited, the types of browsers and operating systems they used, the top ten unique visitors by country of origin, websites that the visitors have just visited and websites they are going to visit. Webmasters can use this tracking information to help create better sites – but they can also use it to assemble demographic information to sell to advertisers.

Finally, **web bugs** can also trace people's paths through a website. Web bugs get their name not in reference to an error in a program, but instead from the term to "bug" as in "to wiretap". Essentially, a web bug is a piece of HTML code placed on web pages or in email messages that can be used to silently gather information about people, track their internet travels, and even allow the creator of the bug to secretly read a person's email.

Originally, a web bug was a small (usually 1×1 pixel) transparent GIF (an image of the same colour of the background) that is embedded in an HTML page, usually a page on the Web or the content of an e-mail. Modern web bugs now uses the HTML iframe, style, script, input link, embed, object, and other tags to track usage.

Whenever the visitor opens the page with a graphical browser or graphical e-mail reader such as Yahoo mail, Gmail or Hotmail, the image or other information (such as a cookie) is downloaded. This download allows the website to be informed of when the HTML page has been viewed. However, when web bugs are used in e-mails, they serve difference purposes. In the case when a web bug is embedded in an email, the transparent image (unknown to the user) is requested when the user reads the email for the first time and can also be requested every time the user loads the email again. The server implementing the web bug now knows certain information about the user's computer such as its IP address. Spammers using this technique know whether the email address is an "active" one and when the email has been read.

If the recipient of the email sends the email along with it, the whole process starts all over. Now, however, when the web bug sends the contents of the email message, it contains the recipient's comments, so the email has effectively been wiretapped. This can keep continuing so that every time a new person gets the message, the wiretap continues.

Precautions You can Take

Most browsers, including Internet Explorer allow you to control whether you want to accept cookies or not. For instance, in Internet Explorer, you can go to: Internet Option\Privacy\Advanced to select whether you want to accept cookies by default or to give you the option to decide whether you want to accept the cookie from a particular website or not. However, you will have to bear with the inconvenience of each time answering a question of whether you want to accept the cookie or not. Even websites such as Google Search, Gmail, Hotmail and Yahoo mail load first party cookies to your PC.

Since surfing the web and searching on the internet leaves you vulnerable to privacy invasion by some of the web sites that you visit, you can consider surfing anonymously. Setting up a proxy server when surfing means the web site sees a proxy server instead of your PC. However, in setting up a proxy server, bear in mind that surfing on the web will become much slower and if the proxy server setting is not done correctly, you may not be able to reach your ISP to connect to the web.

Many modern e-mail readers and Web-based e-mail services such as Gmail, Yahoo mail, etc will not load images automatically when opening an HTML e-mail from an unknown sender or that is suspected to be spam mail. The user must explicitly choose to load images. Not displaying images while still using HTML is usually not enough since other techniques such as iframes can still be used to track email viewing. To totally avoid web bugs in email, the only choice is to use pure text-based email reader such as Pine or Mutt that do not interpret HTML or show images.

Finally, clear your internet cache memory including browsing history, cookies regularly and if possible, after each web surfing session.

The writer is the Principal Consultant & Director at Innovar Pte Ltd (www.innovar.com.sg).