Office of the
Privacy Commissioner
of Canada

What an IP Address Can Reveal About You

A report prepared by the Technology Analysis Branch of the Office of the Privacy Commissioner of Canada

May 2013

Table of Contents

Introduction – What we set out to explore and why

Over the past decade, the Government of Canada has tabled various iterations of so called lawful access legislation.

The latest one identified six specific elements of subscriber information which would be made available to law enforcement and national security authorities without prior judicial authorization; specifically, one's:

- name;
- address;
- telephone number;
- electronic mail address;
- Internet protocol address; and
- local service provider identifier.

(A brief description of some of these elements (i.e., IP address, e-mail address and local service provider identifier) appears in Annex A.)

> *"…. the findings lead to the conclusion that, unlike simple phone book information, the elements examined can be used to develop very detailed portraits of individuals providing insight into one's activities, tastes, leanings and lives."*

Proponents of previous attempts at such legislation have described such subscriber data as being similar to "phone book" information.[1] This document presents findings from a technical analysis conducted by the Office of the Privacy Commissioner of Canada (OPC) examining the privacy implications of subscriber information elements which are not found in a phone book: email address, mobile phone number and Internet Protocol (or IP) address.

Research associated with this analysis concluded December 19, 2012. It was performed in accordance with the Office's mandate to support, undertake and publish research into privacy issues and to promote public awareness through the preparation and dissemination of research findings for use by the general public, federal government institutions and private sector organizations.

Further, the analysis was conducted in order to provide OPC staff the ability to speak to the issues raised by previously proposed legislation, and advise Parliament accordingly on the basis of firsthand knowledge. It is not intended to be a commentary on, or reflect, current or future law enforcement practices or

procedures. It is simply intended as an example of the "state of the possible".

In general, the findings lead to the conclusion that, unlike simple phone book information, the elements examined can be used to develop very detailed portraits of individuals providing insight into one's activities, tastes, leanings and lives.

Acknowledgement

This analysis is not the first of its kind. Prior to this work by the OPC, which began as the latest incarnation of federal lawful access legislation, Bill C-30, was still on the Parliamentary agenda, a similar analysis was performed by Christopher Parsons, a PhD candidate in the Department of Political Science at the University of Victoria.

His analysis, done in the face of a former version of lawful access legislation, was posted to his blog – *Technology, Thoughts and Trinkets* – under the title *"The Anatomy of Lawful Access Phone Records"* on November 21, 2011.[2] It looked at what International Mobile Subscriber Identification and International Mobile Equipment Identification numbers could uncover about individuals.

While these data elements were proposed to be made available to authorities without prior judicial authorization in previously proposed bills, they were not among those included in the definition of basic subscriber information posed by Bill C-30.

Methodology – How we carried out our work

Our research involved carrying out the straightforward task to conduct a simple test to determine what information can be found when starting with an IP address (a similar process can be followed when starting with an e-mail address or phone number). We:

1. used the IP address of the OPC web proxy as well as the IP address of an active contributor to Wikipedia;

2. looked-up the owner of the IP address, including any registration entries, using tools such as WHOIS (an online service used for activities including querying databases that store the registered users or assignees of domain names or IP address blocks);
3. conducted geolocation and network location searches using the IP address;

and

4. used the IP address as a search term in various search engines (e.g., Google, Bing) and examined the web pages returned in the search results looking for examples of web activities (e.g., entries in web server logs, contributions to online forums).

By combining the results of all of these steps, it was possible to build a detailed profile of a person or group associated with the IP address. Some examples are outlined in the sections that follow.

Once the IP address, email address or phone number have been disclosed by the service provider or the subscriber, no special equipment or software is needed to conduct these tests. A variety of services are available on the web for obtaining information about IP addresses, email addresses, and phone numbers. There are also services that allow individuals to look up information about these items, including ownership and geolocation information. Finally, services, such as Google and Bing, can be very powerful when using these pieces of information as search terms.

What can Basic Subscriber Information elements unlock?

The following examples illustrate the types of additional information about an individual that can be discovered starting from knowledge of some element of subscriber information.

As shown, this information can reveal real world locations (in addition to civic addresses), elements of an individual's online activity and possibly lifestyle preferences.

1. Phone number and email address

A phone number (landline and/or mobile) can be used to obtain a variety of other information about an individual, such as:

- names and addresses associated with that phone number (using reverse lookup tools such as www.411.com);
- using open source searches, any public Internet activity or publicly accessible document that includes that phone number, including blog posts, discussion forums, financial or medical records[3], etc.; and

- using domain registration records, any Internet domains associated with the phone number.

Similar to a phone number, an email address can lead to a variety of information about an individual, including:

- the real name, if used in the email address or otherwise associated with the address;
- registration for services using the e-mail address. For some services (e.g., LinkedIn), the e-mail address acts as the username;
- any domains that were registered using the e-mail address;
- Internet activities or documents, including e-mails, that contain the e-mail address and that are subsequently indexed by search engines;
- friends on social network services; and
- previous employers (e.g., if the e-mail address is included in a resume posted online).

What we found …

NOTE: The results of the tests conducted during this analysis were quite revealing and had the potential to lead to the identification of an individual. In order to protect privacy, and reduce the risk of identification or misidentification of an individual, the results presented in the examples that follow were generalized to remove as much identifying information (e.g., IP addresses, website names, specific search subjects, URLs and so on) as possible.

Indeed, as a demonstration, the mobile phone number of an Office of the Privacy Commissioner of Canada staff member was used, with consent, to conduct online searches.

The results revealed:

- the individual's full, real name;
- the individual's mobile telecommunications service provider;
- two personal web sites and their domain registrations;
- an affiliation with a university;
- contributions to online discussion forums concerning Internet broadcasting, security and professional conferences; and
- participation in a local interest group on technical issues.

2. IP Address - General remarks on IP address functionality

"...even non-commercial Internet activity, such as reading documents on web pages, invariably requires the transmission of IP address information that can identify what one reads online."

Knowledge of an IP address allows a searcher to obtain other information about a network, device or service. Specifically, one can:

- determine who owns and operates the network. Searching the WHOIS database using an IP address can provide a range of information about the individual[4] (which could, in turn, reveal organizational affiliations) or organization to which the address is assigned, including a name, phone number, and civic address[5];
- perform a reverse lookup (the resolution of an IP address to its associated domain name) to obtain a computer name[6], which often contains clues to logical and physical location;
- conduct a traceroute (a computer diagnostic tool for displaying the route (path) of packets across an IP network) to find the logical path to the computer, which often contains clues to logical and physical location;
- determine the geolocation of the computer, with varying degrees of accuracy. Depending on the lookup tool used[7], this could include country, region/state, city, latitude/longitude, telephone area code and a location-specific map;
- search the Internet using the IP address or computer names. The results of these searches might reveal peer-to-peer (P2P) activities (e.g., file sharing), records in web server log files, or glimpses of the individual's web activities (e.g., Wikipedia edits). These bits of individuals' online history may reveal their political inclinations, state of health, sexuality, religious sentiments and a range of other personal characteristics, preoccupations and individual interests; and/or
- seek information on any e-mail addresses used from a particular IP address which, in turn, could be the subject of further requests for subscriber information.

According to Electronic Frontier Canada[8], even non-commercial Internet activity, such as reading documents on web pages, invariably requires the transmission of IP address information that can identify what one reads online.

What we found …

To illustrate the process, a simple test was conducted using, as a starting point, the IP address of the web proxy of the Office of the Privacy Commissioner of Canada.

A WHOIS lookup revealed that the IP address was assigned to Public Works and Government Services (PWGSC), with an address of 350 KEDC (this is the King Edward Avenue Data Centre), Ottawa, ON, K1A 0S5. The technical point of contact is listed in this entry, including full name, email address, and phone number.

Using the IP address as a search term yielded more than 240 ”hits.” The results revealed that individuals working behind the IP address had visited sites dealing with, for example:

- search engine optimization training;
- Canada's advertising and marketing community;
- web governance;
- identity management;
- privacy issues;
- legal advice related to insurance law and personal injury litigation;
- a specific religious group;
- fitness;
- online photo sharing;
- the revision history of a Wikipedia entry; and
- specific entertainers which, in turn, exposed a variety of usernames.

3. IP Address - Information about individuals

It should be noted that the above information was based on the online activity of a group of computers, not an individual work station. Having said that, the process used to derive these results applies equally well to the case of a residential subscriber. The specific information that can be retrieved however depends on how active the subscriber is online and how the websites he/she visits treat IP addresses (i.e., do they expose them to indexing by search engines).

To show what an IP address can unlock about an individual, a similar analysis was undertaken using IP addresses more representative of an individual subscriber.

What we found …

Starting with people who were active contributors to Wikipedia, we found that conducting searches using the IP address shown by this site often reveals a detailed profile of an individual's activities.

For example, the IP address of one individual Wikipedia contributor[9] revealed that the person has:

- Edited hundreds of pages on Wikipedia about television shows, both North American and international. The interest in TV shows was extensive and specific, but the details are not included here for privacy reasons;
- Edited dozens of pages on Wikipedia related to history topics;
- Participated in a discussion board about a television channel; and
- Visited a site devoted to sexual preferences following an online search for a specific type of person.

For the purposes of the research undertaken by the OPC, the above traits were gained only by looking at one IP address. These examples, however, give a glimpse into the kind of portrait that authorities could be able to paint of individuals without needing to obtain prior judicial authorization as has been proposed in previous legislation introduced at various points over the last decade.

"…the FBI was able to obtain this information using administrative subpoenas, or they may have been able to use National Security Letters, neither of which require prior independent judicial approval. Similar information could be obtained without prior independent judicial approval under previously introduced Canadian federal legislative proposals."

The Petraeus Incident – Demonstrating what Basic Subscriber Information has unlocked and led to

Another example of the information that can be determined using an IP address as the starting point for an investigation is the widely-publicized Petraeus case in the U.S.

This case started as an investigation into harassing emails but eventually resulted in the revelation of an extramarital affair by the Director of the CIA, David Petraeus, and other compromising details, which resulted in his resignation.[10]

As best as can be determined from publicly available media sources, the following appears to be the sequence of events:

- **a)** An individual received a number of "anonymous" harassing e-mails and asked the FBI to investigate. Copies of the e-mails were made available to the FBI;
- **b)** Although the messages were sent from an anonymizing service, the IP addresses from which they were sent were available in the e-mail headers;
- **c)** From knowledge of the source IP address(es), the FBI was able to identify the organization to which the IP address(es) had been allocated (typically a telecommunications service provider(s);
- **d)** Upon receipt of administrative subpoenas[11], which are issued by law enforcement authorities without judicial oversight, the telecommunications service provider(s) then provided subscriber information about the IP addresses used to access the originating e-mail account, as well as any other e-mail accounts that were accessed from the same IP address(es). It has been reported that Google gave the FBI information about every IP address used when accessing that account[12];
- **e)** The ISP associated the IP addresses with various locations, including hotels;
- **f)** Knowing the physical locations from which the e-mails were sent, the FBI was able to obtain lists of people who were at those locations when the messages were sent through the use of administrative subpoenas[13];
- **g)** One name kept appearing in guest lists during the times the messages were sent, so this individual was considered the most likely suspect; and
- **h)** It was at this point that the FBI sought and obtained a warrant to get access to the contents of the anonymous email account.

The FBI was able to obtain the following information without having to obtain a warrant:

- **a)** The IP address(es) from which the harassing e-mails were sent;
- **b)** The names of the telecommunications service providers to whom those address(es) were assigned;

c) The subscriber information associated with the e-mail account used to send the e-mails, along with information about other e-mail accounts that were accessed from the same IP address(es);

d) The organizations – in this case hotels – to whom the telecommunications service provider had assigned the IP address(es); and

e) Lists of guests who were registered at those hotels at the time the emails were sent.

According to several public sources[14], the FBI was able to obtain this information using administrative subpoenas[15], or they may have been able to use National Security Letters, neither of which require prior independent judicial approval. Similar information could be obtained without prior independent judicial approval under previously introduced Canadian federal legislative proposals.

"As information technologies become more and more common in our lives, and the more they become an extension of our very selves, the more sensitive and revealing subscriber identification information becomes."

Summary – What this all means

As demonstrated in the above case studies, knowledge of subscriber information, such as phone numbers and IP addresses, can provide a starting point to compile a picture of an individual's online activities, including:

- Online services for which an individual has registered;
- Personal interests, based on websites visited; and
- Organizational affiliations.

It can also provide a sense of where the individual has been physically (e.g., mapping IP addresses to hotel locations, as in the Petraeus case).

This information can be sensitive in nature in that it can be used to determine a person's leanings, with whom they associate, and where they travel, among other things. What's more, each of these pieces of information can be used to uncover further information about an individual.

As information technologies become more and more common in our lives, and the more they become an extension of our very selves, the more sensitive and revealing subscriber identification information becomes.
Referring to such data as being on par with what one would find in the white pages of a phone book grossly misconstrues and underestimates what can ultimately be gleaned from such information.

As such, it is truly more than just "phone book" information.

Annex A

Internet Protocol Address

An Internet Protocol (IP) address is a numerical identification and logical address that is assigned to devices participating in a computer network utilizing the Internet Protocol. Although IP addresses are stored as binary numbers, they are usually displayed in a more human-readable notation, such as 208.77.188.166. The Internet Protocol also has the task of routing data packets between networks, and IP addresses specify the locations of the source and destination nodes in the topology of the routing system.

The IP address is assigned, or leased, to an individual by an Internet service provider and is an essential element to accessing the Internet itself. IP addresses identify where data originates from and where it should be sent towards. IP addresses can either be static or dynamic. A static IP address is one that is assigned to a network-connected device that needs to have a permanently assigned address (e.g., a server, firewall or router). Alternatively, a dynamic IP address is one that is assigned to a network-connected device on a temporary basis, which is typically the case in the consumer space. It should be noted that the duration of an IP address assignment can vary from a few days to a few months, depending on a number of factors such as the size of the pool of IP addresses available to the ISP, the number of subscribers and the relative stability of the network.

Most telecommunications service providers impose limits on the amount of data a subscriber can download in a given period of time, depending on the plan that a subscriber purchases (e.g., Rogers permits 20 GB of data per month for their "Lite" Internet access package) and levy surcharges for any amounts in excess of the plan limit. In order to do this, telecommunications service providers must be able to accurately associate download traffic with a subscriber and this can be done by keeping a record of the IP address or addresses assigned to that subscriber during that time period. How long a particular Telecommunications Service Provider keeps these records depends on relevant legislative or regulatory requirements or their particular business practices.[16]

E-mail Address

An e-mail address identifies an e-mail box to which e-mail messages are delivered. The general format of an e-mail address is jsmith@example.org. It consists of two parts: the part before the @ sign is the *local-part* of the address and the part after the @ sign is a *domain name* to which the e-mail message will be sent.

The local part of the address is often the username of the recipient (jsmith). This is certainly true in the Government of Canada and in most enterprises, which typically adopt a standard convention for e-mail addresses (i.e., FirstName.LastName@).

However, the local part of the address could also be a pseudonym. Although some web-based e-mail service providers (e.g., Google's Gmail, and Microsoft's Hotmail) require that the subscriber enter a name, address and so on when creating an e-mail account, they do not necessarily verify that the information is real. The domain name portion of the address will reveal the user's organizational affiliation (e.g., @priv.gc.ca) or it will identify the e-mail service provider (e.g., @rogers.com, @gmail.com).

E-mail addresses may be tied to particular accounts, or they may be general-purpose addresses. Individuals may also have more than one e-mail address, perhaps one e-mail address for a web forum, another for purchases online, and yet another for personal correspondence. In fact, this is considered good practice from a security and privacy perspective.

Local Service Provider Identifier

The Local Service Provider Identifier, sometimes referred to as Local Service Provider Identification (LSPID), is a unique number assigned to service providers so that telecommunications switch owners and service providers can enter financial relationships for the purposes of carrying traffic. The number identifies the company that 'owns' the account associated with the traffic. This helps to identify the subscriber using a particular service (e.g., a Rogers subscriber using a Rogers mobile phone on the AT&T network) in order to ensure that the use of the service (in this case, the AT&T network) will be billed to the proper individual.

Endnotes

[1] CBC News "Opposition jumps on surveillance bill confusion" dated February 20, 2012.

[2] Parsons, Christopher, "The Anatomy of Lawful Access Phone Records", posted to the "Technology, Thoughts and Trinkets" blog on 21 November 2011. See also: "The Issues Surrounding Subscriber Information in Bill C-30", posted 28 February 2012.

[3] A search based on an element of the basic subscriber information, such as phone number or e-mail address, can return financial or medical records if those records contain the search string and have been indexed by a search engine.

[4] As more and more individuals register their own domain names (e.g., johnsmith.com), an IP address lookup in WHOIS could directly reveal the individual's name, address, etc. – without having to go through a service provider.

[5] The WHOIS system originated as a method that system administrators could use to look up information to contact other IP address or domain name administrators (almost like a "white pages"). For an example of the kind of information returned in response to a Whois query. See also http://whatismyipaddress.com.

[6] A computer name is used to help identify or locate a computer on a network. Computer names need to be unique so that computers can be accurately identified for communication purposes.

[7] There are a number of tools available for looking up IP addresses and associated information including, but not limited to, IP Lookup, IP Tools, and WHOIS.

[8] Taken from a joint Electronic Frontier Canada/Electronic Freedom Foundation submission dated 17 December 2002, in response to a Department of Justice consultation paper released 25 August 2002.

[9] There are two ways in which to make or edit contributions to Wikipedia. The first is to create an account and then log in to that account prior to making or editing a contribution. The other is to contribute anonymously, in which case Wikipedia logs the IP address of the computer used to access Wikipedia. For the purposes of this research, we started by selecting the "Recent Changes" link (on the left hand side of the main webpage) and then looked for entries that included IP addresses. By clicking on the IP address, we were able to see a list of contributions by that user. We then selected a user with a high level of activity. At the bottom of that page are tools such as WHOIS, traceroute and geolocate for deriving more information about that user.

[10] There has been extensive media coverage of the Petraeus incident including but not limited to:

a) NBC News, Engel, R., "Petraeus' biographer Paula Broadwell under FBI investigation over access to his e-mail, law enforcement officials say", dated 9 November 2012, accessed 5 December 2012.

b) WIRED Magazine (online edition), Zetter, K., "Email Location Data Led FBI to Uncover Top Spy's Affair", dated 12 November 2012, accessed 5 December 2012.

c) USA Today, Leinwand Leger, D., Alcindor, Y, "Petraeus and Broadwell used common e-mail trick", dated 13 November 2012, accessed 5 December 2012.

d) Klosowski, T., "How CIA Director David Petraeus's Emails Were Traced (And How to Protect Yourself)", dated 13 November 2012, accessed 5 December 2012.

e) American Civil Liberties Union (ACLU), Sogohian, C., "Surveillance and Security Lessons from the Petraeus Scandal", dated 13 November 2012, accessed 5 December 2012.

f) BBC, "How email trail aided Petraeus case", dated 14 November 2012, accessed 5 December

2012.

g) Sanchez, J., "Collateral damage of our surveillance state", Reuters (US Edition), dated 15 November 2012, accessed 17 December 2012.

h) Schneier, Bruce, "E-mail security in the wake of Petraeus", entry on Schneier on Security blog, dated 19 November 2012, accessed 17 December 2012.

[11] See, for example, Sanchez, J., "Collateral damage of our surveillance state", Reuters (US Edition), dated 15 November 2012, accessed 17 December 2012. See also Ambinder, M. "What the heck, FBI?", The Week, dated 13 November 2012, accessed 17 December 2012.

[12] USA Today, Leinwand Leger, D., Alcindor, Y, "Petraeus and Broadwell used common e-mail trick", dated 13 November 2012, accessed 5 December 2012.

[13] Leonard, A., "Paula Broadwell's big mistake", Salon, 16 November 2012, accessed 28 January 2013.

[14] See, for example, Sanchez, J., "Collateral damage of our surveillance state", Reuters (US Edition), dated 15 November 2012, accessed 17 December 2012. See also Ambinder, M. "What the heck, FBI?", The Week, dated 13 November 2012, accessed 17 December 2012.

[15] There are a number of types of "subpoenas" recognized by US law. The three most recognized are: an Administrative Subpoena (i.e. a subpoena from a government agency with the authority to issue such a process), a trial subpoena (sometimes referred to as an Administrative Law Judge subpoena), and a Grand Jury subpoena. An administrative subpoena is most likely what the FBI used to get certain preliminary information in the Petraeus case. See, for example, Rothacker, R. and Ingram, D., "Identity of second woman emerges in Petraeus' downfall", Reuters, 12 November 2012 (accessed 14 January 2013), which specifically quotes an unnamed US government official who stated "the FBI investigation into the emails was fairly straightforward and did not require obtaining court orders to monitor the email accounts of those involved, including the personal email account of Petraeus". Also see Leonard, A., "Paula Broadwell's big mistake", Salon, 16 November 2012, accessed 14 January 2013.

[16] Webmail providers like Google, Yahoo and Microsoft retain login records (typically for more than a year) that reveal the particular IP addresses a consumer has logged in from. See American Civil Liberties Union (ACLU), Sogohian, C., "Surveillance and Security Lessons from the Petraeus Scandal", dated 13 November 2012, accessed 5 December 2012.

You can find this publication online at:

http://www.priv.gc.ca/information/research-recherche/2013/ip_201305_e.asp