

Offensive and Defensive Techniques for Determining Web User Identity

Zachary Zebrowski

zakz@zakz.info

Approved for Public Release: 12-3046.

Distribution Unlimited

All materials are licensed under a Creative Commons "Share Alike" license

See <https://creativecommons.org/licenses/by-sa/4.0/>

Welcome

- I'm Zak Zebrowski.
- I'm a Forensic Database Engineer / "Data Miner" / and Perl guy.

Outline

- Characteristics of internet connections
- Ways of determining web user identity
- Preventing a web user identity from being determinable from an end user's perspective
- Finish

Ethics

- Reminder: Use for good not evil.
- I am not a lawyer.
- What you do is at your own risk.

Why Bother

- To determine who visits your website.
- Validate you are connecting from a known location versus an unexpected country.
- To hide your identity on the internet
- Note: Everything in this presentation is old. New technologies and techniques evolve over time which may be more (or less) intrusive.

Characteristics of Internet Connections

- What devices connect to the internet?
- When you connect to the ineternet.
- How you connect to the internet.
- Why bother : You can't escape these characteristics regardless of what you do.

What devices connect to the internet?

- Easier to say what devices don't connect at this point. Largely, things that don't have power. Most everything else, like your car, ebooks, refrigerator, dishwasher, etc, especially newer models with higher end features, can and do. (I just got notified my dishwasher got done washing dishes.)

When do you connect to the internet?

- Most people go to sleep.
- You can get an approximate location of someone, by noting times when people login or not.
- See https://census2012.sourceforge.net/images/geovideo_lowres.gif . This is an image of the (relative) number of devices that were turned on over the course of the day in 2012. Note that the number of devices go down (denoted by the light blue color) over nighttime and the number goes up (orange red) over the course of a day.

How you connect?

- If you think outside of the modern economy, not everyone has a cell phone... Or if they do, it may be slower performance or faster performance depending upon where someone lives (city versus country... and different countries as well with better economies.)
- Also, different regional differences (such as hurricanes, where infrastructure is not reliable) could redesign the average connection to the internet.

Other identifying techniques

- IP Geolocation
- What your browser exposes when you connect to the internet

IP Geolocation

- IP Geolocation is the name for the technology to go from an IP Address to a physical location. Companies, open source projects, provide this information in various formats to download and use within your applications (internal or external). Further companies can provide API's as well to various websites through techniques.
- Note that companies often provide additional attributes as well as IP geolocation, such as domain name, isp, org, and other calculated attributes. These are a by-product of the analysis they perform to refine calculations.
- As location information is valuable, other projects attempt to refine location as well, through bluetooth or wifi devices, cell phone coverage, or other attributes and then allows the browser to perform an approximate location lookup based upon the devices found and their known locations. See <https://beacondb.net> as an example data source. Applications (such as firefox, or geoclue), can make requests to this api for an approximate location response.

What are unique features about IP Geolocation Databases

- An IP Geolocation database is often offline. As such, I can tell from logs, approximately where you are located, without performing online queries.
- Though IP Geolocation, combined with other resources, I can generally identify your ISP provider (verizon or similar) and then target you based upon that information. I could also potentially identify fraud, if your profile does not match where you are expected to login from.

How does one create a IP Geolocation database?

- (Caveat, only open source IP Geolocation techniques are discussed.)
- Ask the user. Hostip.info was (past tense) a source that simply asked the user, for a given IP Address, what country they were located in.
- Query registerys. Registrys assign IP Addresses to countries, and provide files of IP addresses to ISP. Generally, these ISPs are known to operate within a given country. Software77.net (a now defunct website) did this, as well as ipasn.com , which provides ip to country information.
- RFC 8805 (and supplement 9632) specifies an open way for ISPs to designate locations for ip address ranges. The supplement specifiys how to identify the locations of the 8805 files for a given ISP.

Differences between open and commercial sources

- Open sources have known algorithms and reues.
- Commercial sources: Claims of higher accuracy, and dedicated staff to further look into the data.

IP Geolocation Caveats

- You can get only to the most visible IP Address.
- Generally IP Geolocation files are delayed, and then need to be ingested into your environment. This may take time.
- There is generally a 1-2% drift per month for location information over time.

IPv6

- IPv6 addresses are used in various countries, and may not be readily resolvable depending upon how ip addresses are resolved.

Lab 1