How to hack the identity of a Tor user

Filip Pynckels
D rector-genera Ict at FOD B nnen andse Zaken

Pub shed on L nkedIn
January 17 2016

The goal of this article is not to be a 100% correct tutorial that will permit you to identify terrorists, paedophiles or drug addicts on the dark web, nor to encourage you to do so. I should warn you that applying some of the techniques in this article is even illegal in certain countries. So you should see the below information as a warning that absolute anonymity does not exist, not even when you use anonymity networks like I2P, FreeNet or Tor. And what could be a better warning than explaining to you how 'bad guys' can find your identity by combining different hacking techniques.

Links to Dark Web sites and their IP addresses will not be mentioned, to stay true to the goal of this article. If and where possible, imaginary names will be used for the same reason.

How many internets are there?

The Surface Web, the Deep Web and the Dark Web

Most of us are not aware that there are lots and lots of web servers out there that you can't find using Google, or by addressing them with an url ending on an official suffix like .com, .be or .gov

Rough estimates say that only 4% to 10% of all web servers, web services and web information are visible for the average internet user by means of for instance Google or a direct URL. All these servers together are called the *Surface Web*.

There is a lot of information out there that is hidden in databases, protected by a web services layer, behind local search forms, and so forth. The information in question is neither illegal nor immoral, but simply not directly addressable. All this information together is called the *Deep Web*.

And then, there is the back ally of the internet, where terrorists, paedophiles and drug dealers dwell, where you can find the WikiLeaks and Edward Snowden publications, and where the members of Anonymous and Lizard Squad exchange information. The list of activities on this part of the internet is very long. Selling drugs, illegal arms, stolen identities and duplicate credit cards are only the tip of the iceberg. This part of the internet is called the *Dark Web*. Some call it the dump of the internet and others call it the ultimate privacy tool. Probably, the truth lies in between, but nobody can deny that almost all activity on the Dark Web is immoral and most of the time illegal.

Rough estimates say that there are about 4000 or something sites on the Dark Web, that change their names frequently. But in reality, numbers are hard to estimate due to the obscure nature of the Dark Web.

Anonymity networks

Connecting your computer to the Dark Web, whether it is to satisfy your curiosity, to try to reach a hacking collective or to hire a hitman is something you don't want your mother in law or your government to know (especially if the target of the hitman is your mother in law). So you will need a method to connect to the information on the Dark Web without someone being able to find out it's you, nor what information you connect to or send. And that is exactly the role of *anonymity networks*.

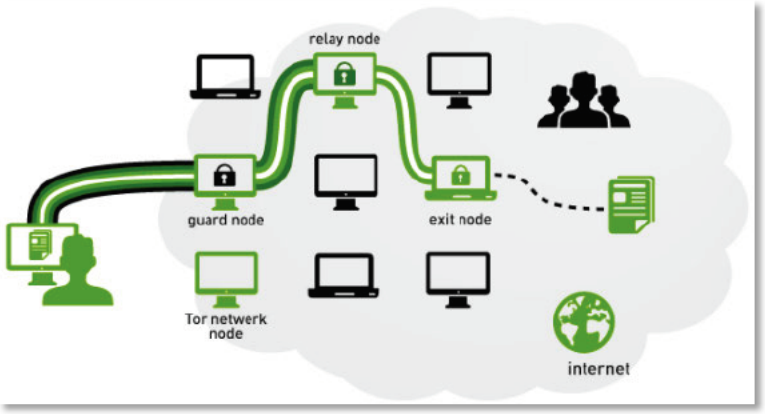
In theory, when you connect to an anonymity network, your provider doesn't see what you are connecting to, the anonymity network servers don't know the combination of who you are and what you are connecting to, and the server you're connecting to doesn't

know who you are. Although, that's the theory. The goal of this text is to make you aware that there is a major difference between theory and practice.

There are a number of well known anonymity networks like I2P and FreeNet, but the best known is the *Tor network*. So I opted to focus on the Tor network for the rest of this text. It must be clear that similar methods can be easily applied to the other networks.

The Tor anonymity network

Now we know what Tor is, we must find out how it's functioning. We already know that, on the one side of the Tor network, we are sitting with our pc in our living room, in an internet café or in the library. On the other side of the Tor network, connected to the internet, are the servers with the information we want to reach. And Tor is in between. So, how do they do it?



We connect our pc to a *guard node*, using the Tor protocol. That's a node that is chosen at the beginning of the Tor session, and that stays stable during the entire session. Since the guard node is a server amongst more than 7000 servers worldwide (maintained by volunteers, open source organisations, governments, etc.), and the guard node has nothing to do with the Pirate Bay or other sites of the internet, it seems to your provider like you are connecting to a completely harmless site. Since, each session again, you can choose a different guard node, it seems like nothing fishy is going on. The list of all available guard nodes is maintained by *Tor network nodes*. The guard node only sees triple encrypted information coming in, and double encrypted information going out. Hence the alternative name for Tor: the Onion network, since information is encapsulated within different layers of encryption.

When setting up a connection, a fixed guard node is chosen, and a first *relay node* and *exit node* are chosen at random. During the connection, the relay node alternates, and so can the exit node. This means that, even if the relay node and/or the exit node are maintained by a government or intelligence service, they only see part of the traffic. The relay node sees double encrypted traffic coming in and single encrypted traffic going out. The exit node sees single encrypted traffic coming in and unencrypted traffic going out during a short period of time. After this time interval, the exit node is switched out in favour of an other exit node. The lists with relay nodes and exit nodes are also maintained by Tor network nodes.

At that moment, your computer is connected to a server with all data traffic transiting by

the guard node, a relay node and an exit node. The server that you connect to only sees data traffic coming from the exit node, and doesn't (in theory) know who you are, or where your computer is connected to the internet.

If you like to have a secure connection between the exit node and the server you're connecting to, you should use a *VPN through the Tor network*. In that way, you encapsulate another layer of encryption within the Tor encryption layers.

So, are we anonymous and cool? Well, cool maybe, but certainly not entirely anonymous...

Hacking the identity of a Tor user

Now that we now where the information on internet is to be found, and how the Tor anonymity network functions, we can start finding (purely hypothetical of course) people looking for illegal arms and living in the environment of Falador (which is my favourite hypothetical town, and home of the equally hypothetical White Knights).

A man with a plan

And how exactly do we start? It's best to first list up the roadmap that we will be following to catch the people we are looking for. We will use a number of techniques like setting up honeypots, using browser functionalities, data mining to collect electronic fingerprints, social engineering and/or phishing, doxing, and of course plain old hacking.

The roadmap we will follow is the following:

- Set up one or more honeypots
- Get electronic fingerprints of the user computer and user browser
- Find the real IP address of the user computer
- Filter the gathered IP addresses
- Find more info on the infrastructure and the person behind the IP addresses

Setting up one or more honeypots

What is a honeypot? A honeypot is an infrastructure that attracts those elements that we want to investigate. Those elements can be viruses, criminals, good and devout people that we want to hack, etc. In our case, we want to attract people who are just that little bit less devout and want to buy illegal arms. And what better way than to set up a flourishing illegal website that offers illegal arms for Bitcoins (see https://en.wikipedia.org/wiki/Bitcoin for more information). Or at least, a make believe website. Even better, set up as much websites as possible, in order to augment the chance that the people we look for will come to our sites instead of to sites that we don't control.

As already mentioned, the Dark Web offers lots of services that can be used in a legal or less legal way. Setting up virtual servers and websites are amongst them. And so our

journey begins. First we find a number of websites that already sell illegal arms. We can copy/paste information and images from those websites, since copyrights don't exist on the Dark Web. We can study the way the competitions websites are structured, do a SWOT analysis of these sites, and make a design for the best ever Dark Web e-commerce sites selling illegal arms.

Then we find some free hosting websites on the Dark Web with a (no longer existing) Dark Web address such as http://nr6juudpp4as4gjg.onion/ We create our honeypot websites and publish them on our free hosting sites, taking care that the websites are using the Tor protocol as a default. We have chosen hosting sites that permit us to do host side programming and also permit us to store data in a virtual database.

There is even a possibility to use legitimate open source software such as the *Social-Engineer Toolkit (SET)* to help in protecting the created websites by analysing them.

An alternative is to install real servers, and to branch them to an anonymity network by VPN. However, in doing so, we are looking at a large extra workload to keep our servers anonymous, safe and hacker-free.

Of course, what goes for the Surface Web goes even more for the Dark Web: if people don't know that our sites are available, they won't visit them. And so we start advertising our sites on the Dark Web. We publish our sites addresses on hidden wiki's, we start talking about our sites on anonymous blogs and chat boards like 4chat and 8chat (again: old names that don't exist anymore), etc. If we do a good job, our visitor numbers will flourish, and we're in business. Our honeypot is waiting for the right 'ants'. Here too, SET can be used to do some anonymous bulk mailing to make people aware of the existence of our sites. Of course, we only send this bulk mail to people living in Falador, since they are our target audience.

In order to get positive comments by different users on our offered service (shipping arms to your front door within the next 72 hours), we can even channel the users (after analysing their information) to 'reliable' sites that offer the same service. For those users, it will seem like they are using our site, and that our site is reliable. And hopefully, they will spread the word…

Get electronic fingerprints of the user computer and user browser

And then, we wait... until our first victims come walking in, and we can start playing a game with them.

The game goes as follows: we know that our 'customers' will not buy from our sites during their first visit. They are aware of the dangers of the Dark Web, and they will take their time to explore. Since they use Tor, they will come in different times using different exit node addresses. And still, we have to know it's them.

So how can we find out that connections from different exit nodes are, with high probability, sessions from the same person? Well, we use electronic fingerprinting. We find whatever information we can from the computer of the visitor, and we combine it into what is called an electronic fingerprint. Of course, most browsers block for instance the computer name, but almost none block things like: mouse speed, screen dimensions, cpu speed, etc. We can find all this information by means of some JavaScripts on our web pages.

Further information can be found at
http://jcarlosnorte.com/security/2016/03/06/advanced-tor-browser-fingerprinting.html
and at
https://en.wikipedia.org/wiki/Canvas_fingerprinting
A simple example of a fingerprinting java script can be tested at
http://jcarlosnorte.com/assets/ubercookie
A more advanced example of what information your browser returns on demand can be found at
https://firstpartysimulator.org.

The major challenge at this point is that we want the user to start browsing through the Tor protocol, but that we want to lure him into using the http or https protocol, with a standard browser that is configured to use java script. In fact, this is easier than it seems. All we have to do is get the user to a home page with an url extension .onion, and afterward redirect the user to a "very interesting page" using the browser extension .html. From that page on, we can lure him to an even more interesting page that "needs JavaScript" to offer him the potential to examine the general technical specification of the goods he wants to procure. And why don't we, once he needs detailed technical specifications, show him a text that his browser doesn't support cookies and that he needs to use Safari, Chrome or what have you. Experience shows that more than half the users go along with such a scenario.

Once we have the user on a normal web browser, with cookies and java script enabled, we can fingerprint him even more. A normal browser gives you the type of browser, the IP address of the user, the computer name, and much, much more.

When we manage to achieve the above, we have a collection of computer characteristics and personal information about the user

Find the real IP address of the user computer

The next step is bringing the user back, by telling him at the end of the fingerprinting that there is a problem with his IP address, or his user ID, or his mail address, or so on… and

that he has to reconnect. With high probability, the Tor user will change exit node and come to the site again, which permits us to collect verification info.

Once we have this fingerprint verified, we can ask the user to make a login ID, being his mail address and a password of his choosing. Afterward we send him a confirmation mail, and ask him to click on a reply url. His mailing software is most probably configured to use a normal browser, and when he clicks on the link in the mail, he is using his normal browser to access our site, and giving us a second opportunity to collect more personal data through his default browser.

When receiving the connection of the user, we show him a web page that states that his user ID can not be verified, but that he can reply on the received mail with the phrase: "PLEASE ACTIVATE MY USER ID" (or whatever other text you think he will believe) in the subject. Again, with high probability, the user will reply on our mail, and in the mail meta data, we find a wealth of information on the user, on his mail provider, etc.

All the received info together will permit us to collect a number of different user IDs used by the user, his IP address, the name of his computer, characteristics of his computer, etc.

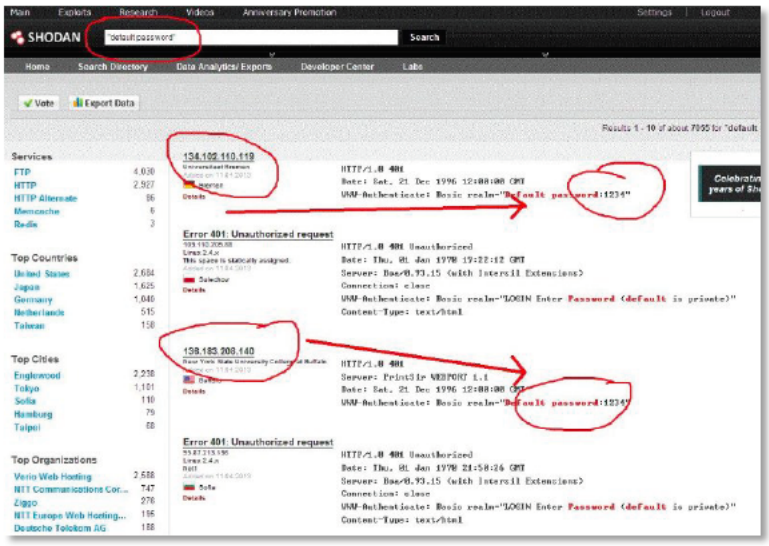
Filter the gathered IP addresses

Using one of the many legal internet services, we can check if the user's IP address originates in Falador. Some examples of manual websites to find the location of an IP address are https://www.iplocation.net and http://www.ipfingerprints.com. But we will use a programmable service of course.

Based on the information we get, we can decide to store the user IP addresses within the area we are interested in, and skip the others. We also store the other information we found on the specified user.

Find more info on the infrastructure and the person behind the IP addresses

Using programmable web services comparable to https://www.shodan.io or http://ip-report.it, we can find a plethora of things on the infrastructure on to which the user computer (IP address) is connected. Besides that, there are all kinds of information (possible attack vectors) that can be gathered using port scanners, analysing the reply on a ping (see for instance http://www.kellyodonnell.com/content/determining-os-type-ping), and so on.



Here we come within the realm of general hacking techniques that need no further explanation, since our goal is not to provide you with a manual. But it's clear that, once the computer of the person who is accessing our honeypot is hacked, we are in possession of all the info we need.

Alternatively, if we don't want to hack, we can use further social engineering techniques to get more information on the person whose browser and computer fingerprint we have collected. A more common way to do this is doxing (which, by the way, is a completely legal technique) by using the information that we already have (such as different user IDs, IP address, …). For more information on doxing, you can look at
https://www.youtube.com/watch?v=5c5Lm0H49DY
or at
http://theblackcathacker.tumblr.com/post/48768791317/methods-of-doxing.

Conclusion

The goal of this text was to give you an idea how easy it is for people with no hacking knowledge at all, and limited IT knowledge to get around the security of an anonymity network like Tor. They can use simple techniques and tools that can be found all over internet, to find your real identity. Now is that frightening or what?