Dưới đại dương Internet – Web chìm

1. Mở đầu:

Internet giống như một đại dương lớn. Đại dương được lấp đầy bởi những châu lục rộng lớn và những hòn đảo mà con người ghé thăm. Cái châu lục rộng lớn sẽ là Google, là hòn đảo sẽ là những site mới cho báo địa phương. Người ghé thăm hằng ngày những châu lục và hòn đảo sử dụng trình duyệt web, hoạt động như một chiếc thuyền đi tới các địa điểm trong internet. Mặc dù trên thực tế những châu lục và hòn đảo này chỉ chiếm 4% của Internet, phần còn lại của Internet được tạo bởi Web chìm, thứ nằm dưới đại dương. Deepweb được sử dụng với cả mục đích tốt và xấu, trong khi có thể một số người cho là nó được sử dụng với những mục đích phi pháp. Việc sủ dụng Internet tiếp tục phát triển và Web chìm là một phần lớn trong đó

1.1. Cách sử dụng Interrnet:

Mọi người trên khắp thế giới sử dụng internet hằng ngày. Hiện này có trên 3 triệu người sử dụng Internet, hơn 1 triệu websites và 3.5 triệu lượt tìm kiếm Google một ngày, có 500 nghìn tweets gửi trong một ngày. Những con số này tăng lên đáng kể trong 10 năm qua và sẽ tiếp tục tăng lên khi Internet phát triển và việc sử dụng nó mở rộng ra. Số lượng lớn người sử dụng Internet tương tự như số lượng lớn cần sử dụng nó. Theo như Top10Base, top 10 tác vụ cho mục đích sử dụng Internet là:

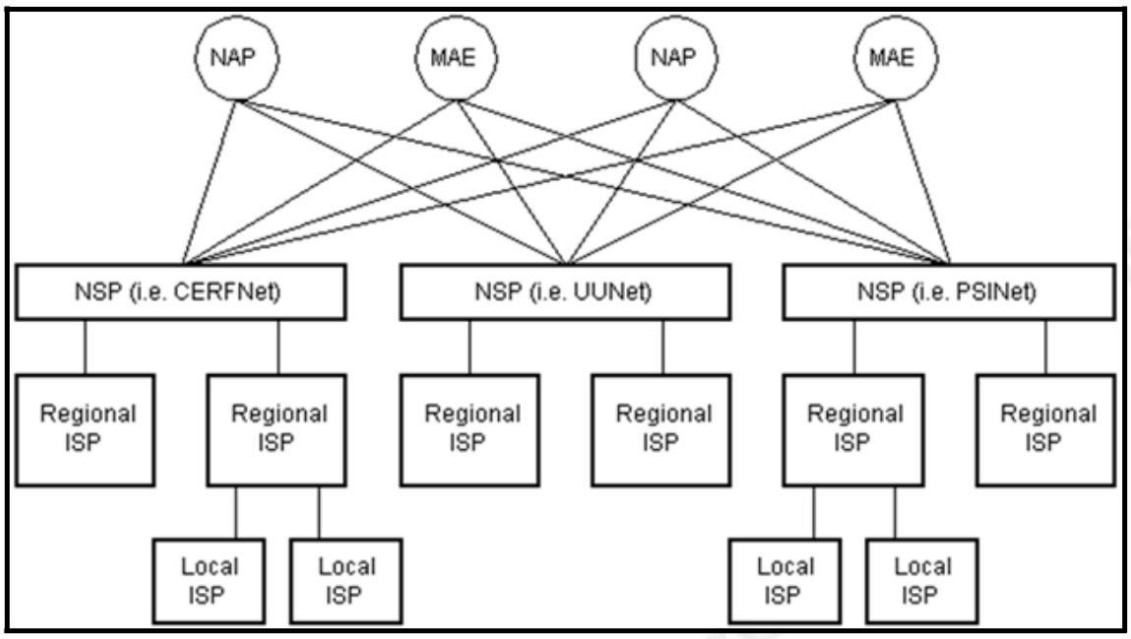
1. Email
2. Nhạc và phim ảnh
3. Tìm kiếm
4. Mua vé
5. Mua sắm

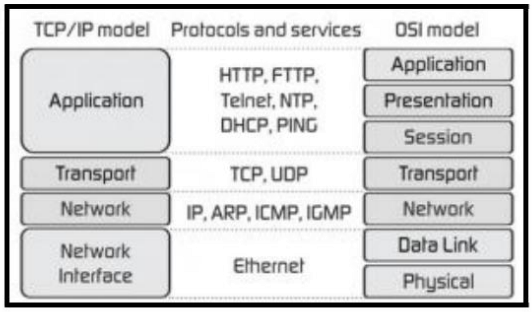
Truyền thông xã hội ở ngoài top 10. Ngoài danh sách tác vụ liệt ở trên, có bao nhiêu người trong số họ tương tác online? Câu trả lời là gần như tất cả trong số họ. Bởi bạn có thể sử dụng những tác vụ này thông qua trình duyệt web, chẳng hạn như Google Chrome, hoặc bạn có một vài mảnh phần mềm để sử dụng tác vụ như iTunes để tải nhạc và phim. Do việc sử dụng internet tiếp tục thay đổi và phát triển, thiết kế Internet cũng sẽ được thử nghiệm

1.2. Thiết kế Internet:

Internet được thiết kế độc đáo như một mạng kiến trúc mở, cho phép giao tiếp và cộng tác. Trích từ "Sơ lược lịch sử Internet":
"Trong mạng kiến trúc mở, những mạng cá nhân có thể được thiết kế tách biệt và phát triển, mỗi thứ có thể có giao diện đồng nhất của riêng nó mà có thể cung cấp cho người dùng và/hoặc những nhà cung cấp khác, trong đó có những nhà cung cấp Internet. Mỗi mạng có thể được thiết kế phù hợp với môi trường đặc trưng và yêu cầu của người dùng về mạng đó. Không có giới hạn chung cho một dạng mạng nào có thể bao gồm hoặc trong phạm vi địa lý của nó, dù những lý do thực tế nào đó sẽ điều khiển những gì có lý để cung cấp"
Bạn có thể thấy, Internet được định nghĩa là không gian mở và miễn phí dựa trên khái niệm của nó. Nó được thiết kế một cách mà những hệ thống phức tạp, cho dù khác biệt như thế nào, vẫn có thể giao tiếp tất cả được với nhau. Ngày nay, kiến trúc này vẫn đứng vững, và nó đảm bảo rằng các gói thông tin thông tin của bạn được truyền đến địa điểm của chúng.

Sương sống của internet được cấu thành bởi nhiều mạng nối liền với nhau. Những mạng lưới này được gọi là Mạng Lưới Nhà Cung Cấp Dịch Vụ (NSP). Mục đích của các NSP là truyền tải các gói thông tin thông tin đi tới đi lui. Mỗi NSP có khả năng kiết nối đến Mạng Lưới Điểm Truy Cập phức hợp (NAP), cho phép sự chuyển động của mạng đi lại giữa các NSP. Các NSP cũng kết nối với Vùng Trung Tâm Trao Đổi (MAE), đáp ứng tất cả các mục đích như một NAP. Thuật ngữ nữa được dùng sử cho các NAP và MAE là Điểm Trao Đổi Internet (IXP). Lợi ích của IXP là dễ dàng làm cho một nhóm Nhà Cung Cấp Dịch Vụ Internet (ISP) kết nối với nhau. Các ISP lấy được băng thông mạng từ các NSP. Sau đó, băng thông này được truyền đến khác hàng của họ để kết nối Internet





Hình 1 ("AS Computing – Unit 2 The Internet") – Sơ đồ thiết kế Internet

Bạn có thể thấy, thiết kế của Internet vẫn rất giống với lịch sử, bắt đầu từ cuối những năm 1960

1.3. Lịch sử của Internet

Khái niệm về Internet được sáng chế bởi Larry G. Roberts cuối những năm 1960 và được gọi là ARPANET, viết tắt của Mạng Lưới Cơ Quan Nghiên Cứu Đề Án Cao Cấp. Đầu những năm 1980, ARPANET bắt đầu sử dụng mô hình Bộ Giao Thức Internet (TCP/IP) được phát triển bởi Robert Kahn và Vinton Cerf. Mô hình này định nghĩa những chuẩn mực để dữ liệu phải được truyền giữa những mạng lưới khác nhau. Mô như thế này vẫn được sử dụng cho tới ngày nay, và được biểu thị trong hình 2 dưới đây:

Hình 2 ("Advantages and Disadvantages Of Tcp/Ip model") – The TCP/IP model

Đến năm 1990, một sự kiện then chốt nữa diễn ra trong lịch sử Internet, đó là phát minh ra World Wide Web (WWW) bởi Tim Berners-Lee. Đó là những gì mà Internet được công nhận ngày nay và giúp tăng tính phổ biến của Internet tới công chúng

1.4. Web Nổi vs Web chìm:

Internet gồm có hai mảng. Hai mảng đó là Web nổi và Web chìm. Web nổi là khu vực Internet mà tất cả những người thường có thể ghé thăm, giống như ghé thăm Facebook, Google, Amazon hay YouTube. Những khu vực này có thể được truy cập bằng mảng phần mềm thông thường, như trình duyệt web. Khu vực khác của Internet gọi là Web chìm. Web chìm được tạo bởi Web đen, cơ sở dữ liệu Web chìm và nhiều thứ khác nữa. Bạn phải cần đến phần mềm chuyên dụng hoặc

truy cập để tương tác với Web chìm. Sự phân biệt giữa hai khu vực này của Internet là rất quan trọng

1.4.1. Web nổi:

Web nổi là khu vực của Internet được biểu thị bằng những công cụ tìm kiếm, như Google. Tên gọi khác của khu vực này của Internet là Web hữu hình (visible web), Lightnet, Indexed Web, Clearnet hoặc Indexable Web. Đúng theo viễn cảnh, hiện nay đã có hơn 4 triệu web biểu thị.

Hãy nhìn vào cách mà công cụ tìm kiếm như Google biểu thị các trang web, Chúng sử dụng các mảng phần mêm gọi là web crawlers, mục đích chính của chúng là khám phá ra các trang web trong internet. Bạn sẽ biết rằng Google web crawlers là lui tới site của bạn để thấy "Googlebot" trong đại diện xâu ký tự của người dùng. Tất nhiên, đại diện xâu ký tự của người dùng có thể bị lừa bởi một kẻ tấn công. Một web crawler ghé thăm một trang web, nó sẽ tìm kiếm bất kỳ đường link nào của trang đó và ghé thăm những trang kia. Trong lúc ghé thăm các tran kia, dữ liệu được thu thập và gửi tới Google. Google có phần mềm để xác định những site nào được lui vào, tần xuất lui và số lượng trang được nhận từ các site. Nó sẽ đưa ra thêm sự chú ý cho site mới, vừa thay đổi hay không còn tồn tại nữa. Những trang đó được biểu thị một cách chính thức bởi Google để những site đó nhận được kiểu cách hiệu quả và chính xác khi cần. Sự biểu thị của Google bao gồm thông tin về các từ khóa khác nhau và nơi chúng nằm. Khi bạn tìm kiếm một thứ gì đó, Google tìm những thuật ngữ bạn tìm được biểu thị khớp với trang web đó. Đó là khu vực của Internet, cho dù Google không thể tiếp cận được thông qua tìm kiếm khi nó không được biểu thị

1.4.2. Web chìm:

Web chìm là khu vực của Internet không được biểu thị bằng công cụ tìm kiếm và không kết nối với tất cả các trang thuộc Web nổi. Những tên gọi khác cho khu vực này là: Deep Net, Hidden Web (Web ẩn), và Invisible Web (Web vô hình). Phần này chiếm 95% của Internet, tức là lớn hơn rõ ràng đáng kể so với Web nổi.
Có rất nhiều lý do mà trang web không thể bị lui tới. Trang web có thể có mật khẩu bảo vệ, sẽ ngăn chặn web crawler truy cập nó. Một cách khác nữa là trang web có thể chỉ được cho phép truy cập một

số lần nhất định, rồi trở nên vô hiệu. Nếu ngưỡng cửa đó được tiếp nhận trước khi crawler tiếp cận được trang web, nó sau đó không thể bị lui tới được. Một cách khác là trang web không thể bị lui tới nếu file robots.txt của site đó nói rõ là không được lui tới. File robots.txt nằm ở trong gốc của website và sẽ để các web crawler biết điểm đến không cho phép được lui tới trong site của nó và nguyên tắc của đại diện người dùng áp dụng. Cách cuối cùng là sẽ tạo cho web đó không thể lui tới được, nếu trang chỉ đơn thuần ẩn hoặc không liên kết với bất kỳ một trang khác trong website. Ai đó sẽ cần đến kiến thức trước đó về con đường để truy cập nó để vào một trang "ẩn". Người sử dụng Internet bình thường sẽ không sử dụng Web chìm, vì việc sử dụng của nó sẽ bị cho là đáng ngờ.

2. Web Chìm:

Web Chìm là một khu vực phức tạp và bí ẩn của internet. Có rất nhiều lý do để nội dung của nó được truy cập hoặc sử dụng cho mục đích hợp pháp hoặc phi pháp. Có rất nhiều nội dung sẵn có trên Web chìm, ví dụ như Web đen, các dịch vụ ẩn và cơ sở dữ liệu Web chìm. Phần mềm đặc biệt như: Tor, được yêu cầu để truy cập Web chìm. Chi tiết của tất cả các mặt của Web chìm gồm ở các mục dưới đây:

2.1. Tại sao phải để xuống dưới nước?

Việc sử dụng Web chìm được chia ra thành hai loại: hoạt động hợp pháp và hoạt động phi pháp. Ví dụ của hoạt động phi pháp là bán thẻ tín dụng được lấy trộm. Ví dụ của hoạt động hợp pháp là sử dụng máy Wayback để xem các phiên bản trước của một trang web. Cho dù việc sử dụng hợp pháp hay phi pháp, Hành động truy cập Web chìm là hành động mang tính quốc tế.

2.1.1. Hoạt động hợp pháp:

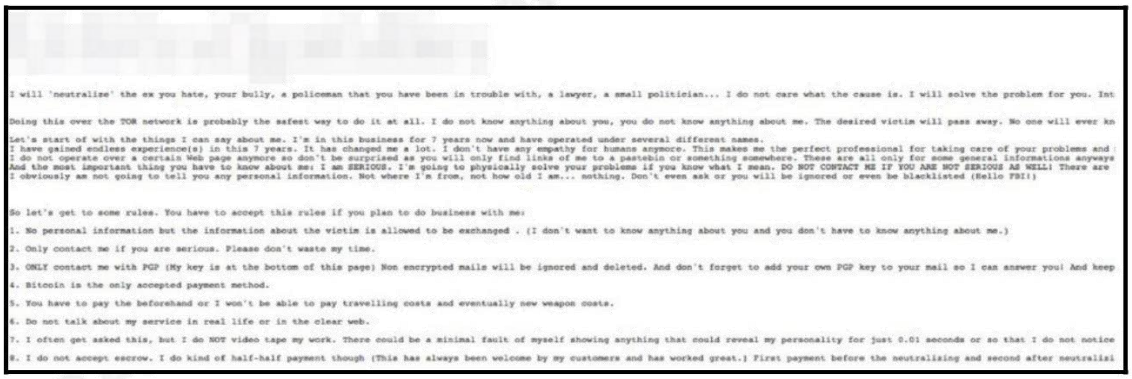
Tin nó hay không thì cũng có rất nhiều hoạt động hợp pháp ở dưới Web chìm. Web chìm có thể là một nguồn sử dụng hữu ích vì thông tin thừa thãi. Ví dụ, có rất nhiều công cụ tìm kiếm cho phép bạn tìm kiếm cơ sở dữ liệu của thế giới không được biểu thị bởi Google hay Bing. Những cơ sở dữ liệu này bao gồm thư viện học thuật ảo hoặc phiên bản cũ của trang web.

Có một vài tác vụ hợp pháp hoàn toàn để sử dụng dưới Web chìm, và bạn có thể không nhận ra dữ liệu đã được truy cập thật sự đang trú ngụ ở đó. Khi một ai đó kiểm tra khái quát về một cá nhân, nó tìm ra một vài cơ sở dữ liệu trong Internet có thông tin. Thông tin

này được tìm kiếm thật sự trong Web chìm. Việc sử dụng khác cho Web chìm là nếu một người muốn được nhận nuôi muốn cố gắng tìm kiếm cha mẹ ruột của họ. Cơ sở dữ liệu lưu giữ thông tin về việc nhận nuôi này vào Web chìm. Bạn cũng có thể sử dụng Web chìm để nghiên cứu về cự chiến binh hay tìm kiếm lịch sử phả hệ của bạn. Nghiên cứu hợp pháp cũng có thể được quản lý bởi Web chìm cho vài trường hợp. Nếu bạn là sinh viên, bạn cũng có thể sử dụng Web chìm. Một vài cơ sở dữ liệu học thuật bạn có thể tìm kiếm qua những chủ đề, chẳng hạn như các báo khoa học. Dù có những hoạt động hợp pháp dưới Web chìm, vẫn có những hoạt động phi pháp dưới đó.

2.1.2. Hoạt động phi pháp:

Không có gì ngạc nhiên, khi có hoạt động phi pháp xảy ra rõ ràng rành dưới Web chìm. Web chìm là nơi bạn có thể giấu tên để đến. Theo đúng nghĩa, đây là một địa điểm phổ biến cho tội phạm mua và bán thông tin. Một vài thông tin có thể được bán dưới Web chìm: mã số an ninh xã hội, bệnh án, số thẻ tín dụng và những thông tin nhận diện cá nhân khác. Bạn có thể mua đủ thông tin về một ai đó để trộm sự nhận diện của họ một cách dễ dàng. Web chìm có thể được sủ dụng để mua ma túy, chiếu phim khiêu dâm trẻ em, buôn bán vũ khí và thuê sát thủ. Hình 3 dưới đây cho thấy ví dụ về một bài đăng của sát thủ dưới Web chìm



I will 'neutralize' the ex you hate, your bully, a policeman that you have been in trouble with, a lawyer, a small politician... I do not care what the cause is. I will solve the problem for you. Int

Doing this over the TOR network is probably the safest way to do it at all. I do not know anything about you, you do not know anything about me. The desired victim will pass away. No one will ever kn

Let's start of with the things I can say about me. I'm in this business for 7 years now and have operated under several different names.
I have gained endless experience(s) in this 7 years. It has changed me a lot. I don't have any empathy for humans anymore. This makes me the perfect professional for taking care of your problems and
I do not operate over a certain Web page anymore so don't be surprised as you will only find links of me to a pastebin or something somewhere. These are all only for some general informations anyways
And the most important thing you have to know about me: I am SERIOUS. I'm going to physically solve your problems if you know what I mean. DO NOT CONTACT ME IF YOU ARE NOT SERIOUS AS WELL! There are
I obviously am not going to tell you any personal information. Not where I'm from, not how old I am... nothing. Don't even ask or you will be ignored or even be blacklisted (Hello FBI!)

So let's get to some rules. You have to accept this rules if you plan to do business with me:

1. No personal information but the information about the victim is allowed to be exchanged . (I don't want to know anything about you and you don't have to know anything about me.)

2. Only contact me if you are serious. Please don't waste my time.

3. ONLY contact me with PGP (My key is at the bottom of this page) Non encrypted mails will be ignored and deleted. And don't forget to add your own PGP key to your mail so I can answer you! And keep

4. Bitcoin is the only accepted payment method.

5. You have to pay the beforehand or I won't be able to pay travelling costs and eventually new weapon costs.

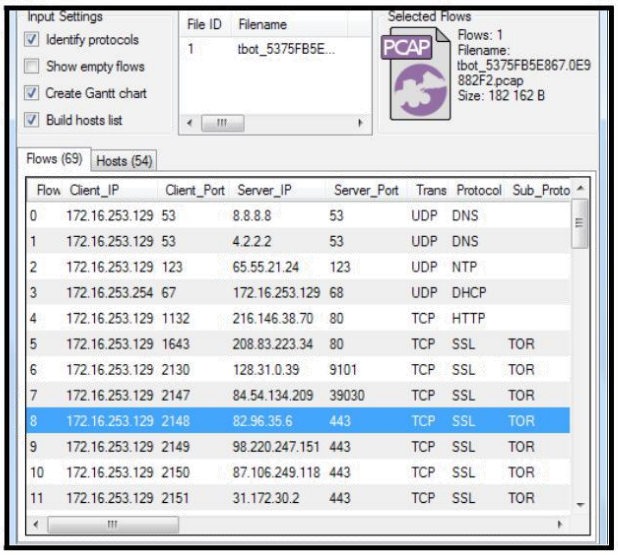
6. Do not talk about my service in real life or in the clear web.

7. I often get asked this, but I do NOT video tape my work. There could be a minimal fault of myself showing anything that could reveal my personality for just 0.01 seconds or so that I do not notice

8. I do not accept escrow. I do kind of half-half payment though (This has always been welcome by my customers and has worked great.) First payment before the neutralising and second after neutralizi

Hình 3 – Ví dụ về bài đăng của sát thủ

Như đã miêu tả ở đây, rõ ràng là có thừa những hoạt động phi pháp, nhưng nhưng làm cách nào để tìm và nhận ra hành động này trong mạng lưới của nó?
Cách phổ biến nhất là giao tiếp với Web chìm bằng Tor. Vì thế, bạn sẽ muốn có khả năng tìm ra giao dịch của Tor trong mạng lưới của bạn. Khi phân tích sự giao dịch của Tor trong gói thông tin thông tin tóm gọn, nó sẽ trông như giao dịch HTTPS thông thường. Tuy nhiên, nếu bạn để ý đến những chứng thư được sử dụng trong giao dịch Tor, bạn sẽ thấy người phát hành và chủ ID được sử dụng ngẫu nhiên để đặt tên miền, chỉ rõ chúng là những chứng thư đáng ngờ.



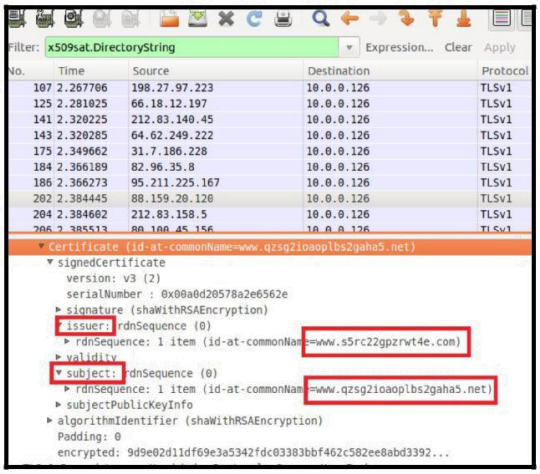
Flow	Client_IP	Client_Port	Server_IP	Server_Port	Trans	Protocol	Sub_Proto
0	172.16.253.129	53	8.8.8.8	53	UDP	DNS	
1	172.16.253.129	53	4.2.2.2	53	UDP	DNS	
2	172.16.253.129	123	65.55.21.24	123	UDP	NTP	
3	172.16.253.254	67	172.16.253.129	68	UDP	DHCP	
4	172.16.253.129	1132	216.146.38.70	80	TCP	HTTP	
5	172.16.253.129	1643	208.83.223.34	80	TCP	SSL	TOR
6	172.16.253.129	2130	128.31.0.39	9101	TCP	SSL	TOR
7	172.16.253.129	2147	84.54.134.209	39030	TCP	SSL	TOR
8	172.16.253.129	2148	82.96.35.6	443	TCP	SSL	TOR
9	172.16.253.129	2149	98.220.247.151	443	TCP	SSL	TOR
10	172.16.253.129	2150	87.106.249.118	443	TCP	SSL	TOR
11	172.16.253.129	2151	31.172.30.2	443	TCP	SSL	TOR

Hình 4 (Reese, 2016) – Mẫu gói thông tin tóm gọn của giao dịch Tor

Bạn có thể thấy, để tìm ra giao dịch Tor, nó quan trọng để có thể phát hiện ra những chứng thư ngẫu nhiên.Bro script cho phép bạn làm việc này, tìm tại: https://raw.githubusercontent.com/sethhall/bro---junk----drawer/master/detect----tor.bro. Một phương pháp khác để tìm ra giao dịch Tor là thông báo liên lạc với dịch vụ Tor đã xác định. Tuy nhiên, bạn sẽ không chỉ muốn dựa vào việc tìm ra giao dịch đáng ngờ qua server Tor, bởi vì những server này có thể được sử dụng với nhiều mục đích khác nhau, một trong số chúng với mục đích chính đáng. Bạn có thể tìm danh sách của server Tor đã xác định tại: https://www.dan.me.uk/torlist/. Sử dụng kết hợp giữa Bro script để phát hiện ra chứng thư ngẫu nhiên và thông báo liên lạc với server Tor đã xác định, cho phép phát hiện đúng đắn và hiệu quả của hoạt động Web chìm trong mạng của bạn.

Phát hiện giao dịch là một chuyện, nhưng nhận diện giao dịch thế nào khi phân tích gói thông tin tóm gọn? Để có thể nhận diện thành công giao dịch Tor khi phân tích các gói thông tin, bạn cần phải phân tích số liệu trong giao thức để thấy được những sự khác biệt trong các hệ thống sử lý SSL khác nhau. Công cụ có thể sử dụng là CapLoader, dùng "Cổng Giao Thức Dò Tìm Riêng Biệt". Khi tải pcap file vào CapLoader, nó có khả năng nhận diện giao thức Tor, dựa vào phương thức số liệu của nó để tìm ra giao thức, mà không cần đến số của cổng.

Hình 5 dưới đây:



Hình 5 (Hjelmvik, 2013) – Cách sử dụng CapLoader để nhận diện giao dịch Tor

Có rất nhiều phương pháp đưa ra có thể giúp phát hiện và nhận diện giao dịch Web chìm. Bây giờ hãy chú ý đến nội dung sẵn có trong Web chìm

2.2. Có những gì trong đại dương:

Có rất nhiều nội dung sẵn có trong Web chìm. Vài trong số đó là tốt và số còn lại xấu. Nội dung xấu bao gồm những thứ như Dịch Vụ Ẩn Web Đen, Hidden Wiki và Silk Road. Mặc dù Silk Road đã bị đánh sập, vẫn còn một số site phổ biến đang chạy và hoạt động để ghé thăm. Những mảng nội dung này sẵn có trong Web đen

2.2.1. Web đen:

Web đen là một khu vực nằm trong Web chìm. Một số người bị nhầm lẫn giữa Web chìm và Web đen, nghĩ rằng chúng là một. Nó rõ ràng không phải như vậy. Web đen chủ yếu được truy cập qua phần mềm khách gọi là Tor, sẽ được bàn luận chi tiết thêm trong bài viết này. Tor là một trình duyệt đặc biệt cho phép bạn truy cập đến Web đen. Một cách sử dụng phổ biến của Web đen là liên hệ với mã độc. Một lượng lớn mã độc sử dụng Web đen để giao thiệp với các server Chỉ Huy và Kiểm Soát (C&C). Một ví dụ của mảng mã độc là SkyNet.

SkyNet là một Trojan có khả năng tấn công DDoS hoặc đào Bitcoins. Nó dùng Dịch Vụ Ẩn cung cấp bởi Tor để giao thiệp nặc danh với các C&C server của nó. Một lợi ích của việc sử dụng các Dịch Vụ Ẩn cho việc giao thiệp C&C là mã hóa, cho nên nó che giấu được nguồn gốc, điểm đến và trọng tải. Một lợi ích khác là người sở hữu các C&C server có thể quay vòng chúng, đến khi

chúng có thể tự tái sử dụng được chìa khóa cá nhân cho Dịch Vụ Ẩn. Việc sử dụng Dịch Vụ Ẩn trong Web đen rất có hiệu lực.

2.2.2. Dịch Vụ Ẩn Web đen:

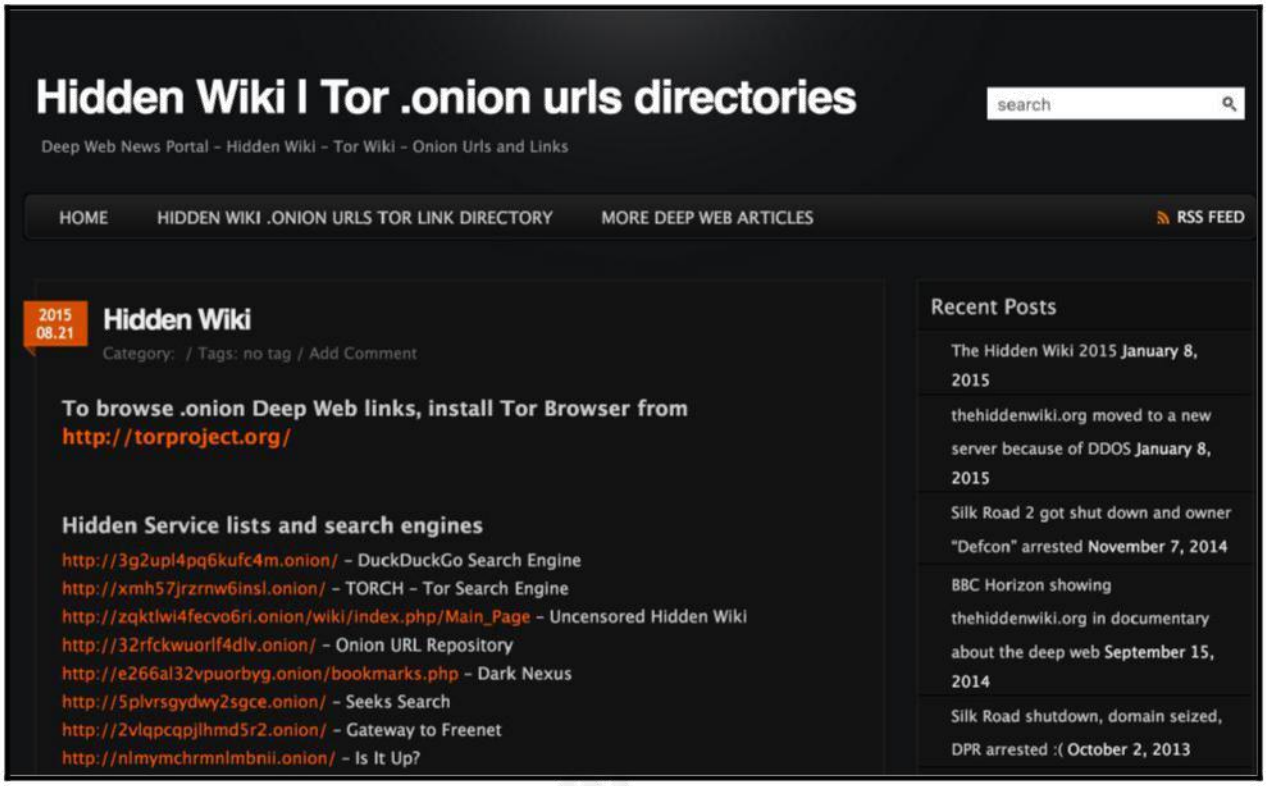
Dịch Vụ Ẩn trong Web đen được sử dụng để cung cấp các dịch vụ đa dạng cho người dùng Web đen, trong khi danh tính các người dùng ở chế độ nặc danh. Một vài loại dịch vụ: tài chính, liên lạc, kinh doanh, tin tức, phim khiêu dâm, công cụ tìm kiếm, lưu trữ dữ liệu và Dịch Vụ Ẩn Chỉ Dẫn và Cổng. Nhiều dịch vụ riêng biệt kết nối với những loại này. Ví dụ, nếu bạn muốn sử dụng một vài Dịch Vụ Tài Chính Ẩn, bạn có thể dùng Bitcoin Fog hay BitBlender. Việc sử dụng Dịch Vụ Liên Lạc Ẩn có thể được đòi hỏi bằng việc sử dụng TorChat hoặc RiseUp. Một vài Dịch Vụ Ẩn cũng được sử dụng cho việc kinh doanh. Nó luôn liên kết vơi Darknet Market. Đôi ví dụ về Dịch Vụ Kinh Doanh Ẩn là Assasination Market và AlphaBay Market. Nếu bạn sử dụng Dịch Vụ Tin Tức Ẩn, bạn có thể sử dụng DeepDotWeb hoặc Wikileaks. Bộ đôi Dịch Vụ Tìm Kiếm sẵn có là The Pirate Bay và Sci-Hub. Free Haven là một trong những Dịch Vụ Ẩn phổ biến nhất cho Lưu Trữ Dữ Liệu và bạn sử dụng The Hidden Wiki là Dịch Vụ Chỉ Dẫn Ẩn phổ biến nhất. Có hàng tấn Dịch Vụ Ẩn sẵn có và chi tiết nó hoạt động thế nào thì phức tạp.

Để xuất bản một Dịch Vụ Ẩn, bạn cần làm cho nó sẵn sang trên mạng lưới Tor để người dùng có thể kết nối với nó. Đầu tiên, người sở hứu dịch vụ sẽ cần ghim điểm mở đầu và xây dựng các đường tròn Tor vòng quanh nó. Điểm mở đầu là rơ le Tor, về cơ bản nó là router. Dịch Vụ Ẩn của bạn có thể chọn đến 10 điểm mở đầu. Dịch Vụ Ẩn của bạn cang phổ biến, thì càng cần nhiều điểm mở đầu. Sau khi lấy các điểm mở đầu, bạn sẽ cần phải quảng cáo Dịch Vụ Ẩn của mình như là “điều gì đó.quan điểm”. Dịch Vụ Ẩn sẽ tạo ra một ký hiệu sẽ bao gồm chìa khóa công cộng của nó và sơ lược các điểm mở đầu được sử dụng bởi nó. Dịch Vụ Ẩn sẽ đánh dấu ký tự này với chìa khóa cá nhân của nó. Ký tự đó được gửi đến bảng băm phân phối, cũng có thể hiểu như một cơ sở dữ liệu. Một khi nó hoạt động, Dịch Vụ Ẩn chính thức được thiết lập và người sử dụng có thể truy cập nó bằng việc yêu cầu tại “thứ gì đó.quan điểm”

Giờ Dịch Vụ Ẩn Danh đó được thiết lập, hãy chú ý đến cách bạn có thể truy cập nó. Đầu tiên, bạn cần phải biết địa chỉ ".onion" đặc trưng đang tồn tại, cũng giống như bạn cần phải biết "google.com" tồn tại trước khi ghé thăm nó. Một khi bạn có địa chỉ ".onion" mà bạn muốn truy cập Dịch Vụ Ẩn của nó, bạn nỗ lực kết nối với nó qua phần mềm khách, như Tor. Người dùng cố tải ký tự cho Dịch Vụ Ẩn từ bảng băm phân phối. Ký tự này sẽ cho người dùng biết các điểm mở đầu và chìa khóa công cộng cần để sử dụng. Trong quá trình này, người dùng cũng đang tọa ra vòng tròn Tor tới một rơ le Tor ngẫu nhiên se được sử dụng như một "điểm gặp gỡ". Một khi bạn có một ký tự và một "điểm gặp gỡ" được tạo lập, người dùng sẽ gửi thông điệp được mã hóa bằng chìa khóa công cộng của Dịch Vụ Ẩn, qua vòng tròn Tor tới một trong các điểm mở đầu, gồm cả vị trí "điểm gặp gỡ" và bí mật không lặp lại. Một khi Dịch Vụ Ẩn nhận được thông điệp, nó sẽ mã hóa lại. Rồi nó sẽ tạo ra một vòng tròn Tor đến "điểm gặp gỡ" và gửi bí mật không lặp lại. Cuối cùng, "điểm gặp gỡ" thông báo cho người dùng kết nối thành công. Một khi nó diễn ra, người dùng có thể giao thiệp với Dịch Vụ Ẩn qua các vòng tròn Tor tới "điểm gặp gỡ". Một trong những dịch vụ được sử dụng phổ biến là Hidden Wiki.

2.2.3. Hidden Wiki:

Hidden Wiki là một site gồm link liên kết với nhiều Dịch Vụ Ẩn khác nhau sẵn có trong Dark Web. Bạn có thể thấy hình chụp ở Hình 6 dưới đây, đây là một đoạn trích ngắn của trang chủ Hidden Wiki.



Hình 6 – Trang chủ của Hidden Wiki

Hidden Wiki gốc được tạo ra một ít trước tháng 10 năm 2011 và có thể truy cập được qua Tor. Nó được chạy qua một tên miền ảo cao cấp nhất .onion. Tên miền ảo cao cấp nhất là tên miền không tham gia vào DNS chính thức. Trong khoảng tháng 8 năm 2013, site làm chủ trong Freedom Hosting, nó là một trong những web chủ dịch vụ lớn nhất được sử dụng bởi những Dịch Vụ Ẩn vào lúc ấy. Trong tháng ba năm 2014, Hidden Wiki bị hack và chuyển tới một site gọi là Doxbin. Doxbin là site được sử dụng để mở PII ra. Một khi nó diễn ra, nội dung trong Hidden Wiki bắt đầu bị phản chiếu bởi một vài địa điểm khác. Chính vì điều này, không còn một Hidden Wiki đơn lẻ nữa. Site phổ biến nhất là Hidden Wiki nối với Silk Road.

2.2.4. Silk Road:

Silk Road là một nơi giao thương trực tuyến trong Web đen được sử dụng để bán sung, ma túy, dữ liệu cá nhân, mã độc và nhiều hơn nữa. Có rất nhiều chợ đen Amazon và được tạo bởi Ross William Ulbricht. Silk Road hoạt động trong khoảng 2 năm trước khi nó bị đánh sập vào năm 2013. Trang này có cảm giác giống như Amazon vậy. Bạn có thể mua được nhiều loại, tìm kiếm sản phẩm và liên hệ với người bán. Một khi vạn tìm thấy một thứ gì mà bạn muốn mua, bạn cứ thêm nó vào rỏ và thanh toán. Tiền

dùng để trả cho những món hàng là Bitcoin, loại tiền vẫn sử dụng

tới ngày nay. Bitcoin là loại tiền ảo được tạo ra và dự trữ kiểu điện

tử không có dấu vết của giấy và về cơ bản không phát hiện ra

được. Mỗi người dùng Silk Road được yêu cầu một địa chỉ

Bitcoin. Những địa chỉ này được lưu trữ trong server của Silk Road

trong một cái "ví". Cho tới khi bạn nghi ngờ về việc giao hàng, nó

vẫn rất không ăn khớp. Gói thông tin hàng của bạn có thể hoặc

không thể bị chặn bởi việc thi hành pháp luật, những đó là rủi ro

2.2.5. Cơ Sở Dữ Liệu Web Chìm:

Một số lượng cơ sở dữ liệu dưới Web chìm có thể được sử dụng cho những thông tin hữu ích. Hãy bắt đầu với người nghiên cứu cơ sở dữ liệu. Những cơ sở dữ liệu này có thể được sử dụng để tìm kiếm khái quát và cũng để tìm một vài trang phổ biến như: pipl.com và spokeo.com. Ảnh chụp được từ pipl.com ở hình 7 dưới đây:

Hình 7 – Trang chủ của pipl.com

Những dạng site này có thể được sử dụng để tìm kiếm thông tin cá nhân như số điện thoại của họ, địa chỉ, thân nhân và nhiều thứ khác. Cơ sở dữ liệu khác có thể được tìm kiếm do những đứa trẻ được nhận nuôi muốn tìm kiếm cha mẹ ruột. Một vài cơ sở dữ liệu khác được tìm kiếm trong Web chìm là: cơ sở dữ liệu phả hệ, ghi chép về mồ mả, cơ sở dữ liệu về lịch sử xã hội và file lưu trữ. Giờ chúng ta đã biết trong đại dương có gì, hãy xem cách truy cập nó nhé.

2.3. Xuống dưới đáy đại dương:

Có rất nhiều nội dung dưới Web chìm, nhưng bạn cần phải có cách để tiếp cận những nội dung đó. Có hai cách để bạn tiếp cận Web chìm. Bạn có thể sử dụng Web nổi để truy cập nội dung qua công cụ tìm kiếm Web chìm hoặc bạn có thể sử dụng một mảng phần mềm gọi là Tor.

2.3.1. Tor:

Định Tuyến Hành (TOR) là tình duyệt web miễn phí, biến thể của Firefox. Bạn có thể chạy nó trong tất cả các hệ điều hành thông thường như: Windows, Mac OS X, Linux. Tor dùng để kết nối tới Web chìm, trong khi vẫn giữ tính nặc danh. Ghé thăm:

https://www.torproject.org/download/ một cách đơn giản để tải Tor, Mục đích của Tor là để cung cấp giao thức mạng có thể giúp dữ liệu truyền tới nó một cách ẩn danh. Khi sử dụng Tor, các gói thông tin của bạn đi qua một vài server được mã hóa trước khi đến địa điểm của nó. Tất cả các server này gọi là rơ le Tor, cấu tạo như các routers. Có một hàng nghìn server chạy trên thế giới. Khi các gói thông tin của bạn được gửi qua mạng lưới Tor nó giảm những mảng của đoạn đầu bao gồm thông tin có thể nhận diện nơi gói thông tin sẽ đến hoặc nơi gói thông tin sẽ đi. Khi những gói thông tin của bạn đi từ rơ le tới rơ le, nó giải mã đầy đủ chính xác dữ liệu để nhận biết gói Tor rơ le từ đâu và điểm đến tiếp theo là gì. Nó khoog giải mã thêm bất cứ thông tin gì. Cách khác để truy cập nội dung dưới Web chìm là thông qua Công Cụ Tìm Kiếm Web Chìm.

2.3.2. Công Cụ Tìm Kiếm Web Chìm:

Công Cụ Tìm Kiếm Web Chìm được sử dụng để truy cập nội dung của Web Chìm từ Web nổi. Lưu ý rằng nội dung mà bạn có thể truy cập bằng những công cụ tìm kiếm này bị hạn chế so với nội dung mà bạn có thể truy cập bằng Tor. Một vài công cụ tìm kiếm mà bạn có thể dùng cho mục đích đó là The WWW Virtual Library, SurfWax và IceRocket. Ảnh chụp từ The WWW Virtual Library ở hình 8 dưới đây:

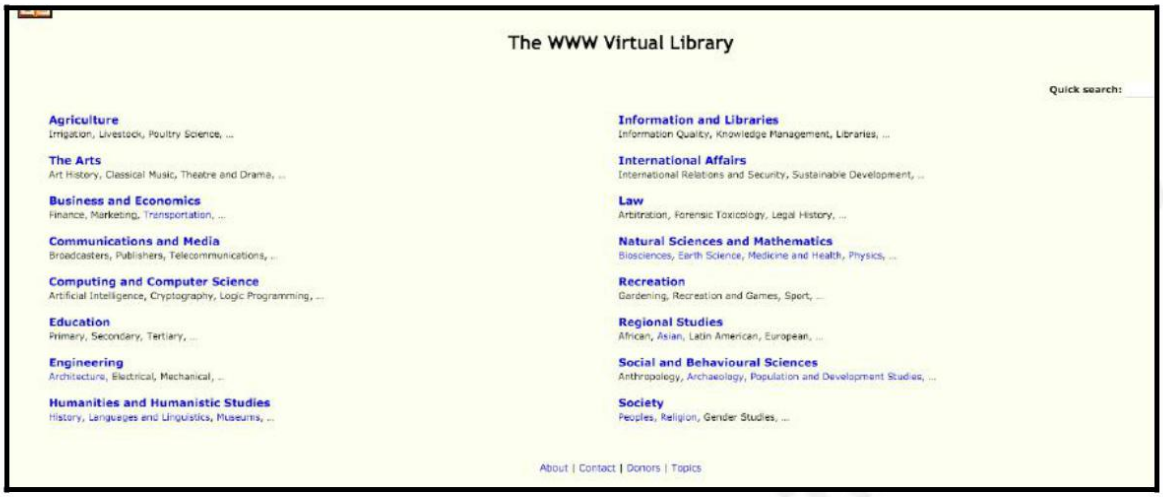
Hình 8 – Trang chủ của The WWW Virtual Library

Những công cụ tìm kiếm này có khả năng thông báo tới Dịch Vụ

Ẩn Web Chìm qua Tor và rơ le của nó, phân giải địa chủ .onion,

rồi đưa nội dung trở lại trình duyệt web thường dùng trên Web nổi.

Có cả kẻ tốt lẫn xấu khi sử dụng một trong những cách này để truy



The WWW Virtual Library

Quick search:

Agriculture
Irrigation, Livestock, Poultry Science, ...

The Arts
Art History, Classical Music, Theatre and Drama, ...

Business and Economics
Finance, Marketing, Transportation, ...

Communications and Media
Broadcasters, Publishers, Telecommunications, ...

Computing and Computer Science
Artificial Intelligence, Cryptography, Logic Programming, ...

Education
Primary, Secondary, Tertiary, ...

Engineering
Architecture, Electrical, Mechanical, ...

Humanities and Humanistic Studies
History, Languages and Linguistics, Museums, ...

Information and Libraries
Information Quality, Knowledge Management, Libraries, ...

International Affairs
International Relations and Security, Sustainable Development, ...

Law
Arbitration, Forensic Toxicology, Legal History, ...

Natural Sciences and Mathematics
Biosciences, Earth Science, Medicine and Health, Physics, ...

Recreation
Gardening, Recreation and Games, Sport, ...

Regional Studies
African, Asian, Latin American, European, ...

Social and Behavioural Sciences
Anthropology, Archaeology, Population and Development Studies, ...

Society
Peoples, Religion, Gender Studies, ...

About | Contact | Donors | Topics

2.4. Cư dân Web chìm:

Có một vài dạng người khác nhau truy cập Web chìm. Họ đều có những động cơ và lý do khác nhau. Thế nhưng, họ có thể được chia làm 2 loại: tốt và xấu

2.4.1. Những người tốt:

Có những người không phải là tội phạm truy cập Web chìm. Ví dụ, một ai đó chỉ là muốn truy cập Web chìm một cách ẩn danh. Họ không có động cơ hay ý định phạm tội, nhưng họ chỉ không muốn để mọi người biết hoặc vì đó là việc tư. Một số khác sử dụng Web chìm là quân đội, cảnh sát và nhà báo. Một những nhà báo có tiếng về An Ninh Thông Tin sẽ sử dụng Web chìm cho việc nghiên cứu để viết bài là Brian Krebs. Trang Web của Brian Krebs: krebsonsecurity.com, và cũng là một blog online nổi tiếng về tin tức An Ninh Thông Tin. Quân đội và cảnh sát sử dụng Web chìm để giám sát các hoạt động đáng ngờ, như buôn bán ma túy, PII và súng. Nhà thi hành pháp luật cũng tương tác với những kẻ buôn bán những thứ đó, lật tẩy các đầu mối để bắt chúng. Những kẻ buôn bán những thứ này được phân loại thành "kẻ xấu"

2.4.2. Kẻ Xấu:

Những kẻ dùng Web với ý định hiểm độc và phạm tội là những "kẻ xấu". Một vài hoạt động phổ biến nhất là chúng tham gia vào việc buôn bán hàng cấm như ma túy, mã độc và những thứ khác. Tội phạm mạng còn dùng Web chìm để giao thiệp một cách ẩn danh với người khác để bán những thông tin bị ăn trộm như số thẻ tín dụng và bệnh án.

3. Kết Luận:

Web chìm là phần lớn nhất trong Internet, vậy mà đa số lại không hề biết về nó hay chưa từng truy cập nó. Nó có thể được dùng cho mục đích tốt và xấu, hoạt động hợp pháp và phi pháp. Quan trọng để hiểu là nó không phải xấu hoàn toàn. Có rất nhiều thứ tốt về Web chìm, gồm quyền cá nhân khi truy cập Internet. Việc hiểu về Web chìm và khả năng của nó là cần thiết cho tương lai của Internet và mong rằng bài viết này giúp đạt được mục đích đó.