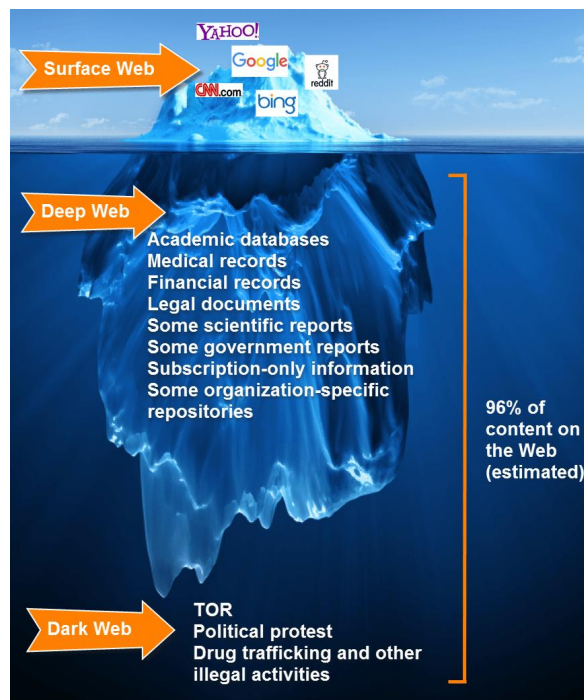


Security Tools Lab 2

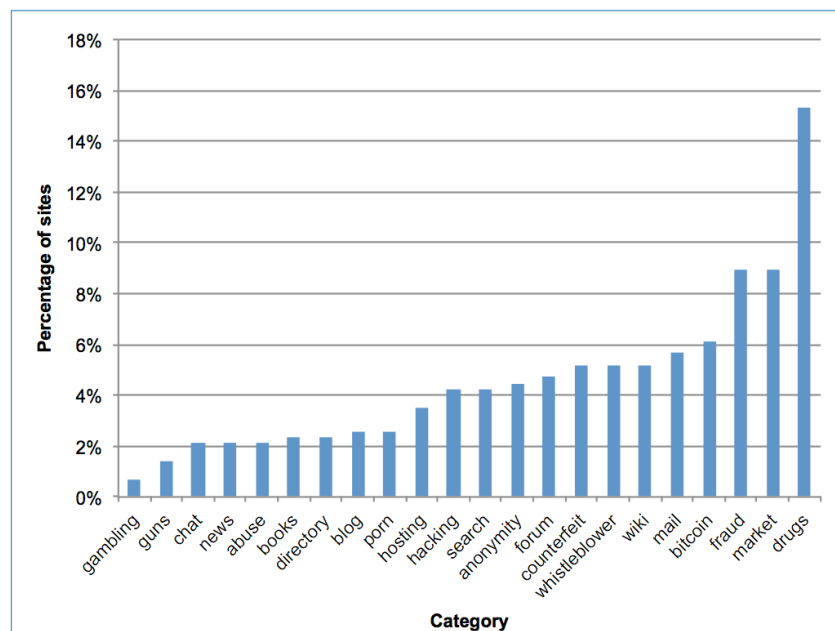
Project 3 – Dark Web Analysis (46 pts)

This project can be done by a single student.

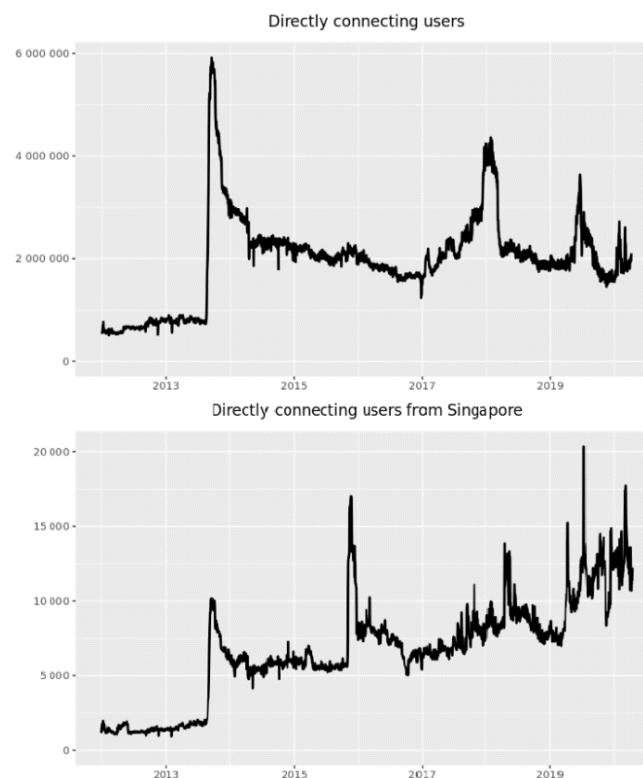
The Dark Web is part of the Deep Web which is an unindexed portion of the Internet. It is invisible to everyday users because its pages and elements cannot be reached using typical search engines. It is often associated with TOR, Freenet, and other anonymizing networks. It requires specialized tools or equipment to access. The part of the Deep Web used as a conduit for illegal and often dangerous activities has been called the Dark Web.



In the public view, the dark web — as its name may suggest — is shrouded in mystery, and often seen as an anonymous haven for criminality. About one third of the dark web may be categorized as dealing with Drugs, Markets, and Fraud, three activities that depend on the anonymity provided by the dark web when skirting around the law.



On average there's about 2 million users per day on the TOR network, one of the largest and the most famous network on the Darkweb according to TOR Metrics (<https://metrics.torproject.org/>). As you can see from the below graph even in Singapore there's around 10,000 users daily. TOR (The Onion Router) is the software used to access this anonymity network, or in other words direct Internet traffic through a free, worldwide, volunteer overlay network consisting of around six to eight thousand relays to conceal a user's location and usage from anyone conducting network surveillance or traffic analysis. You can find more about how TOR works [here](#). URLs in TOR are known as onion links. An example is Facebook on the TOR network which is available at this link : <https://www.facebookcorewwwi.onion/>



Global peak usage always seems to occur when the newsworthy happens, like when major markets are seized and disabled by law enforcement. The peak in 2013 happened when the infamous Silk Road 1 marketplace, the “eBay of illegal drugs” was shut down by the FBI. At that time, it comprised of “more than 70% of the online drug market” and was extensively covered in the news giving rise to the extreme curiosity leading to the highest numbers ever seen.

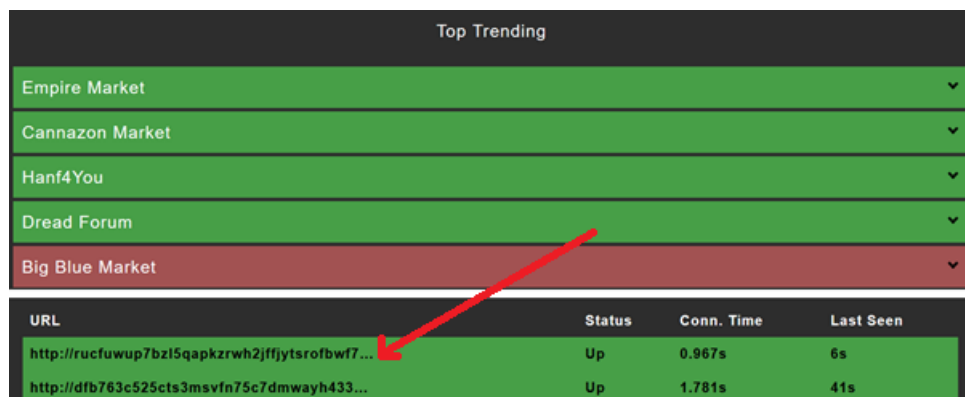
Around the end of 2017, United States and Dutch law enforcement launched Operation Bayonet where they brought down two of the most prominent dark web marketplaces, AlphaBay and Hansa. Before Operation Bayonet, English-speaking cybercriminal activity mainly took place on online dark web marketplaces such as Alpha Bay and Hansa, where hundreds of thousands of vendors and buyers were doing an estimate of over \$1 billion in illegal trade. In May 2019 an internationally coordinated operation led to the takedowns of two more dark web marketplaces, Wall Street Marketplace and Valhalla Marketplace.

Materials provided are for informational and technical training purposes only. It is intended to familiarize you with some of the methods, tools, and services used to surf the deep/dark web. We do not encourage or support using the information presented in this guide for illegal or unethical purposes. A new world will be revealed to you (if you are not already aware of it) during this project and it is only to be for research purposes rather than subscribing to unethical and illegal behaviour which might be very tempting.

Accessing TOR isn't possible from SUTD network unless you use a proxy. It will be easier if you use your home network or your phone hotspot.

Accessing the TOR network

1. Use any Ubuntu/Kali VM obtained during our Labs.
2. Download the Tor Browser (<https://www.torproject.org/>) on your chosen VM.
3. It will be downloaded as a TAR file which you'll need to extract to a folder.
4. Once extracted, you can click on 'Start Tor Browser' in the folder.
5. Because of their popularity for the wrong reasons, Darkweb markets are constantly targeted by law enforcement (DDos attacks, etc) and they often change their Onion links frequently. **On your TOR browser** navigate to <https://onion.live/> to find some of the latest links and most popular sites/markets. Click on any of the green links (red arrow below) to access the darkweb site. If it doesn't connect it most probably means that it has been taken down. In such case try other green links (mirros) until you can access one.



Top Trending			
Empire Market			▼
Cannazon Market			▼
Hanf4You			▼
Dread Forum			▼
Big Blue Market			▼
URL			
http://rucfuwup7bz15qapkrwh2jffjytsrofbwf7...	Up	0.967s	6s
http://dfb763c525cts3msvfn75c7dmwayh433...	Up	1.781s	41s

6. With the dark web marketplace seizures mentioned above, the positive impact has been a breach of trust in dark web criminal trade. This has caused criminals to consider new ways for generating trust in the underground and fortunately there hasn't been much success yet. While dark web markets, such as Empire, certainly still exist, no market has yet risen to the prominence of Silk Road, AlphaBay, or Hansa. New criminal marketplaces continue to crop up, but they struggle to grow or decide to tread lightly with the growing fears of law enforcement disruptions and takedowns. To grow, these criminal marketplaces need a solid reputation, financing to scale, security to maintain current users, and trust to gain more traction. As of Apr-2020, below are the currently most popular markets.

Market	Listings	Vendors	Year Started	Availability	Items	Status
Empire	35k	4309	2018	99%	Various	Active
Cannazon	1.5k	198	2018	92%	Drugs	Active
Versus	6k	226	2019	92%	Various	Active
Monopoly	0.5k	80	2018	72%	Various	Active
Apollon	34k	1,214	2018	79%	Various	Exit Scam*
Berlusconi	120k	2,173	2016	0%	Drugs	Seized [#]

*Running away with users deposited money | #By Italian police in Aug-2019

The Project

Online markets hosted as Tor hidden services provides escrow services between buyers & sellers transacting in Bitcoin or other cryptocurrencies, usually for drugs or other illegal/regulated goods; the most famous DNM was Silk Road 1, which pioneered the business model in 2011. Over the years there has been many data scrapes mostly for research purposes and these scrapes covered information ranging from vendor, article, price, reputation, feedback, images and anything which could be scraped from a normal ecommerce website.

Some of the well-known data scrapes are :

- Gwern Branwen data : Scraped from 2013-2015 (1.5TB of data) by a freelance researcher and comprehensively used in many research and articles to show the dominance of illicit drug trade on the darkweb.
- Agora 2014-2015 : Made by reddit user "usheep" who threatened to expose all the vendors on Agora to the police and Agora shut down a few months after. It contains vendor, category, item, description, price, origin, destination and ratings.
- Kilos : Data released by the search engine from 6 markets (Apollon, CannaHome, Cannazon, Cryptonia, Empire, & Samsara) on 13-Jan-2020 containing 200,000+ reviews of transactions from Feb-2018 to Jan-2020 where we know the vendor, the timestamp of the review, the tone of the review (+ve, -ve, neutral) and the value in bitcoin.

The project is to analyse the Kilos data (provided) since it's the most recent one to understand the current and latest trends in the Darkweb. The data format is :

site, vendor, timestamp, product score, value_btc, comment

Timestamp is in the format of number of seconds since Unix epoch. All reviews from Cryptonia market have their timestamp as 0 due to an issue with the scrape. Score is 1 for positive review, 0 for neutral review, and -1 for negative review. value_btc is the bitcoin value of the product being reviewed.

The current data scrape as it is, is too simplistic and needs to be augmented to generate more features like 'Product Category', 'Vendor Popularity', 'Dollar Value'. For example, if you at Column 5 in the data, the vendor is 'DrunkDragon'. There's only around 1500+ unique vendors in the whole dataset. Doing a simple search on Recon under 'vendors' shows that this vendor deals mostly in credentials and data. Due to the number of records, there's no way to do it manually and you'll most probably have to write a script to obtain the information for you to determine the product type of the vendor. Within the data scrape itself you can aggregate all entries (through a script of course, not manually) with 'Drunk Dragon' to find out his total numbers of feedbacks and his popularity based on the number of +ve, -ve or neutral reviews.

This is just an example and the more features you can generate the better it'll be for your analysis. You're free to choose the way you wish to analyse and below are some ideas :

- Category Sales & Popularity
- Top Vendors and sales over time across categories
- Expensive items v/s Cheaper items comparison across categories
- Vendor lifetimes
- Specific category sales over time (ransomware)
- Traits of Exit scams (Apollon is officially declared one and its data is available)

Marking Rubric

Problem Statement (What you're trying to achieve) – 5

Feature Generation & Methods – 10

Analysis – 15

Report – 10

Discussion – 6