2013

Determining What Characteristics Constitute a Darknet

Symon Aked
Edith Cowan University, secau.2013@tanstaafl.com.au

Christopher Bolan
Edith Cowan University, c.bolan@ecu.edu.au

Murray Brand
Edith Cowan University, m.brand@ecu.edu.au

DOI: 10.4225/75/57b561bfcd8e0
Originally published in the Proceedings of the 11th Australian Information Security Management Conference, Edith Cowan University, Perth, Western Australia, 2nd-4th December, 2013

DETERMINING WHAT CHARACTERISTICS CONSTITUTE A DARKNET

Symon Aked[1], Christopher Bolan, Murray Brand
[1]School of Computer and Security Science, Edith Cowan University, Perth, Australia
secau.2013@tanstaafl.com.au, c.bolan@ecu.edu.au, m.brand@ecu.edu.au

Abstract
Privacy on the Internet has always been a concern, but monitoring of content by both private corporations and Government departments has pushed people to search for ways to communicate over the Internet in a more secure manner. This has given rise to the creations of Darknets, which are networks that operate "inside" the Internet, and allow anonymous participation via a de-centralised, encrypted, peer-to-peer network topology.

This research investigates some sources of known Internet content monitoring, and how they provided the template for the creation of a system to avoid such surveillance. It then highlights how communications on the Clearnet is fundamentally different to that of a Darknet, and examines the characteristics that determine whether or not a network could be classified as a Darknet. Selection of said characteristics is based on how the network was developed, what its intended goals were, and how it implemented technical measures to meet said goals. Five characteristics were found that could be used to determine if a network is to be be classified as a Darknet.

Keywords
Darknet, encryption, anonymity, classification.

INTRODUCTION

Connectivity to the Internet has increased almost exponentially since its introduction to the mainstream (International Telecommunication Union, 2013). In December 2012, there were approximately 12,161,000 Australian Internet subscribers (Australian Bureau of Statistics, 2013b) out of a total population of approximately 22,880,000 (Australian Bureau of Statistics, 2013a). Globally, around 2.7 billion people were online in early 2013 – around 39% of the world's population (International Telecommunication Union, 2013). Such usage is accompanied by a significant flow of data.

In 2011, approximately 30.73 Exabytes (approximately $3.30*10^{10}$ Gigabytes) of data was transmitted across the Internet, and it is estimated that this quantity of data will grow to 110.28 Exabytes by 2016 – a growth of approximately 358% in just 5 years (Cisco, 2012). The capability and capacity that such interconnectivity affords is altering not only the way in which the technology is used, but also users physical lives with internet based activities firmly entrenched in daily routines (Sandvine, 2013).

Given the popularity of Internet communications, it was perhaps always likely that both private and public entities would take an interest in what information is being exchanged. It could also be said that those that place a high value to their security and anonymity would seek ways to ensure their communications could not be monitored. To this end networks referred to as "Darknets" were created to ensure that data exchanged could not be intercepted, altered or read by an outside party.

Investigating what characteristics constitute a Darknet will help formalise what is currently a fairly fluid definition. A formal definition will aid in its correct use in both academic and other literature.

MONITORING INTERNET CONTENT

It is typical for Governments to enact laws that require the national mail carrier and ISPs (Internet Service Providers) to have provisions for content control. Such provisions exist in Australia for mail to be checked for contents that may be prohibited, and in the same way, ISPs must stop access to content that is deemed to be unacceptable (Commonwealth Consolidated Acts, 1997). However, filtering content under the Act has caused collateral damage, and could be seen as a way to implement a wider Internet content filter (LeMay, 2013). Some countries appear to have already taken this monitoring a step further - in Russia, Skype calls are able to be monitored and the geographic location of the participants able to be divulged (Sergina, Nikolsky, & Silonov, 2013).

People are able to use services such as email and VoIP, as well as social networks such as Facebook, Twitter and LinkedIn to easily communicate with friends, family and members of the public across the world. With near instant delivery of electronic communications, disseminating information that is of interest to the intended audience is both simple and effective. Although many messages may be of low significance to people outside of certain circles, often the ability to get accurate information out about local events is of a global significance. It was the ability of citizens of Arab nations to be able to communicate via electronic means the details of revolutionary uprisings in 2010, known as Arab Spring, that led to an accelerated timeframe of political upheaval (Stepanova, 2011).

With the Internet playing such a significant role in the lives of their citizenry, many governments are searching for ways to control, monitor and restrict the usage of such services under the banner of protecting their citizens (Warf, 2010). Such control is seen as an anathema to the widely espoused freedom that such technologies have afforded previously (Cardozo, Cohn, Higgins, Hofmann, & Reitman, 2013). The secret PRISM (Russell, 2013) and XKeyscore (Greenwald, 2013) programs ran by the NSA (National Security Agency) of the United States show that global Internet data is being analysed. Such security concerns may not completely unfounded though; there is a significant record of the Internet being used to facilitate crime and ideas that may be seen to threaten the security or safety of the general populous (Chambliss, 2011). Such concerns, which coupled with the ideologies of more restrictive regimes, have lead countries such as China to construct electronic perimeters around their citizens that are able to monitor and filter which Internet services they are permitted to access (Chen, 2012). This ensures said citizens do not access material that may be seen as detrimental to the Government.

Perhaps one of the reasons that Governments are seeking such control is due to the existence of websites such as Wikileaks (Wikileaks, 2013) and The New Yorker's Strongbox (The New Yorker, 2013) that cater to corporate and Government whistle-blowers, allowing them to divulge information that they cannot afford to be traced back to them. Whilst seen by a section of the community as a means of dissemination of information, Governments may view the leaking of confidential information as a threat to national security and undermining the secrecy that Governments may take for granted. Actions that previously may have been unthinkable, such as the Watergate scandal, almost pass without comment. It is not uncommon for Governments to acquire information that would previously have been unobtainable for them(Sherman, 2013) (Gallagher, 2013).

Companies that create media and other digital content are also eager to implement restrictions, to prohibit their creations from being illegally transmitted over the Internet. To help facilitate this, the DMCA (Digital Millennium Copyright Act) was passed in 1996 (United States of America Government, 1996). This United States copyright bill was designed to prevent the circumvention of content protection methods. It also provides a way for copyright holders to enforce the removal of content from Internet sites that violates their intellectual property copyright, via DMCA takedown notices. The Government of

some countries, including New Zealand (New Zealand Government, 2011) and France (Legifrance, 2009), have legislated that should intellectual property copyright be infringed multiple times, then criminal prosecution may occur. As this step requires that ISPs divulge personal information regarding leases of IP addresses during certain timeframes, concealing one's identity may become more of a concern.

Not being able to be have personally identifiable information be easily harvested is also a concern to those that have no interest in divulging confidential information. Advanced Internet data collection and correlation techniques are being used to relate a person's actions across the Internet, often without their knowledge or consent. This gives rise to data that is of value to both those interested in identify theft crime, and companies that seek to sell data about data about people that access a service on the Internet. Many people are not comfortable with this, which has given rise to web browser extensions that attempt to block such activity, but do not affect other applications (Abine, 2013). This application prevents website being able to identify an individual, and gain data about them and their browsing habits.

With the background of such concerns, a number of people are actively searching and building methods to retain their online freedom. The ability to communicate in a way that allowed for the free, anonymous, and secure exchange of information, while not being at the mercy of any single organisation, was the impetus for the creation of Darknets. These are encrypted networks that operate "inside" the Internet, sending data via the same communications infrastructure that is used by the rest of the Internet.

WHAT IS A DARKNET?

Although the term "Darknet" has no formal definition, it was popularised in the paper "The Darknet and the Future of Content Distribution" (Biddle, England, Peinado, & Willman, 2002), which defined a Darknet as "…a collection of networks and technologies used to share digital content.". The usage of "Darknet" has evolved over time to refer collectively to all encrypted communication networks that allow anonymous participation, and do so using a de-centralised, peer-to-peer network topology, running inside the Internet. Some examples of popular Darknets are Tor (The Onion Router), I2P (Invisible Internet Project) and Freenet. To differentiate between the regular Internet and Darknets, the term "Clearnet" was coined, which refers to the normal, publically accessible Internet at large (Egger, Schlumberger, Kruegel, & Vigna, 2013). Darknets function inside the Internet, and are dependent on it to function, but their design philosophies and goals dictate that they operate under fundamentally different principles than that of the Clearnet. This may be illustrated by outlining how the two networks communicate.

Clearnet communications overview

When a Clearnet user at home opens an URL (e.g. google.com.au) in their Web browser, data is initially sent to their router. This is forwarded to their ISP's routing infrastructure, which in turn forwards to the next router (Kozierok, 2005). This continues until the last hop, which is the server that hosts the requested page, which then services the request – in this case, displaying the Google Australia homepage. The path that the data takes from the home PC to the Web server can be seen by running tools such as tracerout (InetDaemon, 2013e) on the originating computer. Although the exact route that data takes may change over time, based on the routing policy of the network infrastructure, it is likely that subsequent paths will not vary greatly. All data sent via this method is transmitted unencrypted, allowing for data observation, interception or alteration, as can be seen in the figure below.

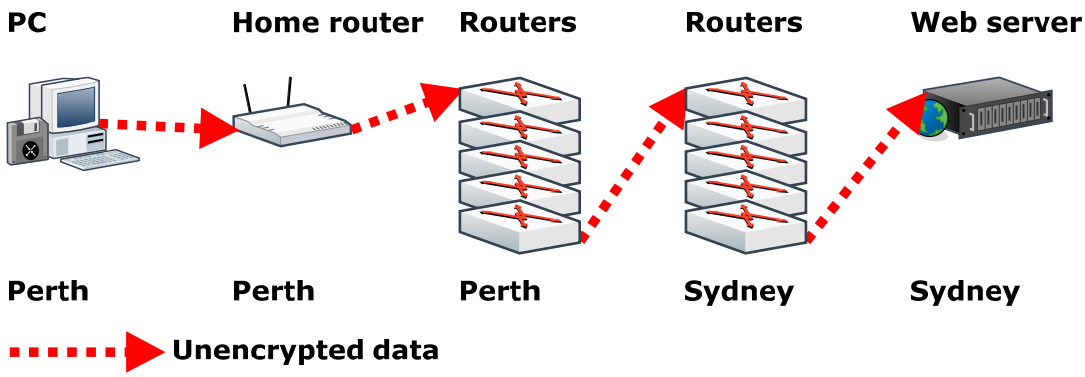

Figure 1 - Data routing across the Clearnet

Darknet communication overview

When a user of the Tor Darknet accesses the same website, data is initially sent to the Tor routing software running on their PC. This software generates encrypted tunnels to one or more Tor entry relays, of which there were approximately 3,500 as of June, 2013. Data is routed from the Tor software through a tunnel to the Tor entry relay, which itself has encrypted tunnels to other Tor relays. Data is forwarded to a number of other Tor relays, before it arrives at a Tor exit node, which is an exit points from Tor into the Clearnet. From here data will traverse Clearnet paths to the server, which then services the request. The exact path that the data took inside Tor cannot be predicted – only the Clearnet route that the data took into Tor and the route it took when it exited Tor can be seen. All data sent via Tor is encrypted in such a way that the original data source and destination cannot be determined, and the data cannot be viewed or changed, as seen in the figure below.

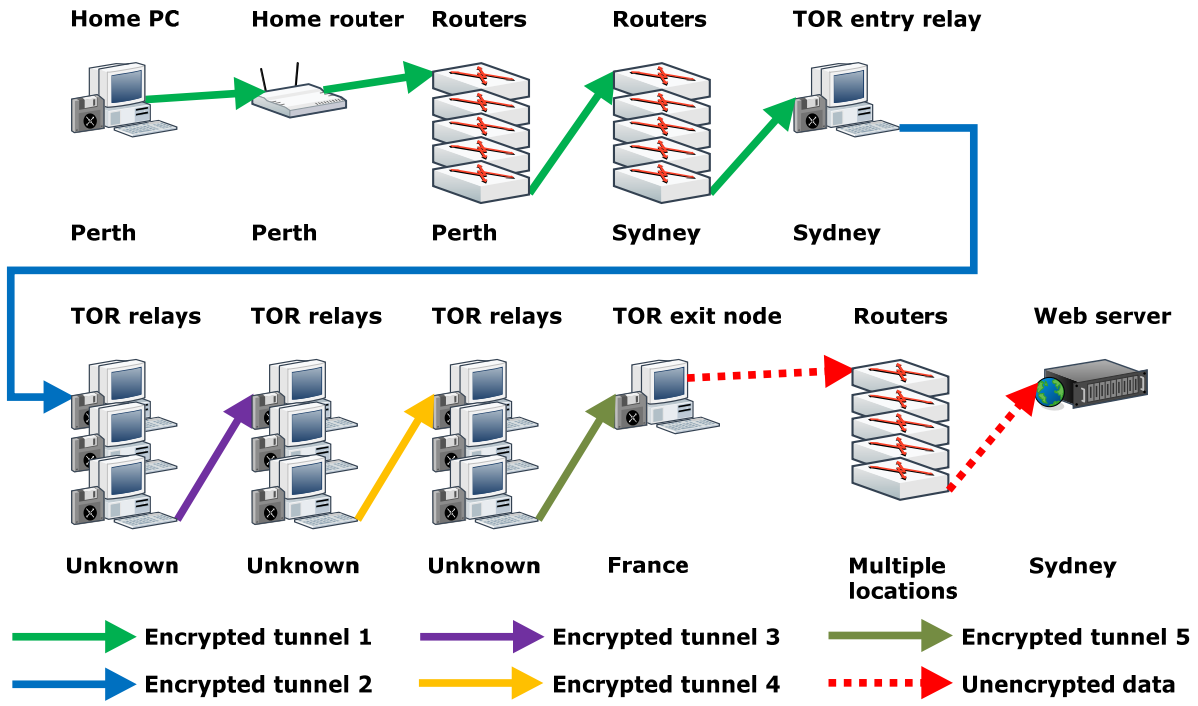

Figure 2 – Data routing across Tor

As well as anonymising their access to Clearnet services, Darknet users may also elect to access websites and other applications that are only available to Darknet users, also known as “hidden services”. These

hidden services are configured to not be available to Clearnet users, and instead to only listen for connections made via a Darknet. Hidden services afford both the host and the client of the application anonymity. Being able to communicate and exchange data anonymously and securely has many legal applications for Darknet users. However, it also allows for illegal services and activities to take place that may not be easily taken down, compared to if the same services were available on the Clearnet.

DARKNET CHARACTERISTICS

There appears to be a general misunderstanding and misuse of the label "Darknet". Terms that have been identified as being used interchangeably include Dark Internet, Dark Address Space, Deep Web, Deepnet, Dark Web, Invisible Web, Undernet, and Hidden Web. However, these terms are not synonyms for "Darknet", and have their own, generally accepted definitions. It is therefore important to determine which criteria constitute a Darknet, and compare it to the Clearnet.

Encryption

Encryption is often not employed on the Clearnet, and its use is mostly confined to situations that require sending authentication credentials, such as logging into an online banking website. Handling encrypted traffic adds extra processing load to the destination's infrastructure, which would add to the costs of offering the service. To ensure that data transmitted cannot be intercepted, changed, observed or read by anyone other than the intended source and destinations, all data that traverses Darknets is encrypted. Although using differing encryption algorithms and implementations, the types of encryption used in Darknets has been proven to take a very long time to decrypt, even for a very well-funded Government organisation, and may be considered secure for the foreseeable future (Smart et al., 2012) – although the United States NSA (National Security Agency) is working to change this (Bamford, 2012). As Darknet participants form part of the routing infrastructure, the extra processing required to encrypt the data stream is shared amongst all participants.

Anonymity

It is usual that, every time a website is accessed on the Clearnet, the source of each request made is logged by the destination. This log contains the IP address of the visitor, which pages they visited, and some characteristics about their operating environment, such as operating system, web browser, and their respective versions. This IP address can be correlated to a household, business or Government agency by their respective ISP. Accessing other services, such as instant messaging or email, may also be logged by the destination. These logs make it possible for law enforcement agencies to piece together a user's Internet activity. In Darknets, a core aim is to preserve every user's anonymity. Each data stream is encrypted and routed in such a way that the source and destination of the request cannot, outside of user or program error, be determined. To prevent other forms of personal information from being leaked, it is common for Darknet application to deliberately mask or sanitise any identifiable information that is sent, such as information commonly provided by web browsers (The Tor Project, 2013b). There are also mechanisms available to applications that run on Darknets for users to maintain a consistent identity, should they wish to do so, thereby enabling users to be pseudo-anonymous. These aliases allow for social networks to be established and utilised.

Routing

Clearnet data takes a path to its destination based on routing rules set up by a user's ISP, their upstream provider, and eventually the preference of the destination. This is facilitated by using technologies such as BGP (Border Gateway Protocol), which governs which of the available telecommunications carriers is used to send data to the next hop of its destination. There are a number of factors that determine which route the data takes, but they are primarily based on latency, cost, geography or congestion of the source and destination of the data. The core Internet routing infrastructure is comprised of dedicated,

expensive hardware, and designed to handle a massive amount of traffic on a continuous basis, with route that data takes can be seen by an end user by running tools such as traceroute (InetDaemon, 2013). For Darknet traffic however, the path that data takes is designed to be untraceable. The route may be based on a number of factors including link stability, latency and bandwidth, but will change frequently and randomly to ensure that analysis cannot determine either the source or destination of the request, thereby assuring the anonymity of both the source and destination of the data (The Tor Project, 2013a). Darknet participants, which usually also act as routers, are themselves unaware of where the data originally came from, knowing only the source and destination of the preceding and subsequent hops (Acquisti, Dingledine, & Syverson, n.d.).

Applications

Once a network connection has been established with an ISP, applications on the host computer can exchange data directly with their intended endpoint via the host operating system's network stack. There are a vast array of applications that exchange data via the Clearnet, such as web browsers, email clients, file sharing tools and games. In Darknets however, after an Internet connection is established, separate software must be installed and ran to form an entry point into a Darknet. Once connection to the Darknet is established, applicable applications must be told to route their data into the Darknet via this conduit. Not all applications are able to communicate on Darknets, which has led to specialised applications that are designed to solely communicate via Darknets, and are generally preconfigured to route data via them.

Visibility

It is trivial for an ISP or other entity with access to the relevant Internet infrastructure to see what sort of activity a Clearnet user is participating in. This information may be logged as part of a transparent proxy, or other "default on" logging system. Should a user be identified for further analysis, then packet capture software can be employed to get a more detailed view of the user's online activities. Connecting to an encrypted Clearnet services can hide the contents of the data transmitted, but the source and destination are trivial to obtain. Although there may be indicators in network traffic that a user is connected to a Darknet (Barker, Hannay, & Szewczyk, 2011), a user's participation may easily be overlooked, provided that there is no other cause to examine the participant, such as excessive bandwidth usage, a higher-than-expected amount of encrypted traffic, or a high amount of traffic on non-standard ports.

The above criteria may be summarised as seen the table below.

Table 1 – Comparison of Clearnet and Darknet attributes

	Clearnet	**Darknets**
Encryption	Sometimes	Always
Anonymity	Users traceable	Users anonymous
Routing	Based on capacity, cost and geography	Based on anonymity, stability and capacity
Applications	Many	Few
Visibility	Obvious	Hidden

CONCLUSION

The Internet is a heavily used resource, replacing many traditional methods of communication. It is therefore perhaps not surprising that many public and private entities have a growing interest in monitoring the content that is being transmitted across it. Given the privacy concerns voiced by people not content to allow said monitoring, being able to avoid having their data be captured and analysed is becoming a pressing issue for some.

Five Darknet characteristics were suggested as the basis for identifying whether or not a network could be classified as a Darknet. When compared against Clearnet attributes, said characteristics show that, although intertwined, the two networks operate under very different parameters. Although further investigation is needed, it is suggested that, should a network meet the presented criteria, then it could be a candidate for being classified as a Darknet.

REFERENCES

Abine. (2013). Stop companies from tracking you. Retrieved 15/05/2013, from http://abine.com/dntdetail.php

Acquisti, A., Dingledine, R., & Syverson, P. (n.d.). On the Economics of Anonymity. http://freehaven.net/doc/fc03/econymics.pdf

Australian Bureau of Statistics. (2013a). Population clock. Retrieved 28/03/2013, 2013, from http://www.abs.gov.au/ausstats/abs%40.nsf/94713ad445ff1425ca25682000192af2/1647509ef7e25faaca2568a900154b63?OpenDocument

Australian Bureau of Statistics. (2013b). Subscribers by states and territories by ISP size, for ISPs with more than 1,000 subscribers. (21/03/2013). http://www.abs.gov.au/ausstats/subscriber.nsf/log?openagent&81530do003_201212.xls&8153.0&Data%20Cubes&094D7ADB9BE886C7CA257B4700137934&0&December%202012&09.04.2013&Latest

Bamford, J. (2012). The NSA Is Building the Country's Biggest Spy Center (Watch What You Say). Retrieved 15/05/2013, from http://www.wired.com/threatlevel/2012/03/ff_nsadatacenter/all/1

Barker, J., Hannay, P., & Szewczyk, P. (2011). Using traffic analysis to identify The Second Generation Onion Router. Retrieved from Research Online website: http://cloud.github.com/downloads/exceptttheweasel/tor-detection/paper.pdf doi:10.1109/EUC.2011.76

Biddle, P., England, P., Peinado, M., & Willman, B. (2002). The Darknet and the Future of Content Distribution. http://msl1.mit.edu/ESD10/docs/darknet5.pdf

Cardozo, N., Cohn, C., Higgins, P., Hofmann, M., & Reitman, R. (2013). Who Has Your Back? Which companies help protect your data from the government? https://www.eff.org/sites/default/files/who-has-your-back-2013-report-20130513.pdf

Chambliss, W. J. (2011). Crime and Criminal Behavior - Internet Crime. http://knowledge.sagepub.com.ezproxy.ecu.edu.au/view/criminalbehavior/n12.xml doi:10.4135/9781412994118

Chen, C. (2012). Beyond “The Great Firewall”: A Closer Look at Online Public Discourse in the People’s Republic of China. http://wesscholar.wesleyan.edu/cgi/viewcontent.cgi?article=1854&context=etd_hon_theses

Cisco. (2012). Cisco Visual Networking Index: Forecast and Methodology, 2011–2016. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360.pdf

Commonwealth Consolidated Acts. (1997). Telecommunications Act 1997 - Sect 313. Retrieved 21/05/2013, from http://www.austlii.edu.au/au/legis/cth/consol_act/ta1997214/s313.html

Egger, C., Schlumberger, J., Kruegel, C., & Vigna, G. (2013). Practical Attacks Against The I2P Network. http://wwwcip.cs.fau.de/~spjsschl/i2p.pdf

Gallagher, R. (2013). FBI Pursuing Real-Time Gmail Spying Powers as “Top Priority” for 2013. Retrieved 21/05/2013, from http://www.slate.com/blogs/future_tense/2013/03/26/andrew_weissmann_fbi_wants_real_time_gmail_dropbox_spying_power.html

George Mason University. (2010). Implications of Declining Mail Volumes for the Financial Sustainability of the Postal Service. from https://www.uspsoig.gov/foia_files/rarc-wp-10-006.pdf

Greenwald, G. (2013). XKeyscore: NSA tool collects 'nearly everything a user does on the internet'. Retrieved 02/08/2013, from http://www.theguardian.com/world/2013/jul/31/nsa-top-secret-program-online-data

InetDaemon. (2013). Traceroute Example. Retrieved 23/05/2013, from http://www.inetdaemon.com/tutorials/troubleshooting/tools/traceroute/example.shtml

International Telecommunication Union. (2013). ICT Facts and Figures. http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013.pdf

Kozierok, C. M. (2005). The TCP/IP Guide. Retrieved 05/09/2013, from http://www.tcpipguide.com/free/t_TCPIPRoutingProtocolsGatewayProtocols.htm

Legifrance. (2009). Haute Autorité pour la diffusion des œuvres et la protection des droits sur internet. Retrieved 13/05/2013, from http://legifrance.gouv.fr/affichCode.do?idArticle=LEGIARTI000020740264&idSectionTA=LEGISCTA000020740339&cidTexte=LEGITEXT000006069414&dateTexte=20130515#LEGISCTA000020740333

LeMay, R. (2013). Interpol filter scope creep: ASIC ordering unilateral website blocks. Retrieved 21/05/2013, from http://delimiter.com.au/2013/05/15/interpol-filter-scope-creep-asic-ordering-unilateral-website-blocks/

New Zealand Government. (2011). Copyright (Infringing File Sharing) Amendment Act 2011. http://www.legislation.govt.nz/act/public/2011/0011/latest/096be8ed806e2a43.pdf

Russell, J. (2013). PRISM: Here’s what you need to know about the US Internet monitoring scandal. Retrieved 09/07/2013, from http://thenextweb.com/insider/2013/06/07/prism-heres-what-you-need-to-know-about-the-us-internet-monitoring-scandal/

Sandvine. (2013). Global Internet Phenomena Report. http://www.sandvine.com/downloads/documents/Phenomena_2H_2012/Sandvine_Global_Inter

net_Phenomena_Report_2H_2012.pdf

Sergina, E., Nikolsky, A., & Silonov, A. (2013). Российским спецслужбам дали возможность прослушивать Skype. Retrieved 21/05/2013, from http://www.vedomosti.ru/politics/news/10030771/skype_proslushivayut

Sherman, M. (2013). Gov't Obtains Wide AP Phone Records in Probe. Retrieved 21/05/2013, from http://bigstory.ap.org/article/govt-obtains-wide-ap-phone-records-probe

Smart, N., Babbage, S., Catalano, D., Cid, C., Weger, B. d., Dunkelman, O., . . . Ward, M. (2012). ECRYPT II Yearly Report on Algorithms and Keysizes (2011-2012). http://www.ecrypt.eu.org/documents/D.SPA.20.pdf

Stepanova, E. (2011). The Role of Information Communication Technologies in the “Arab Spring” - Implications Beyond the Region. http://ponarseurasia.com/sites/default/files/policy-memos-pdf/pepm_159.pdf

The New Yorker. (2013). The New Yorker Strongbox. Retrieved 21/05/2013, from http://www.newyorker.com/strongbox/

The Tor Project. (2013a). Tor: Overview. Retrieved 22/07/2013, from https://www.torproject.org/about/overview.html.en

The Tor Project. (2013b). What is the Tor Browser Bundle? Retrieved 21/07/2013, from https://www.torproject.org/projects/torbrowser.html.en

United States of America Government. (1996). Digital Millennium Copyright Act. http://www.gpo.gov/fdsys/pkg/PLAW-105publ304/pdf/PLAW-105publ304.pdf

United States Postal Service. (2012). First-Class Mail Volume Since 1926. Retrieved 19/05/2013, from http://about.usps.com/who-we-are/postal-history/first-class-mail-since-1926.htm

Warf, B. (2010). Geographies of global Internet censorship. http://link.springer.com.ezproxy.ecu.edu.au/content/pdf/10.1007%2Fs10708-010-9393-3.pdf doi:10.1007/s10708-010-9393-3

Wikileaks. (2013). What is Wikileaks? Retrieved 14/05/2013, from http://www.wikileaks.org/About.html