# Global Spying:
## Realistic Probabilities In Modern Signals Intelligence

**Jonathan Logan** **Steve Topletz (presenter, editing)**

PREFACE

In this article, we will present insight to the realistic possibilities of Internet mass surveillance. When talking about the threat of Internet surveillance, the common argument is that there is so much traffic that any one conversation or email won't be picked up unless there is reason to suspect those concerned; it is impossible that “they” can listen to us all.

This argument assumes that there is a scarcity of resources and motivation required for mass surveillance. The truth is that motivation and resources are directly connected. If the resources are inexpensive enough, then the motivations present are sufficient to use them. This is visible in the economic effect of supply availability increasing the demand. The effect is that since it is more easily done, it will be done more readily. Another fault in the above argument is that it assumes that there is only all-or-nothing surveillance, which is incorrect.

INDEX

## I. RESOURCE REQUIREMENTS

It is important to break down the resources required and methods available as well as the means of surveillance in order to understand what realistic threat mass surveillance of digital communication is. The resources required are Access, Storage, Traffic, and Analysis. In this paper, we are speaking about digital communications, and these methods do not fully apply to purely analog communication, such as POTS (normal telephone service).

### ACCESS

Surveillence requires access to the communication to be surveilled. Data today is transmitted via copper cable lines, fiber-optics, directed micro-wave communication, broadcast radio (WiMAX,WiFi etc.), satellite, and a few other arcane methods . The most profitable transmission media for surveillance, by far, are fiber, broadcast, directed micro-wave, and satellite. Fiber provides the benefit of large amounts of data from a single "cable." Broadcast radio provides the benefit of non-physical accessibility. Directed micro-wave is easily acquired through classic stand-in-the-middle listening. Satellite provides a very big footprint, where one needs only to be standing near the receiver of the transmission.

Fiber cables provide the most interesting targets for surveillance. Almost all international communication eventually goes over a few particular fiber lines, so this is where the tapping is focused. This is a practice far different from the UK / USA Echelon system of the 1980s, which operated mostly by targeting direct micro-wave and satellite transmissions, because international fiber-optic lines were more rare. Today, tapping into fiber is easily accomplished through a variety of methods:  splicing the fiber-optic line, connecting to the repeaters, or tapping into the endpoint routers, and through even more esoteric methods, like bending the fiber and detecting stray "ghost" photons[1]. Tapping in most cases is purely passive, which means two things. First, the signals are being listened to and not intercepted or modified. Second, surveillance-induced artifacts are non-trivial to detect by the endpoint, which means there is no "click" on the phone to tell you that someone is listening in. This is especially true in digital communications espionage, which is the focus of this paper.

Access to fiber-optic lines is mostly accomplished by connecting to repeaters and tapping endpoint routers. That is what is being performed by AT&T at the request of the NSA. This method is inexpensive in resources and easy to implement, plus it requires very few people to know about it and to operate it. In the case of repeater connections, even the fiber owners may not be aware that their lines are being tapped unless they find the tap during routine maintenance.

Civilians generally assume that the Internet consists of millions of independent lines that would have to be tapped individually for mass surveillance. Luckily for signals intelligence gathering and analysis, this is not the case. To tap into 90% of traffic connecting the Eastern Hemisphere to the Western Hemisphere (GUS / RUS / AFRICA / MIDDLE EAST / EU to US), agencies only need access to either 30 fiber cables[2] or half of the 45 landing points[3]. An alternate method to achieve such access to this traffic is to install access devices in just seven of the correct Internet Exchanges[4] (IXs), which are where ISPs and backbones interconnect at a single location. Rest assured, all of above has happened at various scales[5] as intelligence agencies are pitted against each other to gain power through knowledge.

Competition levels of espionage can be represented as many sets of Nash equilibria[6], where allies and enemies are not in distinct groups. In specific game theory, it can be represented by the classic Arms Race model[7], with distrustful parties engaged in noncooperative escalation games.

A special property of the Internet, which lends itself to accessibility, is resiliency in routing; if you can not tap into a specific route, then you can destroy it to have the traffic rerouted through lines that you have full access to. Accidently drop an anchor on a submarine cable, or have an excavator accidentally cut a line, and then execute a Distributed Denial of Service or Table Poison attack against the routers in question. There is an endless amount of innocuous events which are created or exploited for covert access to fiber-optic communications. For example, one event occurred in November 2005, where the cable between Iceland and Scotland was apparently severed[8], rerouting all traffic through the USA. Such an event could easily be used for purposeful traffic rerouting, a tapping opportunity, or both[9,10]. For tapping subjects that require more surgical precision and shorter time windows than typical dragnet operations, there are additional options like breaking into routers to establish “shadow routes" on IXs and landing points. A shadow route is where the traffic between two interfaces is mirrored and a copy is sent to a third virtual interface, such as a GRE tunnel or IPSec Encapsulated Security Payload[11].

It is important to keep in mind that a surveillance organization does not have to cover all nodes or routes for full access. One must simply select the ones with the most connections or throughput to other nodes in order to succeed. Tapping into the connection at *any* endpoint, transmission line, repeater, or router is enough to obtain the access required for mass surveillance. After you have access, the remaining work for mass surveillance is relatively trivial.

## STORAGE

Storage, as well as traffic, are relatively expensive resources. It makes little sense to tap into communication lines and not be able to store the data that you want. However, if you are able to select, reduce, and compress the data you are interested in, then storage resource requirements decrease. Today, the cost of storage using standard products on the market is high when compared to the total amount of traffic traveling the Internet. The cost of storing a year’s worth of traffic is very high; for 2008 alone, it would cost over $33 billion[A]. However, if you use data reduction methods, then the total storage costs are much lower. For example, it is not necessary to store a copy of all traffic each time someone downloads a movie; it is enough to reference the movie. The same applies for webpages, documents, and other uniform communications. By storing only unique Internet traffic at the data-mining facility, storage costs are reduced to much less than 1% of the original projection, which brings mass surveillance into close reach for many organizations like the NSA, which has a projected yearly budget between $3.5b and $4b[12], excluding covert operations (Black Ops). Italy implemented such a system in 2007, named DRAGON[13], to retain data acquired from the mass surveillance of their citizens. Some countries, like Sweden, not only record traffic destined for their country, but they also record all international traffic that crosses their borders[14]. Think of them as a nosy person who not only reads his own mail, but rifles through his neighbors' mailboxes as well.

## TRAFFIC

Captured data must be transferred from the temporary storage on the tapped line to the aggregate data

stored at the data-mining facility. Therefore, data of interest must be transferred to a collection point. Using the above projections, transferring unique traffic from the tapping point to the data-mining facility costs roughly $40 million annually[B]. This is entirely in the financial reach of both large and small intelligence gathering organizations. Although it is not publicly known if any organization does indeed copy and store all unique traffic on the Internet, game theory suggests that if it is both possible and beneficial, then not only is it likely, but also, capable parties will scramble to do so just to remain on par with their counterparts.

## ANALYSIS

Analyzing the stored data is where real intelligence happens, and it is more demanding than both storage and traffic requirements. Post-tapping analysis and offsite analysis should be differentiated; post-tapping is what selects and reduces the data to that which is unique and of interest[15], whereas offsite analysis is where raw data is turned into intelligence to be acted upon. Post-tap analysis typically occurs directly at the tap, and the resulting data is stored. Very little communication is of interest for realtime surveillance, so data is rarely relayed immediately and is typically cached to be transmitted at a time better suited to both the cost and detectability of the surveillance. For the purpose of this document, we will always assume higher and padded costs in an effort to demonstrate the maximum financial requirements. The reality is that the costs are much lower, especially when equipment is purchased in massive quantities and resources are shared by multiple organizations. The cost of post-tap analysis is approximately $4.5k per Gbps of traffic[C]. This means that the post-tap analysis hardware cost for *all unique* global Internet traffic at full network utilization is roughly $530m per year[D] for hardware costs alone. Hardware costs consist of only 48% of the total cost before traffic, with traffic, datacenter upkeep, energy, and storage maintenance making up the rest. This brings the total post-tap costs to approximately $1.13b per year, not including the installation and maintenance of access components, which is an additional $1.5b per year. Offsite analysis costs vary, and depending on what operations and techniques are performed on the unique data collected from the entire Internet, costs could start at a few million dollars and reach up to a $1.5b in yearly costs[E].

**The total cost of surveilling *all* unique Internet traffic in the world is approximately $4.1b**[F], with a variance of around $500m, depending on what is done with the information. Since the regions of interest are different, with some intelligence organizations focusing on multi-national rather than global surveillance, the cost for non-global mass surveillance of the Internet is less than $1.5b per interested party. Eight particular intelligence-service countries have a strong interest in acquiring total global surveillance; those are the United States, the United Kingdom, France, Germany, Russia, Israel, China, and Australia. Other countries are restrained in their interest, limiting their appetites to those of their domestic and foreign intelligence services[16], which request spying on their citizens and neighboring traffic.

Economically speaking, this is far less than many countries spend on things like military weapons or state police[17], all while providing an invaluable threat and strategic intelligence. This financial estimate assumes that the selection of unique communication is 100%, without regard to protocol, and includes all website, e-mail, and VoIP traffic. This estimate also assumes that it is a single party doing the work, and that resources like taps, storage, and manpower are not being shared. In practice, however, many

allied countries share intelligence resources. One probable example would be the United States and Germany sharing hardware taps of Middle Eastern traffic. As we can see in most "developed" countries, today the actual work is outsourced to private contractors by legislative mandate, such as the EU Data Retention Directive[18], which provides no funding and shifts the burden entirely upon private Data Centers and ISPs. In some countries, ISPs are required to provide the access, storage, and traffic components or do it for their own profit by participating with interested 3rd parties, such as Nokia & Siemens[19]. Given the minimal costs compared to both the budgets and perceived benefits, it is naive to assume that mass surveillance is not being employed.

## II. METHODS OF POST-TAP AND OFFSITE ANALYSIS

A netflow is a relationship between one computer and another one; the word "connection" does not really apply to packet-based networks. One thousand active “professional” Internet users create between 30k-50k concurrent netflows with roughly 80 Mbps to 250 Mbps of sustained bandwidth consumption. Occasional Internet users, the majority, create much less. The numbers appear huge at first glance, but applying professional processing equipment and software can reduce those huge numbers to an easier-to-handle set of information that can readily be acted upon. Communication surveillance analysis uses the Escalation of Surveillance concept, executed by four basic methods: Classification, Interpretation, Reaction, and Selection.

Escalation of Surveillance means that, depending on previous analysis, the computers reserve more resources to spy on a specific target. How they do it depends on the rules given to the Reaction component and can be exceptionally complex. The escalation process does not stop at the post-tap analysis stage, but instead, it "trickles up" to the offsite analysis. Additionally, if a target becomes interesting due to escalation, then other people in connection with the target become more interesting as well. This is because of context classification, and it can be summed up as “guilty by association.” Technology makes it possible to interconnect seamlessly and inexpensively the post-tap installations and the semi-automatic creation and updating of reaction rules. Therefore, escalatation of resources spent on ancillary target groups that are connected to an escalated target can happen almost in realtime[20].

When communications are tapped into, the first step for analysis is **Classification**. The two types of classification are *Content* classification and *Context* classification.

Context classification defines what kind of data (protocol) is transferred and who transfers it. Context classification on IP networks, such as the Internet, is trivial because the underlying protocols provide all required information in a form that is easy for computers to read and understand. With the advent of Deep Packet Inspection, the context classification even touches the application protocols (Layer 7 analysis) and payload (classical deep packet inspection). The result is not just having the conclusion "XY reads a Google page", but also being able to state "XY searched for porn on Google.” The data generated by context classification is ideal for storage and later data-mining. Such data sets are relatively small and have a precise meaning. It is fair to assume that the majority of Internet surveillance focuses on context.

Content classification defines what type of data is transferred and what meaning the data has. In most cases, content classification only considers the *type* of data, such as pictures or movies, but in some cases, the *meaning* of the data is of interest. Content classification is especially effective on unique Internet traffic. The Google logo is transferred millions of times every day, however, it is not unique; it is classified once, put into a reference table, and never revisited. The same goes for most web and p2p content. Combined with context classification, a resulting data set would say "XY downloaded a nude picture of Angelina Jolie from webpage Z". The resulting dataset will be less then 200 bytes, regardless of picture size, and by the time the first 5 to 10 packets are transferred, the connection has already been analyzed[21]. One real-life example of this technique is a Bundeskriminalamt[22] operation under the auspices of stopping child pornography. It references known child pornography images, generates a reference, and then watches to see if those references appear in network traffic. The effect is that they will instantly know if anyone on their network is sending or receiving such images. This technology is not limited to images; a checksum of any dataset can be programmed into their scanner, such as sensitive or politically embarassing documents uploaded to Wikileaks. If the content, however, is not unique, then the Classification method fails, and the next method used is Interpretation.

**Interpretation** of unique data means that the data is translated into a form that data-mining can act upon. For e-mail, that means that the text is analyzed by semantics analyzers[23]. Such an analyzer running over this document will return something similar to "Analysis of Internet surveillance feasibility and implementation." These tools are able to find out the most important words, places, times, subjects, and people mentioned in the communication content in only a fraction of a second. The resulting data set for analysis will be relatively small (around 2-5KB), machine readable, and easy to store. After content/context analysis and interpretation are completed, the result is a data set that can be reacted to. Further automated analysis, such as Writeprint[24], can be performed for profile development of specific targets, including author discovery of anonymous publications.

The next two steps are a dual-factor component, requiring both human presets and computer processing. **Reaction** is programmed into the computer as a rule set, as it requires abilities beyond those that a computer can intuitively choose or measure, reporting back if the traffic is interesting or not. In the **Selection** process the data is combined with a vector that holds "points" for the various interests the spying party associates with it. While the programming is a thing performed by humans, the "interest vector" is computed automatically. Depending on the "interest vector", the data might be thrown away, cached locally to be combined with additional data, or transferred to offsite storage and processing. Both Reaction and Selection are completed very quickly, during which the parties of communication are re-classified as well, which accomplishes Escalation.

Computers can make a lot of sense out of seemingly harmless data. They are able to correlate many communication processes, and they are able to remember things of raised interest. Given the low cost of processing required at different stages and the cheap storage available, it is likely that a historically detailed profile[25] of all communication of an individual is created.

## III. IMPLICATIONS

The result of inexpensive Internet surveillance measures that do not require human intervention is a

collection of data for offsite analysis[26] and reaction. It is entirely possible to automatically create classification, interpretation, and reaction rules that preselect certain communication participants for more in-depth surveillance without any human interaction.

If a person shows an unusual communication pattern, perhaps at the 80th percentile, then this person becomes someone of greater interest to agencies conducting espionage. The communication patterns that are analyzed could be over months, and include online hours, contacts of 1st and 2nd and 3rd degrees, web search terms, and the interpreted content of all communication. The only thing that effectively keeps spy organizations from automatically spying on you is if your total communication profile, and the communication profile of the people in your social environment, are entirely uninteresting to them both now *and* in the future.

It is feasible and realistic to expect that Internet mass surveillance of a certain scale and reach already exists worldwide. The analytic capabilities of current technology is exceptional, and since the long-term memory is inexpensive for data of interest, it is therefore likely to exist. That means that both innocent actions and the actions of those in your social environment can trigger more in-depth surveillance in an automatic fashion. The human and technical resources required for Internet mass surveillance are not only within the reach of many parties, but they also constitute a small fraction of their available resources. If it is assumed that there is *any* motivation for mass surveillance, then all other factors aside, the economics suggest that it is performed on an astronomical scale not only by nation states and their agencies, but also by corporations. Looking at the sales data available for specialized surveillance and analysis equipment offered to the market, it is naive to assume that many bytes of communication escape surveillance.

The distinguishing matter is not if individuals are being spied on by computers (because they certainly are) but if they are also being spied on by people. Signals intelligence always has been a large portion of an intelligence agency's budget, and it is more so after the American tragedy of September 11th. International corporations that try to control information leakage, public image damage, competitive analysis, and outright espionage are also increasing their signals intelligence budget. This is especially true in times of economic turmoil, where there will be globally-heightened competitive intelligence competition. Furthermore, intelligence gathering is the bread-and-butter of many “dot com” companies that provide their services for free, such as Google, Yahoo, and MSN[27]. These companies and their offerings are ubiquitous, so the issue is not *if* or *why* they do it, but *how* you become a person of interest.

## IV. THREAT ASSESSMENT

The specific motivation to select your communication for analysis does not have to be high at all. It is an anticipated future interest and is visible in data retention and other “preventive” measures employed by governments today. The motivation can be anything interesting to an agency, including web searches about tax savings, e-mails from those with unpopular political opinions, interest into certain technological trends, the layout of your stock portfolio, the grade you achieved in your chemistry course, the position you hold in a company, and participation in a group of interest. The list of “interesting” activities is innumerable, and the more interesting your activities, the more elevated you

are as a surveillance target. In fact, anyone reading this paper, especially those reading it online for a longer time or increased frequency, would almost certainly elevate their status as a surveillance target. Staying below the radar can be extremely hard if you are in any way different from the majority of the populus.

When surveillance becomes trivial for an unrestrained party, then it will be done, and sadly, there is no good reason that they should not do it if they are unrestrained. Most of the notions against the reality of mass surveillance are based on "scarcity of resources and motivation" arguments. It has been demonstrated in this document that there is no scarcity of resources to do surveillance or store its results, only to act upon it by human resources. In our current world, there is no scarcity of motivation to do it either. In fact, there is a whole industry and even political parties lobbying on the behalf of surveillance. There are enough power-hungry people that want to stay in power and institutions that exist to self-perpetuate. Someone once said that the Internet is not only the best tool for mass communication but also the best tool for mass surveillance and control ever created. That person was right.

## V. CLANDESTINE INTELLIGENCE GATHERING

Clandestine intelligence gathering is spying performed by agencies and corporations that do not have "lawful interception"[28] privileges, lacking legal authority and legitimate access to infrastructure. This is the traditional idea of espionage, where one country or company is spying on another or a target group. The stages are similar to traditional surveillance; however, the methods used tend to be less traditional since the spying organization involved does not have conventional communications access but also is not confined by the rule of law.

Clandestine intelligence may be as insignificant as one auto dealer spying on another to gain an advantage[29], or as disturbing as a country spying on the government employees of a rival country to cripple their defense infrastructure in preparation for a future war[30].

Data collection for clandestine operations follows the path of least resistance, depending on the objective. Because clandestine data collection is not lawful, it cannot be overtly employed, but instead, it must be covertly deployed using either Open Source Intelligence (OSINT) or "covert intelligence" techniques. Open Source Intelligence gathering "involves finding, selecting, and acquiring information from publicly available sources and analyzing it to produce actionable intelligence... The term *open* refers to overt, publicly available sources"[31] as opposed to covert intelligence which refers to private, classified, or illegal sources.

One example of an Open Source Intelligence gathering source is the Tor Network. The Tor Network is an anonymity network that is participation-based and allows anyone to access communications traffic of its users; however, it also attempts to obfuscate the origins of the traffic in order to render the user anonymous. The inherent weakness of the Tor Network is that each node in the network acts like a miniature IX, routing the traffic of other users and giving easy eavesdropping access to anyone who wants to abuse it. The Tor Network provides an endless supply of interesting traffic, specifically because the users are those who wish not to be observed or identified. Because this traffic is both

suspicious and interesting, it is the natural target of surveillance by both state agencies[32] and hackers[33]. In an Open Source Intelligence gathering model, the spying organization might operate Tor nodes and perform traffic analysis to identify political dissidents[34], capture sensitive government credentials[35], and even to deanonymize[36] and correlate traffic back to reporters, bloggers, and governments agents.

Covert intelligence gathering for clandestine surveillance uses non-traditional methods to acquire communications access. These are typically Black Ops programs which employ trojans[37], bribery, blackmail[38], misdirection[39], and infiltration[40].

## VI. END NOTES

This article exclusively deals with the possibilities and methods for passive surveillance of non-participants of the communication being surveilled. There are numerous other methods of surveillance and data collection existing on the Internet. Those include cookies, spyware, log file aggregation, system fingerprinting, and many other methods.

VII. Q&A

Q: What about using word scrambling to defeat language analysis?

A: The technology used in most word processors is good enough to instantly reconstruct large portions of a scrambled text. The approaches by systems working with semantic analysis, context and subject discovery, as well as whole text probability, are even better. They might not be able to reconstruct every single word, but rather, just enough of the content to make sense of it. The same is true for most if not all "good advice" given by friends. Good security is not that easy. If advice does not include strong cryptography, it is uninformed at best, and disinformation at worst.

Q: Are encryption users more likely to become targets?

A: As mentioned in the article, one of the methods used is to find out unusual traffic and content patterns. Using e-mail encryption is something unusual for the normal population. There have been several cases where the use of encryption increased the interest of investigating agencies. However, we still think that it is a necessary and smart move to encrypt everything you can. Surely you cannot beat context analysis with encryption alone, but content analysis and interpretation can be rendered much less effective or even impossible.

The advice we would give is to encrypt all your communication every time. It is better to have a consistent communication pattern than to only encrypt occasionally because the total amount of valuable data collected will be lower. If you are only encrypting information you think is sensitive, then it is also known which communications should be more heavily analyzed.

Q: Are people using anonymity networks more likely to become targets?

A: Yes. The total number of available anonymization services is small. Just a few thousand computers in total are serving in publicly available anonymity networks. To target all traffic going to or from those computers is trivial. However, only a really big adversary would be able to automatically trace and connect the various relayed packets to each other, and those adversaries surely exist.

Looking at the network layouts of the more popular anonymization networks, it is actually not hard to watch all traffic they relay. Some services make it hard to identify single communication events when watching only a limited set of the total connections that exist; at the same time, this increases the crowding effect (hiding in the crowd). With effectively executed crowding, you will be seen but not necessarily identified.

Q: But company X said they use technology Y.  Won't that protect me from all adversaries?

A: No. It is true that technologies exist to drastically increase your privacy on the Internet. However, none of them protect you against an omnipotent attacker. Most are good for evading nosy marketing

groups, though few are good enough to hide yourself from the eyes of domestic security agencies. However, none will protect you against a motivated attacker with global access to the Internet. If your anonymization service is decent, then they will have a note in their website or documentation that effectively states, "Do not rely on this technology if you require strong anonymity." If they aren't decent, they will say, "We make you 100% anonymous on the Internet."

Q: What can be done?

A: Writing to your congressional representive will not stop spying. Politics and public opinion will not help at all to reduce or even solve this problem, because politics and public naivete created the problem. There are only seven things you can effectively do:

1. Accept that the world is *not* a place where everyone believes others should be free.

2. Use self-defense technology such as adequate anonymity services and best practices.

3. Use encryption on all your traffic, and support programs that employ opportunistic encryption. Even weak and poorly-implemented encryption is better than plaintext, because it cripples spying by reducing it to context analysis.

4. Call up your ISP and tell them you want a dynamic IP address, because static IP addresses are a threat to your privacy. If you work at an ISP, insist that it assigns IP addresses dynamically, not statically.

5. Prepend common data to the first 1k of your data transfers to defeat modern checksum analysis.

6. Fight against any force that wants you to give up your freedoms and privacy.

7. Teach others how to fight for their privacy as well.

Protecting your privacy does not come for free today, and it never has. One last word to the wise: those that shout the loudest that they will protect you or those that do it for free are not necessarily those that have your freedom and privacy in mind. There is no such thing as a free lunch!

## VIII. ABOUT THE AUTHORS

Jonathan Logan works as a communication network consultant for Cryptohippie PA Inc. and Xero Networks AG. He can be reached via email at j.logan at cryptohippie.net (PGP Key: 0xE82210E6) Steve Topletz is the operations advisor for XeroBank, an anonymity service operated by Xero Networks AG. The opinions expressed in this article are those of the authors and do not reflect the views of Cryptohippie PA Inc., Xero Networks AG, their management, or their respective owners. If you want to distribute this article, please contact the authors.

IX. EXHIBITS

Note: Figures used in calculations are designed to be rough and larger than actual costs, in order to demonstrate maximum reasonable costs.

Exhibit A: (http://www.dtc.umn.edu/mints/home.php) 5000 ~ 8000 PB / month. Presume ~85th percentile at 7500 Petabytes * 12 months = ~90 Exabytes (94,371,840,000 GB). Data warehousing costs are approximated to $0.35 / GB / year, ($0.168 / GB hardware, $0.014 / GB power, $0.091 / GB housing, $0.077 / GB maintenance; breakdown derived from classified source, traffic costs not included). 94,371,840,000 GB * $0.35 / GB = $33,030,144,000 USD / year.

Exhibit B: 1% * (94,371,840,000 GB) x $0.02 / GB fiber-optic transfer x 2 destinations (collection and endpoint) = $37,748,736 total fiber-optic transmission costs. Note that although internet traffic doubles, unique traffic does not increase at the same rate, so 1% is a shrinking figure as total traffic increases. Non-unique traffic is typically limited to personal communications such as VOIP, email, and instant messaging.

Exhibit C: IBM BladeCenter PN41, 20 Gbps @ $90,000 = $4.5k / Gbps. Similar costs across the board (90k wholesale, 106k ~ 120k retail) with other DPI / traffic analysis solutions (Narus, Sandvine, LSI, Qosmos, Interphase, Ellacoya etc).

Exhibit D: ~90 Exabytes raw analysis / 1 year = ~24 Tbps (23.36) average usage (20Tbps domestic, 4 Tbps international) @ 20% utilization = 117 Tbps (@ 100% utilization) x $4.5k Gbps = $526,500,000 USD. Hardware has a yearly cost of 48% of costs before traffic (power, housing, maintenance). Costs before traffic are $570,375,000 ($526,500,000 / 0.48 * 0.52), and traffic costs of $37,748,736 bring the total to $1,134,623,736 for all costs post-tap / pre-analysis.

Exhibit E: Maximum 5000 tapping points worldwide x $3,000,000 / tap / year for physical surveillance, compliance, black operations, tap installation, and maintenance, and upkeep costs. In Germany alone, there are 30 major backbone loops, and 10 major IXs, which require multiple taps for total surveillance.

Exhibit F: The cost of Access is $2.027b, consisting of $527m for Traffic Analysis, and $1.5b in Tap Installation and Management (Exhibit E). The cost of Storage is $570m (Exhibit D), favoring the larger cost against the 1% of $33b (Exhibit A). The cost of Traffic is $38m, and the cost of Analysis can reach as high as $1.5b. $2,027m + $570m + $38m + $1,500m = $4,135m.

X. CITATIONS


1. Olzak, Tom (2007, May 3). Protect your network against fiber hacks. Retrieved July 18, 2009, from TechRepublic Web site: http://blogs.techrepublic.com.com/security/?p=222&tag=nl.e036.

2. (2004). Map of U.S. city connectivity. Retrieved July 18, 2009, from TeleGeography Web site: 2. http://www.telegeography.com/ee/free_resources/figures/ib-04.php.

3. (2006). Submarine cable system diagram. Retrieved July 18, 2009, from TeleGeography Web site: http://www.telegeography.com/ee/free_resources/figures/ib-02.php.

4. List of Internet exchange points by size. (2009). In Wikipedia [Web]. Retrieved July 18, 2009, from http://en.wikipedia.org/wiki/List_of_Internet_exchange_points_by_size.

5. Information awareness office. (2009). In Wikipedia [Web]. Retrieved July 18, 2009, from http://en.wikipedia.org/wiki/Information_Awareness_Office.

6. Nash equilibrium. (2009). In Wikipedia [Web]. Retrieved July 18, 2009, from http://en.wikipedia.org/wiki/Nash_equilibrium.

7. Brams, S., & Kilgour, D. (1991). *Game theory and national security*. New York: Wiley-Blackwell.

8. Libbenga, Jan (2005, Nov 28). Iceland left in the cold after cable cut. The Register, Retrieved July 18, 2009, from http://www.theregister.co.uk/2005/11/28/iceland_without_broadband.

9. (2005). Navy commissions spy submarine Jimmy Carter. Retrieved July 18, 2009, from Cryptome Web site: http://eyeball-series.org/mmp/jimmy-carter.htm.

10. (2001). Ships, sensors, and weapons. Undersea Warfare, 3, Retrieved July 18, 2009, from http://www.navy.mil/navydata/cno/n87/usw/issue_11/ship_sensors_weapons.html.

11. Kent, S., & Atkinson, R. (1998). IP encapsulating security payload. Retrieved July 18, 2009, from The Internet Engineering Task Force Web site: http://tools.ietf.org/html/rfc2406.

12. Pike, J. (1996). Intelligence agency budgets. Retrieved July 18, 2009, from Federation of American Scientists Web site: http://www.fas.org/irp/commission/budget.htm.

13. (2007, May 3). HP launches DRAGON to help telecoms manage data in fight against global terrorism . Retrieved July 18, 2009, from PR Domain Web site: http://www.prdomain.com/companies/H/HP/newsreleases/20075440637.htm.

14. O'Brien, D. (2008, June 15). Sweden and the borders of the surveillance state. Retrieved July 18, 2009, from Electronic Frontier Foundation Web site: http://www.eff.org/deeplinks/2008/06/sweden-and-borders-surveillance-state.

15. (2009). NarusInsight is the most scalable traffic intelligence system for capturing, analyzing and correlating IP traffic in real time. Retrieved July 18, 2009, from Narus Web site: http://narus.com/index.php/product.

16. (2008). About BND. Retrieved July 18, 2009, from Bundesnachrichtendienst Web site: http://www.bnd.de/nn_1435078/EN/WirUeberUns/WirUeberUns__node.html.

17. Pike, J. (2009). World wide military expenditures. Retrieved July 18, 2009, from Global Security Web site: http://www.globalsecurity.org/military/world/spending.htm.

18. (2006). Directive 2006/24/EC of the European parliament and of the council. Official Journal of the European Union, 105, 54-62. Retrieved on July 18, 2009, from http://www.ispai.ie/DR%20as%20published%20OJ%2013-04-06.pdf.

19. Krempl, S. (2009, June 7). CCC: Vorratsdatenspeicherung bringt unkontrollierbare Überwachung. Heise, Retrieved July 18, 2009, from http://www.heise.de/newsticker/CCC-Vorratsdatenspeicherung-bringt-unkontrollierbare-Ueberwachung--/meldung/141623.

20. Zetter, K. (2009, June 22). WSJ: Nokia, Siemens help Iran spy on internet users. Retrieved July 18, 2009, from Wired Web site: http://www.wired.com/threatlevel/2009/06/wsj-nokia-and-siemens-help-iran-spy-on-internet-users.

21. Cheung, H. (2006, June 27). ISP heavyweights join forces to fight child porn. Retrieved July 18, 2009, from TG Daily Web site: http://www.tgdaily.com/content/view/27256/118.

22. Bundeskriminalamt. (2006). The Bundeskriminalamt Profile [Brochure]. Bad Homburg, Germany. Retrieved on July 18, 2009, from: http://www.bka.de/profil/broschueren/profile2006.pdf

23. Latent semantic analysis. (2009). In Wikipedia [Web]. Retrieved July 18, 2009, from http://en.wikipedia.org/wiki/Latent_semantic_analysis.

24. Li, J., Zheng, R., & Chen, H. (2008). From fingerprint to writeprint. University of Arizona. Retrieved from http://ai.eller.arizona.edu/COPLINK/publications/CACM_From%20Fingerprint%20to%20Writeprint.pdf.

25. Chaos Computer Clubs. (2009). Stellungnahme des Chaos Computer Clubs zur Vorratsdatenspeicherung [Report]. Germany: Kurz, C., & Rieger, F. Retrieved July 18, 2009, from http://www.ccc.de/vds/VDSfinal18.pdf.

26. Stokes, J. (2009, July 6). NSA's power- and money-sucking datacenter buildout continues. Retrieved July 18, 2009, from ARS Technica Web site: http://arstechnica.com/tech-policy/news/2009/07/r2e-nsas-power--and-money-sucking-datacenter-buildout-continues.ars.

27. (2004, Apr 13). Google's gmail could be blocked. Retrieved July 18, 2009, from BBC News Web site: http://news.bbc.co.uk/2/hi/business/3621169.stm.

28. Lawful interception. (2009). In Wikipedia [Web]. Retrieved July 18, 2009, from http://en.wikipedia.org/wiki/Lawful_interception.

29. Roth, D. (2009, Apr 23). Auto espionage: Koenigsegg dealer caught spying on competing Ferrari dealer. Retrieved July 18, 2009, from Auto Blog Web site: http://www.autoblog.com/2009/04/23/auto-espionage-aston-dealer-caught-spying-on-competing-ferrari.

30. Anderson, N. (2007, Sept 3). Pentago hacked, Chinese army suspected: Report. Retrieved July 18, 2009, from ARS Technica Web site: http://arstechnica.com/security/news/2007/09/chinese-military-accused-of-hacking-pentagon-computers.ars.

31. Open source intelligence. (2009). In Wikipedia [Web]. Retrieved July 18, 2009, from http://en.wikipedia.org/wiki/Open_Source_Intelligence.

32. Soghoian, C. (2007, Sept 16). Tor anonymity server admin arrested. Retrieved July 18, 2009, from Cnet Web site: http://news.cnet.com/8301-13739_3-9779225-46.html.

33. Lemos, R. (2007, Mar 8). Tor hack proposed to catch criminals. Retrieved July 18, 2009, from Security Focus Web site: http://www.securityfocus.com/news/11447?ref=rss.

34. (2009). Who uses Tor?. Retrieved July 18, 2009, from Tor Project Web site: http://www.torproject.org/torusers.html.en#activists.

35. Gray, P. (2007, Nov 13). The hack of the year. Retrieved July 18, 2009, from The Sydney Morning Herald Web site: http://www.smh.com.au/news/security/the-hack-of-the-year/2007/11/12/1194766589522.html.

36. Deanonymizer (2009). http://deanonymizer.com.

37. (2008, May 7). German intelligence caught spying on journalist's emails. EDRI-gram, 6, Retrieved July 18, 2009, from http://www.edri.org/edrigram/number6.9/german-intelligence-emails.

38. Davies, B. (2005). The spycraft manual: The insider's guide to espionage techniques. St. Paul, MN: Zenith Press.

39. Schneier, B. (2008, Feb 5). Fourth undersea cable failure in Middle East. Retrieved July 18, 2009, from Schneier Web site: http://www.schneier.com/blog/archives/2008/02/fourth_undersea.html.

40. Acohido, B. (2009, Apr 9). Q&A on U.S. electrical grid infiltrated by Chinese, Russian cyberspies. Retrieved July 18, 2009, from The Last Watchdog Web site: http://lastwatchdog.com/chinese-russian-cyberspies-lurk-us-electrical-grid.