Dark Side Intelligence

Executive Summary

Over the past several years, a new type of threat has emerged on the Internet – the botnet. Using "Command and Control servers" deployed in the dark corners of the Internet, new criminal enterprises have grown from their roots in the traditional computer viruses to become serious threats to enterprises and consumers alike.

These threats have emerged despite the nearly ubiquitous presence of traditional antivirus and endpoint security products. These security products have proven a poor defense against botnets and have yet to adapt to the changing threat landscape.

Several new approaches have emerged to combat this threat. Each have contributed to the fight, but until now no one solution has emerged with enough promise to warrant widescale deployment. Umbra Data's Dark Side Intelligence is a security service which combines Umbra Data's unique malware intelligence, the best of the thriving security community's research, and a unique vetting process. The result is a targeted, actionable inventory of botnet C&C servers designed to be deployed onboard an organization's existing network routers, switches, firewalls, and other control points. Once deployed, Dark Side Intelligence disrupts communication between the cybercriminal and their digital minions.

Introduction

Everyone involved in modern-day Information Technology is aware of the fact that cybersecurity is something to take seriously. Many of us are also aware that the types of threats and the motives of the perpetrators have changed over time from simple vandalism and fame into financial gain and criminal exploitation. In fact, making the news, or disrupting everyone's digital life, is no longer on the agenda of the cybercriminal.

Fewer of us understand the new technology and business behind this quiet revolution, and how it is thwarting our existing cyber defenses. The technology – which have been coined 'botnets' and 'zombies' – represent a paradigm that is intended to turn infected computers into long-term assets for the new cybercriminal. These criminals are also working in new ways, forming cabals and syndicates, and new markets and monetization strategies resembling virtual back alleys and fences for stolen goods are quickly being established. In short, the world of the virtual criminal is beginning to resemble the world of the "real" criminal.

In this paper, we will examine the changes in our threat landscape, and explore how most current security approaches don't do an adequate job of addressing these threats. Then we'll examine a new strategy that promises to disrupt the workflow of the new cybercriminal and break their cycle of infection.

That was then

In the past, the goal of the cybercriminal was simple. He wanted to prove to the world that he was smart enough to create a piece of malicious code that could spread far and wide. Sometimes, he also had nefarious goals such as vandalizing or damaging the systems that became infected.

Although the methods of infection used were diverse, we can effectively generalize how they worked. A malware author, typically working alone, crafts a threat, often in response to a pre-existing and well understood weakness in an operating system or application. They then release the threat into the wild, and it begins to spread. Sometimes the threats spread by themselves, from computer to computer, without the involvement of people. Often, however, they spread via email or the world wide web, and depended on people 'clicking' them to do their business.

Typically, once these threats were released into the wild, that was it. They spread from system to system doing their thing. Often, another aspiring malware author would modify and release a variant of the threat. Once these threats came to the attention of security vendors, they would write a signature to detect the threat. Once the signature of the threat had been distributed to all of these antivirus clients, the threat would be blocked from spreading to new systems. Thus, the cycle of infection would be broken.

This is now

Today, things have evolved on many fronts. On the technological front, we have seen the decline of the 'fire and forget' type of threats in favor of a new approach – malware that includes a 'phone home for further instructions' capability. After all, why go through the trouble of re-infecting a system with a new threat once you've successfully taken control of it once?

These new threats spread in many of the same ways as older threats, and, of course, leverage new and more creative techniques. But it is the power of the 'phone home' capability that has revolutionized the malware world. Botnets, as they have come to be called, have become invisible armies in the hands of increasingly organized cybercriminals.

The exact size of the botnet problem is difficult to determine. Estimates of the number of "zombies" or "bots" – systems infected with botnet malware – vary greatly depending on your source. Estimates range up to millions of systems. Shadowserver's estimates of bot-infected PCs over the past 3 years have been as high as 7 million systems, and seldom dropped below 1 million. [i]

Now, instead of a hit-and-run impact, malware authors have platforms from which to launch sustained criminal enterprises. This has lead to new priorities and

business models for the malware authors, giving them the ability to monetize their skills in ways that were previously unavailable.

These new business paradigms break down into a few different categories: individual proprietorships, malware service providers, and software publishers.

The individual proprietorship model is a simple extension of the way malware authors have operated for years. These malware authors work alone, set their own hours, and answer to no one. The main difference is that they now have a way to get paid. By using their malware to collect data or launch attacks (more on these later), they are able to raise money by selling this information at wholesale rates to criminals, or extort money directly from organizations or individuals.

In the service provider model, malware authors are simply adopting the increasingly popular software as a service (SaaS) concept and molding it to the dark side of the Internet. Using the same markets as the individuals, they, essentially, rent their botnets by the hour. The larger the net, the higher the price.

Lastly, many malware authors are coming to the conclusion that all of this exploitation of the Internet is just too much work. Creating malware for their own use and going to all of the trouble to create and administer the botnets themselves takes time. Instead, they are doing exactly what many software entrepreneurs are doing today – selling their botnet technology in the dark-side equivalent of the app store. There are dozens of "botnet kits" for sale today, and many of them have lowered the bar for becoming a "hacker" to a simple point-and-click interface. See figure one for a look at a botnet ecosystem.

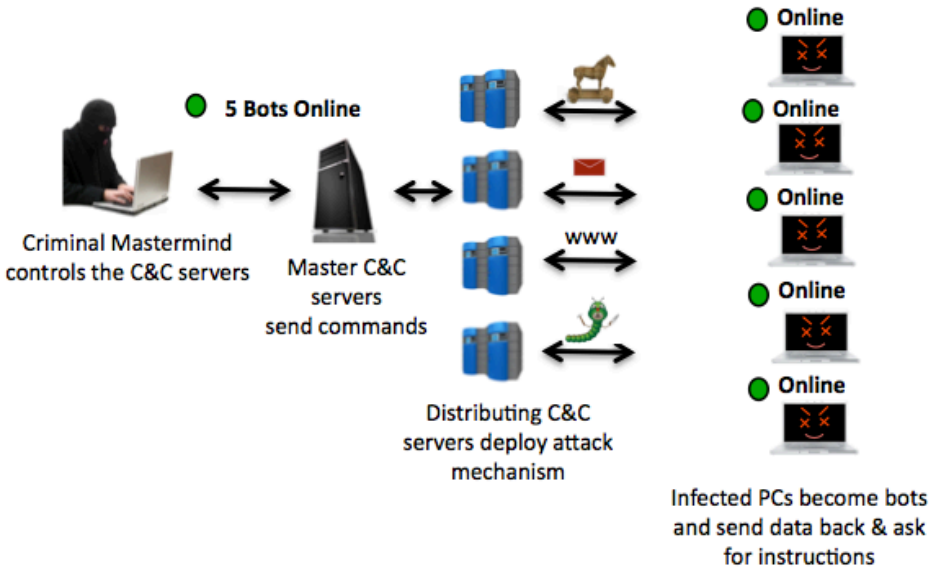

Figure 1: Botnet Command and Control (C&C) Architecture

So, what are all of these hundreds of thousands of "bots" doing? What revenue stream is supporting this branch of the underground economy? The short answer is "lots of stuff". The criminal community is constantly coming up with novel targets for exploitation in the digital world, just as they are constantly finding new niches in the physical world. However, there are certainly clear trends. Not surprisingly, these encompass spam, extortion, and the theft of personally identifiable information (PII). A new, and potentially even more dangerous, trend is also now emerging: very targeted attacks aimed at high value targets such as governments and corporations.

Spam, and its associated phishing attacks, have been around forever. Spam first appeared on the Internet scene in 1995, with phishing following a year later. However, spam has also received more attention than perhaps any other digital problem, since it has such an impact on the quality of everyone's Internet experience. As a result, it has become harder and harder for the spammers to get their messages out. Botnets represent a significant new avenue for them, and they bypass many of the Internet controls put in place to reduce the spread of spam. The top spam botnet is believed to send out over 14 billion spam messages per day.[ii] It is believed that the majority of spam is now sent via botnets[iii].

Extortion has been a recurring theme on the Internet for many years. One of the most notable extortion attempts occurred several years ago and affected online gambling sites.[iv] Although several Russian criminals were sentenced to jail time for this specific attack, this type of attack continues to this day. In 2009, an attack was launched on Techwatch using similar tactics.[v] These attacks leverage a large number of "bots" to simultaneously send huge numbers of messages (typically TCP "SYN"s) to the entity under attack. This overwhelms the target, blocks legitimate traffic, and effectively takes them off the air. These attacks are known as distributed denial of service (DDoS) attacks.

Identity theft is an even more direct way to exploit botnets. Shockingly, a complete individual identity sells for $3 to $15 in wholesale markets.[vi] Bots use keystroke loggers – which collect every key a user presses – and screen scrapers – which collect what the user sees on their screen – to quickly collect the online identity, credit card information, and banking information of their victims. Once this information is harvested via the botnet Command and Control server, criminals quickly sell this information via the online black market.

The extremely targeted use of botnets – narrowly focused attacks against certain types of systems or information – have been predicted for some time, and have only recently begun to appear. The StuxNet worm, for example, exploits a weakness in Microsoft operating systems and a Siemens industrial software package to steal industrial design information from systems it infects. This industrial espionage is disturbing enough – but the systems targeted in these attacks are "used in control

systems including industrial manufacturing, utilities and even nuclear powered aircraft carriers"[vii]. Stuxnet was really only half of a "true botnet" – it harvested confidential information and sent it to the bad guys, but lacked to ability to accept new commands.

Google has also been the subject of a variety of similar attacks directly targeting some of their users. These attacks attempted to install malware to steal information from Chinese dissidents and monitor their on-line activity. These attacks caused an international incident in early 2010. [viii]

Yesterday's security doesn't fix today's problem

So how do we solve this problem? What has the security industry been doing to address this new threat? The answer: not enough. The botnet problem has proven intractable to the current generation of security products. As a result, the security community has launched several non-profit endeavors and new commercial enterprises. Many of these have made significant contributions to solving the problem, but none have adequately addressed the core issue.

The majority of security solutions in the market today are reactive in nature. They rely on the detection of a given threat in the wild, and then the creation of a unique signature the software can use to identify and remove the threat once it appears on a "protected" system. This is a reasonable approach to the old, more static threats. However, the botnet has a huge advantage: it can upgrade and update itself on the fly. In fact, the number of unique threat signatures has risen so fast that the sustainability of this entire approach to Internet security has been called into question. In 2009, Symantec wrote nearly 3 million signatures.[ix]

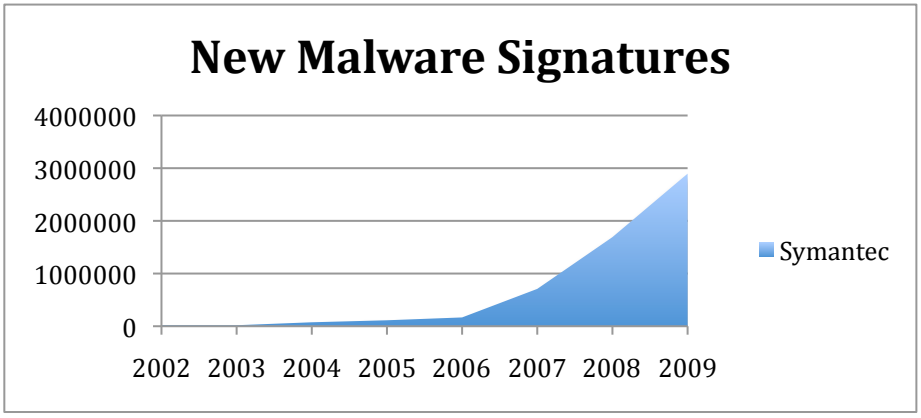

Figure 2: Growth of Malware Signatures

Furthermore, the harm posed by a zombie system can be realized by the cybercriminal almost immediately – long before a signature can be released to stop it. As soon as they are installed, bots will begin sending spam, reporting back with personal information, and launching attacks. And, just in case, many of today's

botnets, such as Conficker, will actually block access to antivirus websites to prevent your software from being updated.

The efficacy of these old-school solutions can be easily measured by looking at the rate of botnet infections over the past few years. A particularly telling example is the Zeus botnet, which is well known to collect banking information from infected computers. Although first discovered in July of 2007, even as of August of 2010, fewer than 50% of Antivirus clients were correctly detecting and removing this extremely dangerous threat. See figure 3.

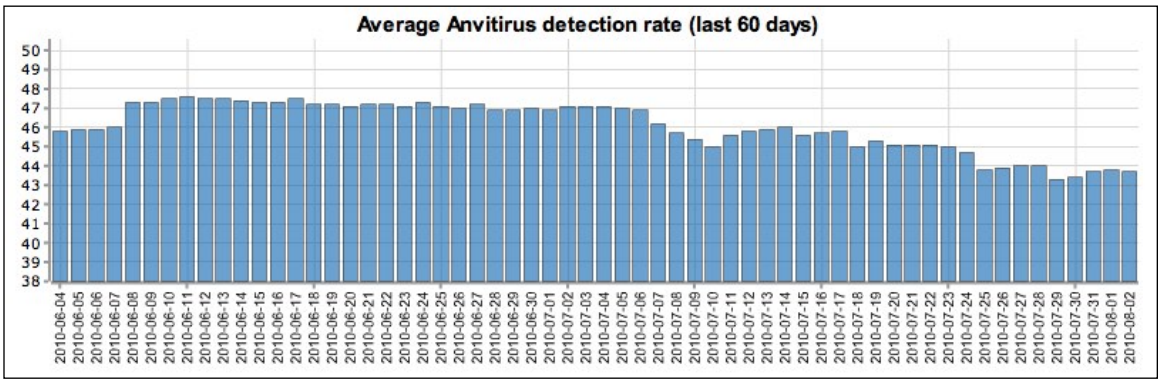

Figure 3: Effectiveness of Antivirus Against Zeus Botnets in August, 2010

First Generation Anti-Botnet Approaches

Because existing security products were not adequately addressing botnets, key members of the security community banded together to form a variety of organizations to help with the problem. These first-generation anti-botnet efforts each had slightly different goals and approaches, but they can be effectively summarized as having a three-pronged mission: raise awareness of the problem, generate intelligence about the scope and specific sources of the problem, and, if possible, hijack the botnet networks themselves.

These organizations have added tremendously to the body of knowledge about botnets, and played key roles in taking down some botnet networks. Key examples include the Conficker Working Group and the Zeus Working Group. These organizations are working to rid the world of Conficker and Zeus, respectively.

While these approaches to the botnet problem have had a dramatic impact, effectively eliminating some botnets and slowing the spread of others, this approach has not realized its potential, especially given the often high quality of the data generated by these efforts.

The reasons for this are twofold – lack of a single, central broker for the network intelligence puts too much of the burden onto the IT staff and makes implementation very difficult. Further, the lack of a consistent process for vetting

suspected C&C hosts keeps the data from being actionable for fear of creating collateral damage.

This "whack-a-mole" approach to botnets is, frankly, too challenging for IT security to take advantage of. Which list of botnet C&Cs should I worry about? One? Two? More? While the Conficker Working Group, for example, may be an excellent source of resources to help IT once they realize they have an infection, the plethora of botnet intelligence is simply too diverse to use proactively, to prevent infection. If we're going to address the problem by depriving the bad guys of their meal ticket, any security solution must be able to shorten the window in which a given botnet can operate effectively to the point where the malware authors give up on this approach.

Another key issue for both service providers and enterprise IT security is concern about the accuracy of C&C identification. How do hosts get on the list, and how do they get off? What assurance is there that the data is accurate, customers won't be blocked from legitimate sites, and enterprises will maintain connectivity with their business partners? The lack of a consistent process for vetting information across all of these information providers means that the data they generate is not actionable as a proactive defense against infection. In other words, they don't raise to a consistent standard.

Second-Generation: Commercial Alternatives

Several commercial products have emerged in the marketplace over the past couple of years aimed squarely at this problem. Although each product has taken a slightly different approach to the problem, they can be generalized as leveraging DNS, in-line traffic sniffing, or a combination of the two. Each of these techniques have merit, but none can provide a complete, deployable solution.

DNS, or Domain Name Service, is the Internet infrastructure that translates human-friendly hostnames and URLs into the IP addresses used to direct traffic on the Internet. DNS traffic can be used in several ways to either detect or block communication between C&C servers and bots.

For example, most of the time, the bots use DNS to translate a hostname into the IP address of their C&C server. This isn't always the case (they can be hard-coded into the code), but it is handy in the event that the C&C is detected and needs to move to a new server. The most straightforward DNS-based solutions simply fail to resolve the host names of already identified C&C servers.

More advanced solutions depend on known botnet behavior patterns. Botnets leverage DNS in several different ways in an effort to protect themselves. One is called "Fast Flux", and it works by quickly changing the mappings between names and IP addresses, effectively changing servers as frequently as every minute. By

sending the traffic to other hosts they have already compromised, the bots relay their information back to the C&C and make it harder to find the C&C infrastructure.

DNS servers or proxies can learn these patterns, and flag them as malicious. Traffic for malicious hosts can then be directed into a black hole, or be redirected to a web landing page for remediation. This eliminates the ability of the malware to communicate with its server to deliver your personal information or get spam to send from your computer. However, a false positive could block a critical network resource. Small changes in behavior of the threat could also render this type of detection unreliable.

However, without being in-line to see the traffic, this strategy has a critical weakness: bots can easily re-configure DNS server information on an infected host. Once this happens, the traffic will no longer be sent to the DNS proxy, and the solution is broken.

This leaves the inline solutions. They do similar analysis of DNS traffic, at the packet level, and can break the botnet communication by dropping (failing to forward) packets to and from known C&C IP addresses.

In general, these types of in-line systems are actually Intrusion Prevention System appliances with a few added 'botnet' features. As such, they suffer from all of the same issues as their brethren IPS's. Deep-packet inspection is expensive. Most software-based solutions demands high-end server hardware and can seldom handle more than 1 Gbit/second performance. Since they do not generally scale gracefully, if more than 1 Gb/sec is needed, networks become expensive and complex.

There is a newer class of hardware-assisted network devices which are capable of doing a more than adequate job of inspecting traffic at line rate. These are variously referred to as "next generation firewalls" or "hardware-based Unified Threat Management systems". Armed with Umbra Data's Dark Side Intelligence, they can be highly effective at combating the botnet threat.

Additionally, organizations can't justify the acquisition of new hardware, with all its indirect costs – rack space, power, training on another management console – in order to combat the threat in such a partial way. The market is waiting for a simple to deploy solution with a higher degree of certainty before stepping up to the plate.

Third-Generation: Dark Side Intelligence

In order to deploy a solution, it needs to be completely actionable. Networkers and security teams already have too many 'detection' systems, log messages, and other system data. It also needs to fit into the existing infrastructure of the network to reduce deployment cost and complexity.

The missing piece to the anti-botnet puzzle is an infrastructure for evaluating potential C&C servers as soon as possible after they come on line. This information can be used to build a database that can be used by existing elements of the network and security infrastructure.

Umbra Data's Dark Side Intelligence is this missing piece. Dark Side Intelligence is the first comprehensive "in the cloud" botnet defense designed to ensure 100% safe, actionable data that can be leveraged using existing networking and security products.

The Umbra Data Difference

Dark Side Intelligence accomplishes this by combining our unique, early insight into emerging botnet C&Cs with the best data from the security community. All of this data is then subjected to a careful and consistent vetting process, and carefully categorized into actionable lists of C&Cs. Figure 4 illustrates this process.

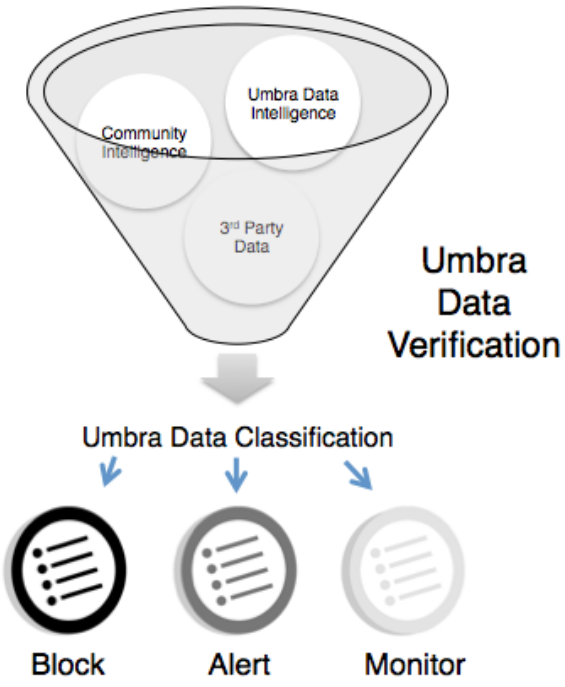

Figure 4: The Umbra Data Classification Process

By generating our own Internet threat intelligence, and combining it with the best resources in the Internet community, Umbra Data is able to build a comprehensive list of potential C&Cs. This list is then put through Umbra Data's exhaustive analysis process, which carefully evaluates each candidate. Not only does the process look for C&Cs, but also other, legitimate, services running on the same servers. Based on this process, Umbra Data produces Dark Side Intelligence. Dark Side Intelligence is

separated into three separate categories: “Block”, “Alert” and “Monitor”. It’s worth taking a detailed look at each component to understand the power of Dark Side Intelligence.

Detection

Early detection of emerging Command and Control servers is a central focus of the Dark Side approach. Umbra Data uses a variety of different techniques to accomplish this.

One of these techniques is domain registration monitoring. By keeping an eye on what domains are being registered, by whom, and where, over a long time horizon, Umbra Data is able to determine with a high degree of accuracy which Internet domains will be used for C&C hosting.

Domain registration data is then combined with other types of DNS-based intelligence to create more candidate C&Cs. This DNS data, unlike many other solutions, is not collected directly in customer networks. Instead, the data is harvested from a variety of root servers and other hubs of DNS activity. This gives a much larger sample set of data, without relying on corruptible endpoint DNS configurations. Threats from these sources are immediately added to our “Monitor” list where they receive continuous attention.

Another valuable source of data is the malware itself. By evaluating samples of bots and C&C systems that are captured in the wild, Umbra Data gains intelligence not only about the names and IP addresses of Command and Control infrastructure, but also the communication methods they use. This information is vital to the evaluation and analysis of C&C candidates, and helps keep Dark Side Intelligence actionable.

Active scanning of suspect sections of the Internet is also used to find new C&Cs. Since the Internet is extremely large, these scans must be carefully targeted if they are to yield results. Umbra Data’s algorithms identify the most likely areas of the Internet and consistently identify new C&Cs.

However, Umbra Data cannot do all the work of C&C detection alone. Many in the security community have spent years gleaning insight into the workings of C&Cs and the criminals behind them. Ignoring these valuable contributions would be ill advised. The security community cooperates to collect data on a variety of threats. Some examples of these community resources include Shadowserver, Cymru, and Zeustracker. Currently eight data sources are combined and included in the C&C candidate list, more are being added.

Each of these sources, by themselves, includes valuable information about C&C servers. However, they each have their own focus, and none of them represent the

most complete view of C&Cs available. Furthermore, many contain false positives or C&Cs that have been taken down.

Data Analysis

Therefore, all data, including Umbra Data generated intelligence, must be completely and accurately vetted in order to make it actionable. This is done via a large, anonymous network, preventing the C&C from detecting the fact that they are being scanned.

Using data gleaned from our own and others' analysis of bots in the wild, a large variety of different techniques designed to mimic botnet behavior are used to evaluate suspected C&Cs. In simple terms, this means that we speak to the C&C servers in the same way that their bots do, ensuring there is no uncertainty about the fact that the target is, in fact, malicious.

If a threat is deemed to be malicious, additional investigation is performed to determine if the IP address also hosts legitimate services such as public web sites or IRC servers. Criminals often hack into legitimate servers to make it harder to block their traffic.

Once data is collected about a target, Umbra Data uses a reputation-driven formula to assign scores to each IP. Factors such as completeness of our information, confidence in our data, the last time a system was scanned, and other factors are included in the score.

The Result: Dark Side Intelligence

Based on the score of each target, it may be placed on one of the Dark Side Intelligence lists. These lists are the core of Dark Side Intelligence, and key to delivering on the promise of always actionable botnet prevention.

Completely malicious C&C servers end up on the "Block" list. This list is 100% guaranteed to contain only hosts dedicated to malicious behavior. This means that none of the hosts on this list have any redeeming value to your business or your customers, and can be blocked without any collateral damage. Umbra Data works continuously to ensure that this list remains 100% accurate.

The "Alert" list also contains 100% vetted C&Cs. However, unlike the "Block" list, hosts on the "Alert" list are also running other non-malicious services. Since these services may be legitimate, completely blocking access to these IPs could result in unintended side effects. The majority of hosts on this list, to date, have been hosting Internet Relay Chat (IRC) channels or some sort of web traffic. In general, we believe that it is safe for enterprises to completely block all hosts on this list, since real, high value, hosts will move quickly to rid themselves of the botnet infection. This decision may be harder for the service provider.

In order to make the lists as actionable as possible, Dark Side Intelligence contains far more than a simple list of IP addresses. It also includes all the relevant information an organization would need to block specific traffic while leaving production services running, e.g. limit blocking to just the ports and applications for botnet traffic. Using this information, the "Alert" list can be kept actionable without blocking the legitimate services running on these IPs. The most straightforward technique would be to utilize the port and application level information in an IDP or other content aware device.

The "Monitor" list contains IP addresses that Umbra Data believes are very likely to turn up as C&Cs in the near future. This determination is made using a variety of our proprietary analysis techniques. No immediate action is recommended for these potential threats, unless an organization wants to input these into their own threat evaluation systems. Once Umbra Data detects malicious activity, the IP will quickly land on the appropriate action list.

In order to keep the lists actionable, Umbra Data updates Dark Side Intelligence every hour, and can even provide near-realtime feeds upon request. In general, frequent updates are extremely helpful in protecting against emerging threats. Our advice is to take feeds as often as practical for your organization's best practices. We offer Dark Side Intelligence in formats designed to be digested by today's network management infrastructure tools, in order to make the process as simple as possible.

Implementation

Knowing where the bad guys are is a good thing, but how does that help keep the Internet safe? The short answer is that breaking the communication layer between the bots and their criminal masters breaks their entire business model. Unless they can collect the personal information, send the spam, make good on their extortion threats, etc., then there is no reason for them to keep doing what they are doing.

This is the core of Dark Side Intelligence: breaking the connection between the bad guys and their paycheck. So how does that work in your network? The short answer is that Umbra Data provides you with the information you need to make your existing infrastructure smart enough to defend against the threat.

Service provider and enterprise networks are already populated with a multitude of devices capable of consuming Dark Side Intelligence. At the most basic level, all that is needed to start the bad guys working on their resume is any device capable of filtering Internet traffic at layer 3 – and nearly every router, firewall, and switch deployed in networks these days is able to do this.

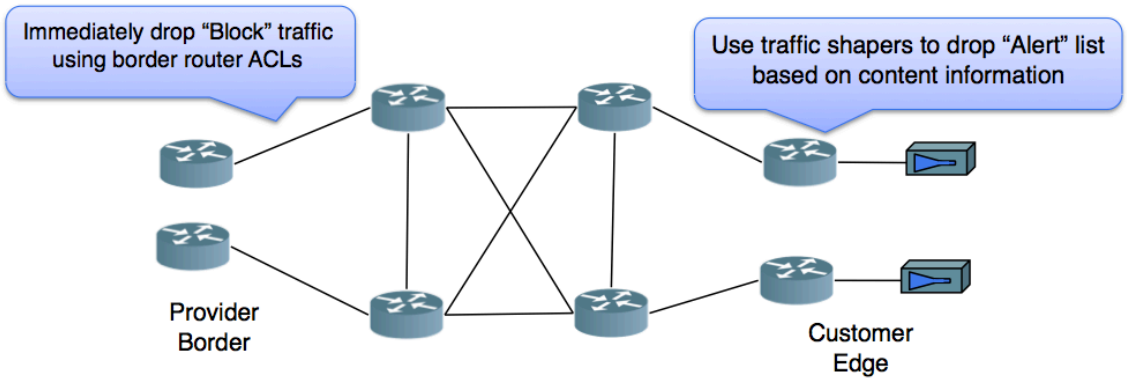

Figure 5: Example Deployment for Service Providers

To fully leverage the protection offered by the "Alert" list, organizations will need to be able to leverage the capabilities of some sort of content inspection network device. Today, these devices are extremely common throughout both enterprises and service providers. These devices provide networkers with the ability to block, for example, malicious IRC traffic while permitting 'legitimate' traffic to pass by unaffected, even if they are using the same server.

See figures 5 and 6 for examples of where Dark Side Intelligence can fit into service provider and enterprise networks. These are only a few of the ways that Dark Side Intelligence can be deployed in existing infrastructure.

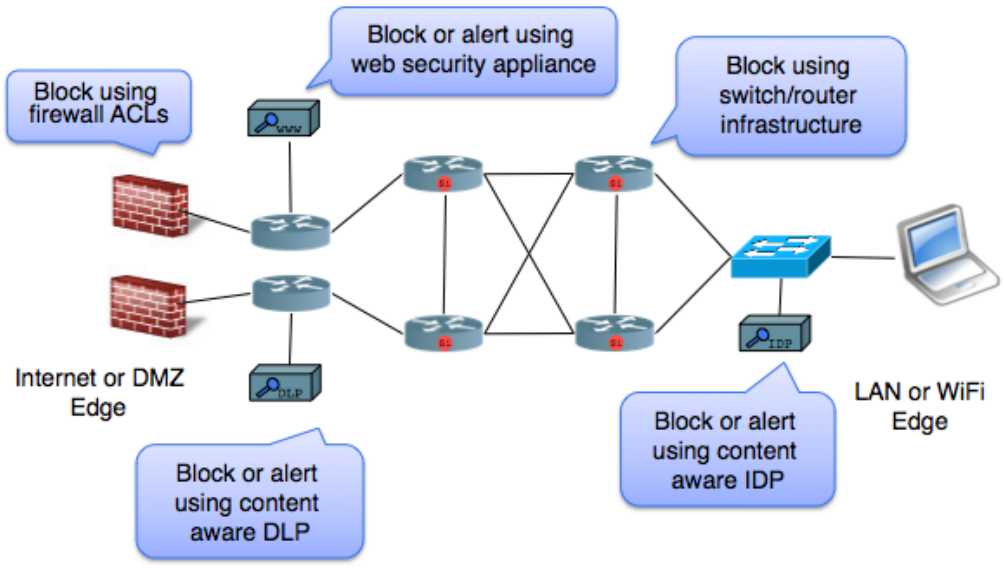

Figure 6: Example Enterprise Deployments

In order to make the implementation of Dark Side Intelligence as simple as possible, Umbra Data provides the information in many different "precompiled" formats for use in a large number of already installed network devices. See table 1 for a snapshot of the ever-growing list of formats in which Umbra Data can provide any of our lists.

Table 1: Dark Side Intelligence Data Feeds

Device Type	Vendors
Routers and Switches	Cisco, Juniper, Brocade
IPS	SNORT
DNS Appliances	DNS Defender
Firewalls	SonicWall, Cisco, Juniper
Any custom application	XML data feed

Conclusion

The threat posed by botnets cannot be emphasized enough. Botnets have become pervasive in today's Internet, and are as realistic a threat to the worlds of ecommerce and social networking as any threat yet seen online.

Yesterday's solutions to Internet threats are not well equipped to handle this new attack on our online lives. Furthermore, most current solutions to the problem represent more of a start than a finish to the threat.

This leaves Umbra Data's Dark Side Intelligence as a logical entry point for service providers who are interested in protecting their customers, and enterprises who are worried about their confidential intellectual property, to embark on the journey of attacking and defeating the latest, but not the last, attack on our online economy.

To learn more about botnets, including an interactive view of where C&C servers are active, visit http://www.umbradata.com.

[i] http://www.shadowserver.org/wiki/pmwiki.php/Stats/BotCount36-Months
[ii] http://www.allspammedup.com/2010/05/top-5-Botnets/
[iii] http://www.m86security.com/labs/bot_statistics.asp
[iv] http://www.sophos.com/pressoffice/news/articles/2006/10/extort-ddos-blackmail.html
[v] http://www.theregister.co.uk/2009/01/30/techwatch_ddos/
[vi] Symantec Intelligence Quarterly January-March 2010, page 4
[vii] http://news.cnet.com/8301-27080_3-20011159-245.html?tag=topImage1
[viii] http://www.pcworld.com/businesscenter/article/194557/report_google_attack_targeted_gaia_password_system.html
[ix] Symantec Global Internet Security Report 2009, page 48