Featured in this issue:

Distributed denial of service attacks – holding back the flood

Distributed Denial of Service (DDoS) attacks have become more sophisticated and capable of carrying out several functions at once. The frequency of attacks has also increased significantly.

But what is the best way to respond to attacks of this nature – should DDoS solutions be deployed on-premises or in the cloud? Dave Larson of Corero Network Security looks at updated approaches to mitigating DDoS attacks, examining the technology available today and how it’s being deployed, before taking a real world example of how this technology can be used in a contemporary business environment.

Full story on page 5…

Detecting and destroying botnets

Botnets are among the most powerful components of a modern cyber-criminal’s arsenal of attack techniques, leveraging the processing power of a multitude of endpoints.

Garrett Gross of AlienVault looks at the history of botnets and what they have evolved into today. And he outlines how security practitioners can detect and defeat bots operating within their own networks through both static analysis and by analysing the behaviour of systems.

Full story on page 7…

Anonymity networks and the fragile cyber ecosystem

The necessity to safeguard privacy and anonymity has never been more urgent, which is why we’ve seen the emergence of privacy enhancing technologies, including TOR, i2p and Freenet.

But how effective are they? Hamish Haughey, Gregory Epiphaniou and Haider M Al-Khateeb analysed 21 recent articles to categorise whether the main emphasis in current research is on attack or defence, what kind of resources the attacker must have, what position the attacker must hold with respect to the victim, and which part of the technology is targeted in order to degrade anonymity.

Full story on page 10…

Ransomware expands, attacks hospitals and local authorities, and moves to new platforms

Ransomware is quickly becoming more prevalent and is moving beyond infecting individual users’ machines. The malware, which encrypts a victim’s files before demanding a ransom, usually payable in bitcoins, is increasingly targeting websites and enterprises.

A new strain of ransomware, dubbed ‘Locky’, is spreading like wildfire, according to security experts. Its typical attack vector is via Microsoft Word attachments in email. Victims are lured into enabling macros in the documents which then download the main malware payload.

Continued on page 2…

Contents

NEWS
Ransomware expands, attacks hospitals and local authorities, and moves to new platforms 1

FEATURES
Distributed denial of service attacks – holding back the flood 5
Today’s Distributed Denial of Service (DDoS) attacks are almost unrecognisable from the basic volumetric attacks of old. And the frequency of attacks has also increased. But what is the best way to respond? Dave Larson of Corero Network Security looks at updated approaches to mitigating DDoS attacks, examining the technology available today and how it’s being deployed.

Detecting and destroying botnets 7
Due to their limitless size and capacity, botnets are among the most powerful components of a modern cyber-criminal’s arsenal. They are made up of compromised workstations distributed over the public internet that leverage the untapped processing power of a multitude of endpoints. Garrett Gross of AlienVault looks at the history of botnets and outlines how security practitioners can detect and defeat bots operating within their own networks.

Anonymity networks and the fragile cyber ecosystem 10
Privacy enhancing technologies, including TOR, i2p and Freenet have emerged to meet users’ needs for anonymity and privacy. But how effective are they? Hamish Haughey, Gregory Epiphaniou and Haider M Al-Khateeb analyse the technical literature to see where the emphasis is being placed in research – on attack or defence. They also look at what kind of resources the attacker must have, what position the attacker must hold with respect to the victim, and which part of the technology is targeted in order to degrade anonymity.

A new month, a new data breach 18
The wave of data breaches washing over enterprises around the world has surfaced an essential truth about IT security – typical enterprise security architectures are vastly outmoded in terms of how modern applications are shared, managed and accessed by users. Paul German of Certes Networks warns against repeating old mistakes and says it is time to think differently about security and embrace cryptographic user-to-application segmentation.

REGULARS
News in brief 3
Reviews 4
Events 20

Come and visit us at www.networksecuritynewsletter.com

ISSN 1353-4858/16 © 2016 Elsevier Ltd. All rights reserved
This journal and the individual contributions contained in it are protected under copyright by Elsevier Ltd, and the following terms and conditions apply to their use:

Photocopying
Single photocopies of single articles may be made for personal use as allowed by national copyright laws. Permission of the publisher and payment of a fee is required for all other photocopying, including multiple or systematic copying for advertising or promotional purposes, resale, and all forms of document delivery. Special rates are available for educational institutions that wish to make photocopies for non-profit educational classroom use.

Network Security
ISSN 1353-4858 March 2016
www.networksecuritynewsletter.com
This in turn encrypts the victim’s files, giving them the file extension .locky, before demanding a ransom of 0.5 to 1 bitcoins ($200-$400). The malware also removes any Volume Snapshot Service (VSS) files, making restoration impossible through shadow copies of files. More than half of the victims seen so far are in the US. At the time of writing, Palo Alto Networks said it had seen more than 446,000 sessions for this malware.

After being hit by ransomware, a hospital in Los Angeles was unable to access patient records, X-rays and other critical data and staff were left having to make notes by hand. According to some reports, a ransom of $3.6m was demanded of the Hollywood Presbyterian Medical Centre, but the organisation’s CEO, Allen Stefanek issued a statement saying that 40 bitcoins ($17,000) had been paid. The FBI is investigating.

Two hospitals in Germany were similarly affected. According to Deutsche Welle, the Lukas Hospital in Neuss and the Klinikum Aarnsberg hospital in North Rhine-Westphalia were both attacked. At the Lukas Hospital, admins responded to pop-up warnings by quickly shutting down the networks, preventing further spread of the malware. The hospital restored its systems from back-ups, but a few hours’ worth of data, including emails, was lost. At the Klinikum Aarnsberg, similarly prompt action confined the issue to a single server that was also restored from back-ups.

In the UK, North Dorset District Council said it had been hit but was refusing to pay the ransom. Around 6,000 files had been encrypted, but the council claimed to have fixed the problem, although it didn’t provide details. This follows a similar attack on Lincolnshire County Council a month earlier which left workers without the use of computers for a significant period.

After a couple of false starts, ransomware has now arrived properly on Apple’s OS X platform. Attackers managed to infiltrate the server hosting downloads of the popular bittorrent client Transmission. They replaced the genuine software with a trojanised version that included ransomware. This encrypts the user’s files and demands a ransom of one bitcoin (around $400). It also attempts to encrypt Time Machine back-up files.

This went unnoticed for around 24 hours, during which time it was downloaded an unknown number of times. The malware has been dubbed KeRanger and appears to be a modified version of the Linux Encoder trojan, says security firm Bitdefender. The infected version of the software was signed with a legitimate developer certificate, issued to someone in Turkey, and so was able to bypass OS X’s Gatekeeper protection. The certificate has been revoked by Apple. Anyone who downloaded the tainted version, 2.90, is likely to have been compromised by now. Nonetheless, the developers of Transmission suggest users should upgrade.

A strain of ransomware known as CTB Locker is targeting WordPress-based websites, rather than Windows PCs. Researchers who discovered the malware warned that it had already infected hundreds of sites. The malware operates via a PHP script. Meanwhile, the Internet Storm Centre (ISC) says that a cybercrime group is using iframe injection techniques on Joomla-based sites that previously had only turned up on sites running WordPress. The injected code is a way of inserting exploit kits on targeted servers – initially the Nuclear exploit kit but more recently Angler.

Ransomware is now the chief security threat facing enterprises, according to Trend Micro. In the last quarter of 2015, 83% of data extortion attempts used ransomware, rather than the threat of Distributed Denial of Service (DDoS) using botnets. Trend found that nearly a third (31%) of those attacks involved the CryptoWall ransomware, making it the leader in the field. Trend’s report can be found here: http://bit.ly/1LTPIwx.

An online poll by Bitdefender found that 44% of UK victims of ransomware have paid to recover their data. And nearly a third of the country’s computer users would be prepared to pay up to £400 to get their files back. Some 18% put recovering personal documents as the most important reason, closely followed by personal photographs (17%), with work documents trailing behind at 10%. The full report is available here: http://bit.ly/1QPeivo.
New flaw drowns SSL

Another flaw has been found (and patched) in OpenSSL. Like other high-profile, high-severity bugs before it, this one has been given a catchy name – Drown ( Decrypting RSA with Obsolete and Weakened eNcryption). The flaw enables a cross-protocol attack in which bugs in SSLv2 can be used to compromise a TLS connection. While many website operators might imagine they were immune from SSL-based attacks, because they were implementing only TLS, it turns out that OpenSSL remained vulnerable. Even with OpenSSL configured to disable SSLv2 ciphersuites, it would still respond to them. Exploiting the flaw is non-trivial – the attacker needs a privileged position on the network – but it is within the reach of malicious actors with only modest resources. “The OpenSSL team estimate that 11% of servers with browser trusted certificates are directly exploitable, a further 11% could be exploited due to secondary services like SMTP” explained Gavin Millard, technical director EMEA at Tenable Network Security. “It would be a worthwhile exercise to identify any service running SSLv2 or v3 and remove support for the outdated versions.” However, as past experiences with vulnerabilities such as Pooldown have shown, while a patch is available it may take some time before all systems are upgraded, potentially leaving a large number of web-facing servers vulnerable. So far, the response by website operators has been slow: one week after Heartbleed was announced, the number of vulnerable sites fell by 93% – with Drown, the figure is 5%, according to Skyhigh Networks’ Cloud Security Labs. The OpenSSL advisory is here: http://bit.ly/1RaTmo1.

Diffie and Hellman awarded

Whitfield Diffie and Martin Hellman, whose innovative work on cryptography underlies much of the encryption used today, have received the annual Turing Award from the Association for Computing Machinery (ACM). Their paper, ‘New Directions in Cryptography’, published in 1976, provided the basis for public key encryption and specifically public key exchange – still known as the Diffie-Hellman exchange – which not only enabled the kind of encryption used in securing websites and remote computer connections but also threatened to get the pair thrown in jail: the US Government treated encryption as a weapon of war under the International Traffic in Arms Regulations and saw public key encryption as a national security issue. Their 1976 paper is available here: http://stanford.io/1RbG90l.

Are exploit kits doomed?

Security firm F-Secure believes that cybercriminals are going to have an increasingly tough time producing exploit kits. Overwhelmingly, these kits – which are used to deliver malware – are dependent on victims running out of date software with known vulnerabilities. And a favourite platform to attack has been Adobe’s Flash, which has had far more than its fair share of flaws. But with even Adobe now deprecating the technology and leading browsers, including Chrome, Firefox and Microsoft Edge removing support for it, exploit kit builders have less to choose from when it comes to exploitable vulnerabilities. In addition, operating systems – even Windows – have become more secure. Since 2015, we’ve already seen a move back to macro-based exploits delivered as email attachments – something that hadn’t been common since the early 2000s, says F-Secure. There’s more information in the company’s ‘Threat Report 2015’, available here: https://secure.f-secure.com/threat_report.

Ofcom breached

UK media regulator Ofcom is investigating a breach of its systems after a former employee offered a massive data cache to his new employer, described as a major broadcaster. According to a report in The Guardian, the data could include as much as six years’ worth of information provided to Ofcom by the organisations it oversees.

Cancer clinic breached

A US cancer clinic, 21st Century Oncology, has been hacked and as many as 2.2 million patient and employee records may have been compromised. The breach occurred in October 2015 and was reported to the organisation by the FBI the following month. But the FBI asked the firm not to go public before it could investigate further. In a statement, 21st Century said it had no evidence the data had been misused, but it is offering customers a year’s worth of free credit monitoring. This is the second time that the organisation has been breached: in 2013, the FBI alerted it to an insider breach linked to a tax refund fraud scheme.

Anti-DDoS firm hacked

Staminus Communications, a California-based company that offers anti-DDoS services to websites, was taken offline by hackers, with its services being down for more than 20 hours. During that time, someone claiming to be one of the attackers posted an online ‘e-zine’ that claimed they had been able to take control of the company’s routers and reset them. They included links to what may be files containing the firm’s customer credentials, support tickets, credit card numbers and other sensitive data. The e-zine also criticised Staminus by offering a number of sarcastic “Tips when running a security company”, such as: “Use one root password for all the boxes”, “Never patch, upgrade the stack and “Store full credit card info in plaintext”.

Cybercrime seen as major threat

In the UK, 88% of people see cybercrime as being as big a threat as offline crime, while 76% believe governments and large organisations are not doing enough to combat hackers and cybercrime, according to new research by security firm Vest. In addition, 93% of people think online crime is getting considerably worse over the next 10 years. When asked what measures they take to avert the threat of computer viruses, hacking or cybercrime, 83% said they use anti-virus software. Only 19% regularly change their online passwords or use a random password generator. And 7% of Internet users in the UK take no personal security measures whatsoever online, while 10% avoid banking and shopping online completely due to fear of cybercrime. Meanwhile, 26% have been confronted with an attempted online scam in the past five years. Over the same period, 8% of people have actually suffered financial losses in an online scam, just 3% less than the 11% who have suffered financial losses in an offline or real-life crime. Another alarming statistic to emerge from the study was that 16% of the adult UK population – approximately 7.84 million people – have been the victim of a virus or hack in the past five years in which email or personal data was compromised. Asked what most concerns them about online crime 37% responded ‘potential financial losses, while for 30% the main worry is ‘potential theft of personal data’, followed by 14% who fear for their family’s safety.

Snapchat hands over personnel data

Snapchat, the online service that allows users to share photos temporarily, has suffered a data breach after a human resources employee fell for a phishing attack. The worker responded to an email that appeared to come from the head of the company and sent files containing information about “some current and former employees” according to a statement. But the firm went on to say that, “we responded swiftly and aggressively. Within four hours of this incident, we confirmed that the phishing attack was an isolated incident and reported it to the FBI. We began sorting through which employees – current and past – may have been affected. And we have since contacted the affected employees and have offered them two years of free identity-theft insurance and monitoring.”
Automating Open Source Intelligence
Edited by Robert Layton & Paul Watters.
Published by Syngress.
ISBN: 9780128029169. Price: €28.95,
222pgs, paperback and e-book editions.
Penetration testers have long
known about the value of open
source intelligence (OSINT). During
the early stages of a penetration test
– or a malicious hack, for that matter
– the attacker builds a detailed picture
of the target, collating everything
they can from publicly available informa-
tion before ever sending a packet
in the target’s direction.

Typical sources for this OSINT will include
the organisation’s website and social network
pages, as well as the social networks of any
known or high-value employees. Then there
are newspaper and magazine articles, online
forums and newsgroups, analyst reports …
the list goes on. The beauty of the Internet is
that so much of this data is readily available
now, if only you know where to look.

It’s the looking that’s the hard part. It can
be very time-consuming to trawl through
hundreds or thousands of sources, amassing,
collating and filtering the information. This
is a clear case for automation.

This collection of essays is primarily aimed
at helping you understand how to develop
algorithms that carry out the key functions
of OSINT – analysis of what’s out there and
linking records to build domain knowledge
and then presenting that as actionable intel-
ligence. These OSINT algorithms will make
great use of the many APIs now available.
But mapping, linking and reducing the data
is as important as finding and fetching it.

It’s not just hackers – whatever the colour
of their hats – who can make use of
OSINT. Companies can use it to build a
deeper understanding of markets that is
constantly refreshed, for example. And a
new breed of ‘data journalist’ has emerged
who can read the significance hidden in
aggregated information.

That said, of the 10 chapters in the book,
three have direct relevance to information
security, and are primarily about using
OSINT to improve defences. These are
the chapters on cyber-attack attribution,
enhancing privacy to defeat open source
intelligence and preventing data exfiltration.

Whether attacking or defending, this book will help improve your understand-
ing of the capabilities of OSINT as well
as developing the overall algorithms and
concepts for exploiting it. And while the
tone and content is primarily academic, the
practical applications are self-evident.

There’s more information available here:

– SM-D

Infosec Management Fundamentals
Henry Dalziel. Published by Syngress.
ISBN: 9780128041727. Price: €28.95,
72pgs, paperback and e-book editions.

The ever-increasing requirement
for information security skills and
capabilities within organisations,
combined with the current shortage of
people with the right skills and experi-
ence, means that there’s going to be a
significant number of people entering
this domain over the coming years.

These people are likely to consist of both
new graduates – some from specialised infos-
ce courses, but also many with more general
technology qualifications – as well as people
moving over from other areas of IT. What all
of them will need is a good grasp of the most
critical concepts and techniques, particularly
when it comes to management issues such as
security policies, the human resources angles,
the acquisition of equipment and services
and incident management.

In addition, there are people already work-
ing in this field who may lack formal training,
having developed their skills and achieved their
positions over a course of years. Information
security has changed a lot over the past few
years, not least with the emergence of ISO and
other standards. Various best practices have
evolved and compliance requirements have
become ever more complex.

This book is effectively a foundation
course in best practice. Built around the
ISO/IEC 27000 series of standards, it pre-
sents a framework for building an informa-
tion security management system (ISMS).
It will help you develop controls covering
areas such as policy, security organisation,
human resources, asset management, access
control, cryptography and security opera-
tions. It also includes sections on physical
and environmental security management,
systems acquisition, development and main-
tenance, managing suppliers and compli-
ance. And for when bad things happen, it
deals with business continuity and security
incident management.

That’s a lot of ground to cover in 72 pages.
So, as you might expect, this is not a book
packed with solutions. It is more about get-
ting you to ask the right questions and to
understand the full range of what’s involved
in keeping an organisation secure. Each
chapter starts with a brief list of questions to
ask – for example, in the chapter on access
management it leads off with “Does everyone
have access to what they need in order to do
their jobs?” and “Can unmanaged devices
attach to our network?”. It then presents
a list of controls you need to put in place, such
as the principle of least privilege, a central-
ised user directory, access reviews and more.

What you have here, then, is a detailed
‘to do’ list for your first day on the job as
a security manager. Even if you’re not in a
position to effect all the controls listed in
the book, you at least know enough from
reading this to spot where weaknesses may
lie. And that suggests another potential
group of readers – senior non-IT managers
and C-level executives, many (some would
say most) of whom still have no idea how
far the domain of information security
extends. For example, the idea that purchas-
ing and hiring might have security implica-
tions simply won’t occur to most managers.

The fact that this book is so concise and
designed to be such an easy, non-technical
read makes it a useful tool for raising security
consciousness within an organisation. It’s
most likely to be bought by people who are,
or are soon to become, security specialists.
But once they’ve read it, they could do them-
selves and their organisations a favour by
passing it around all the senior managers.

There’s more information available here:
http://bit.ly/225Qu0H.

– SM-D
Distributed denial of service attacks – holding back the flood

Dave Larson, Corero Network Security

Today’s Distributed Denial of Service (DDoS) attacks are almost unrecognisable from the basic volumetric attacks of old. In their early days the aim was simply to deny service to the Internet by bombarding it with traffic, designed to overload a server and render it out of action. But today’s attacks have evolved to become more sophisticated and capable of carrying out several functions at once – they might bring a firewall down with a volumetric attack and then infiltrate the network to steal sensitive data while IT teams are distracted.

The frequency of attacks has also increased significantly – we’ve seen DDoS attacks grow by a third in recent quarters, with operators experiencing an average of 4.5 attacks every day.¹ But what is the best way to respond to attacks of this nature – should DDoS solutions be deployed on-premises or in the cloud? Here we look at updated approaches to mitigating DDoS attacks; we will examine the technology available today and how it’s being deployed, before looking at a real world example of how this technology can be used in a contemporary business environment.

A brief history

In the early days of DDoS attacks (c.2000), DDoS mitigation technology focused on the ability to observe when a DDoS attack was occurring, by sampling network statistics from edge routers by interrogating NetFlow. As a result, an operator could spot when the attack was in progress – but had little means of blocking the attack. Without any true solutions available or in place, a network operator would first interpret that an attack was in progress, then manually inject a null-route – sometimes referred to as a black-hole route – into the routers at the edge of the service provider’s network, and block the attack. This null-route effectively blocked all attack traffic headed toward the intended victim.

This approach had some seriously negative implications. For example, null-route injections also blocked all good traffic along with the bad. So the use of the null routes rendered the victim completely offline – which was exactly what the attacker wanted. This approach provided a way of at least blunting the flow of the attack and served as a tool to eliminate the collateral damage to other customers or infrastructure as a result of the DDoS attack, but it did nothing to help the intended victim.

Fast forward several years and we find improvements to DDoS mitigation, and an evolution in protection techniques available to operators. It became clear that null-routing was not a sustainable approach and so the landscape evolved and new techniques emerged. Instead of injecting a null-route when an operator observed a large spike, they were now able to inject a new route instead. By implementing a new route, operators could now gain the ability to redirect all traffic through an appliance or bank of appliances that inspected traffic and attempted to remove the DDoS attack traffic from the good user flows. This approach spawned the existence of DDoS scrubbing-centres and DDoS scrubbing-lanes that are commonly deployed today.

The rise in DDoS attacks over successive quarters, to Q2 2015. Source: Corero.
Human intervention

This DDoS scrubbing approach, while a significant improvement, still relied far too heavily on human intervention. A DDoS attack would have to be detected (again by analysing NetFlow records) then an operator would have to determine the victim’s destination IP address(es). Once the victim was identified, a Border Gateway Protocol (BGP) route update would take place to inject a new route to redirect or ‘swing’ the victim’s incoming traffic to a scrubbing lane. Human analysts would then configure countermeasures on appliances in the scrubbing lane in an attempt to remove the DDoS traffic from the good traffic and forward it to the downstream customer.

In order to forward the good traffic back to the original destination, in most cases an operator would also have to create a Generic Routing Encapsulation (GRE) tunnel from the scrubbing lane back to the customer’s border router. The use of the scrubbing approach was a huge improvement over the previous tactic of ‘blackholing’ traffic; however it also came with several trade-offs – namely, the carrier network topology was made significantly more complex and fragile, and in order to be properly executed, an organisation needed dedicated security personnel to manage the system, which increased operational costs.

Looking at what DDoS attacks have developed into today, we can see a significantly larger problem as attacks have increased in size, frequency and sophistication. The modern attack surface has become enormous thanks to the successes and growth of network operators, with multiple entry points and significant aggregate bandwidth enabling many more damaging and disruptive DDoS attacks than ever before. Developments of this kind are now driving the need for an even more sophisticated approach to DDoS mitigation technology at the Internet transit and peering points are now possible and offer much needed relief from the frequent and damaging attacks that providers are dealing with on a regular basis. Additionally, for providers that prefer a scrubbing-lane approach, modern DDoS mitigation solutions offer enhanced visibility into traffic patterns as well as the ability to scale the scrubbing operation for increased bandwidth, while reducing both capital equipment and operational expenses.

DDoS defences today

New threats from DDoS methods showed us that our current mitigation strategies were too slow to react, too costly to maintain and not adaptable enough to handle shifting and progressive threats – it was clear that modern responses needed to be always-on and instantly reactive. It’s also imperative that they be adaptable and scalable so that defences can be quickly and affordably updated to respond to the future face of DDoS threats – whatever those may be.

“The desirability of these tools is due to the fact that they can be constantly on, with no need for human oversight, and they provide non-stop threat visibility and network forensics”

The increasingly popular method of fulfilling these aims is with inline protection, coupled with dynamic mitigation bandwidth licensing. With this technique, an inline DDoS mitigation engine is employed but the operator pays for only the bandwidth of attacks actually mitigated. The benefit of this approach is that it delivers full edge protection for locations in the network that are most affected by DDoS, at a fraction of the cost of traditional scrubbing centre solutions. The desirability of these tools is due to the fact that they can be constantly on, with no need for human oversight, and they provide non-stop threat visibility and network forensics.

Another aspect of effective DDoS mitigation is security event reporting. One of the Achilles heels of traditional DDoS scrubbing centre solutions is that they rely on coarse sampling of flows at the edge of the network to determine when an attack is taking place. DDoS attackers are well aware of the shortcomings of this approach and have modified many of their techniques to keep their attacks below the detection threshold, keeping them ‘under the radar’ and able to operate without being redirected to a scrubbing centre. Your security posture will only be as good as your ability to visualise the security events in your environment, and a solution that relies on coarse sampling will be unable to even detect, let alone act on, the vast majority of the modern DDoS attack landscape, since modern DDoS attack vectors have been adapted to avoid setting off the alarms that would alert a security team to their activities. A robust modern DDoS solution will provide both instantaneous visibility into DDoS events as well as long-term trend analysis to identify adaptations in the DDoS landscape and deliver corresponding proactive detection and mitigation techniques.

New architectural advances in both software and hardware are making real-time DDoS responses possible, because they leverage advanced analytics, which can examine more data and trends to detect attacks far before a system’s critical threshold is reached.

DDoS mitigation has come a long way from ‘blackholing’ traffic or redirecting to best-effort, slow-responding scrubbing centres, as the mechanisms now exist to filter the good traffic from the bad instantaneously. Organisations can now quickly and economically mitigate the threat – either by acquiring a dedicated solution for themselves or asking their service providers to implement automatic, inline DDoS mitigation for them. In either case, they now have an option that can stop these attacks and prevent the costly downtime that they incur.

Defence in action: Jagex Gaming Studios

A common area targeted by DDoS attacks is the gaming sector, since they
require constant and uninterrupted network availability to maintain a good user experience for their players. DDoS attacks lead to system slowdown or in some cases outages in service for players. This is an extremely frustrating experience for multiplayer competitive gaming where even a millisecond of latency, let alone a prolonged outage, severely impacts the user experience.

Cambridge-based Jagex Games Studio is a multi award-winning games developer and publisher of popular gaming communities such as Runescape, War of Legends and Transformers Universe. With 450 staff onsite serving more than three million unique users per month, it is the largest independent games company in Europe and has built a global reputation for developing popular, accessible, free-to-play games along with providing a first-class community experience for millions of players around the world.

Jagex had been experiencing an increase in DDoS attacks for a number of reasons, including: banning a small number of known DDoSers, which prompted more attacks; attacks driven for pure attention seeking purposes; and those looking for bragging rights. While each of these motivations may seem like small-time nuisances to a business that relies on player accessibility, they pose a serious challenge to game availability – the lifeline of the organisation. In addition, extortion attempts targeting Jagex were starting to become more commonplace. One instance saw bad actors taking to Twitter, threatening to take down the site unless Jagex paid their ransom requests or made changes to the game itself.

Jagex was experiencing both the large, volumetric DDoS attacks and the sub-saturating attacks that probe the network in an attempt to uncover weak spots in security defences – in fact, the latter constituted the majority of the attacks on Jagex. “We currently see 300-400 non-critical attacks on our infrastructure per month, and that’s being conservative,” Jagex’s head of IT Barry Zubel explains. “For the IT personnel, that could mean 10 or more call outs per day for a 24/7 service, which of course also meant paying staff around the clock.”

It became abundantly clear that the company needed an alternative means of defeating its DDoS challenge after the New Year in 2013, when attackers used NTP reflection attacks in attempt to take down every major gaming site in the world and managed to throw all Jagex datacentres offline simultaneously. With five main datacentres scattered throughout the US and Europe, Zubel realised that the legacy DDoS mitigation solution Jagex was using was not cutting it in terms of performing to specifications.

As a result, Jagex decided to deploy an inline, real-time DDoS mitigation solution after testing it in its UK datacentre. This solution utilises advanced analytics tools, which allows Jagex to see clearly when it is under attack and whether its inline solution is managing to mitigate it. The company also chose a solution that was modular and scalable, so capacity can easily be increased as the threat landscape changes.

The results it observed have saved the firm’s security teams several sleepless nights and immediately got the DDoS problem back under control. Roughly 95% of DDoS attacks are now absorbed completely and automatically, with zero impact on player experience. This is because attacks are being blocked before they reach servers, and, as a result, gaming sessions can continue uninterrupted. This means that Jagex can dispense with costly human analysis and configuration of DDoS countermeasures. The success of this solution means that Jagex can focus on enhancing its customers’ gaming experience – and the potential for improved reputation and revenues that this could bring.

About the author
Dave Larson is chief technology officer and VP, product for Corero Network Security and is responsible for directing the company’s technology strategy, which includes providing next-generation DDoS attack and cyber threat defence solutions for the service provider and hosting provider market segments. Larson brings over 20 years of experience in the network security, data communication and datacentre infrastructure industries.

Reference

Detecting and destroying botnets
Garrett Gross, AlienVault

Due to their limitless size and capacity, botnets are among the most powerful components of a modern cyber-criminal’s arsenal of attack techniques. They are made up of compromised workstations distributed over the public Internet that leverage the untapped processing power of a multitude of endpoints, usually to accomplish a malicious agenda. Each of these endpoints, or ‘bots’, typically links back to a command & control (C&C) server so the entire botnet can be used as one tool to engineer data theft/fraud or spam marketing on a mass scale, as well as to power huge Distributed Denial of Service (DDoS) attacks.

One of the most vexing qualities of botnets is their ability to operate under the radar for long periods of time, due to their size and the difficulty involved in detecting them. As an example, the Gameover ZeuS
A botnet (based on the Zeus trojan) caused havoc for quite some time before authorities shut it down. An estimated $70m was stolen through Gameover Zeus before the FBI arrested over 100 individuals associated with the botnet. In early June 2014, the US Department of Justice announced that an international inter-agency collaboration named Operation Tover had succeeded in temporarily cutting communication between Gameover Zeus and its command and control servers.1

Nine years later, Gameover Zeus is still going strong, and has lasted longer than many other pieces of malware. Over time, the number of botnets is drastically increasing and their financial value is too. They are also becoming more sophisticated in their targets, infiltration, anti-detection and attack techniques. This is why now, more than ever, it is crucial for IT security professionals to be highly knowledgeable about a range of botnet detection tools and techniques.

The history of botnets

It is difficult to narrow down the first pieces of malware that started the botnet snowball effect. Some would suggest that 1999 was the magical year with the surfacing of the Sub7 trojan and the Pretty Park worm. Both of these nasty pieces of malware used the victim's computer to connect to an IRC channel and silently wait for malicious commands.

Now let’s jump forward a decade to 2009, when the Ponmocup botnet began its savage reign, infecting 15 million computers throughout its existence. Yes, you read that correctly, 15 million! The Fox IT researchers who reported on Ponmocup explained that its infrastructure is “complex, distributed, and extensive, with servers for dedicated tasks”, making it “one of the most successful botnets of the past decade, in terms of spread and persistence.”2

What is frightening is that Ponmocup is still being developed, and researchers have found 25 unique plug-ins and over 4,000 variants.

“What is frightening is that Ponmocup is still being developed, and researchers have found 25 unique plug-ins and over 4,000 variants”

Ponmocup is very difficult to detect as a result of its use of anti-analysis techniques such as heuristic checks for network and host-based analysis tools, debuggers and virtualised environments. If the anti-analysis checkers spot an attempt to analyse the malware, a fake payload is delivered, injecting easy-to-remove adverts. That fake payload disguises the delivery of a much more serious one. It is thought that this botnet is being used for financial gain and we can only imagine the immense sums of money the perpetrators would have made by now.

Another notorious worm that has been making the rounds since around 2008 is Conficker, also known as Downadup, and it is not planning to stop any time soon. The worm has recently resurfaced attacking not only traditional computers but also a range of other objects ranging from IoT devices to police body cameras. Conficker attacks Windows operating systems, forming a botnet that both infects the machine and spreads throughout a given network automatically. It has been suggested that as many as 20% of all global attacks are due to Conficker.

Microsoft’s own advisory states that the Conficker worm spreads by copying itself into the Windows system folder. Microsoft explains that, “It might also spread through file sharing and through removable drives, such as USB drives (also known as thumb drives), especially those with weak passwords.”

So, clearly botnets have been around for a while and can be notoriously difficult to remove. So what is the best approach to detect and eradicate them from your network?

Bots on your network

IT professionals can pick up on several different clues that can expose a botnet hid-
ing in their network. These tend to appear shortly after botnet infiltration as a compromised machine begins following malicious instructions. Some of the symptoms are:

- Connecting to established C&C servers to receive instructions.
- Creating Internet Relay Chat (IRC) traffic via a specific range of ports.
- Inducing simultaneous identical DNS requests.
- Generating Simple Mail Transfer Protocol (SMTP) traffic/emails.
- Decreasing workstation performance/Internet access to the point of becoming apparent to end users.

These issues appear both at the level of individual, compromised workstations and in the network as a whole. For network managers, that means there are various botnet detection tactics and tools that can be used at both of these levels.

Botnet detection at the endpoint level starts with client-side anti-viral solutions, as the infection itself usually happens through malware. However, anti-viral technology that relies on signature-based detection alone cannot identify new variants of malware simply because that exact code hasn’t been seen or researched before, so you need use other tools and detection techniques as well.

Host-based botnet detection includes checking for activity such as rootkit installations, unexpected pop-ups while browsing over HTTP (this may just be spyware), or any abrupt change to the Windows Hosts file, which can be used (or abused) to restrict outbound server access. Of course, if the default DNS servers are changed, this is probably an indication that traffic is going to unwanted destinations.

**Network detection**

Network-based botnet detection tends to be more complicated. One method is identifying and keeping an eye on IRC traffic, which in a normal situation shouldn’t exist on a business network. Additionally, IRC traffic is sent unencrypted, which means that keywords can be spotted with a packet sniffer. Although the default IRC port is 6667, the whole port range (from 6660-6669 and 7000) can be used by bots.

As discussed earlier, if multiple endpoints abruptly and simultaneously strike one or more external sites, this is often a clue that a botnet-driven DDoS attack is being executed from your network. Similarly, mass outward-bound traffic happening over SMTP implies that spam-mailing may also be occurring. Embrace rules for these signs in your network-based security tools to tune them for botnet detection.

More ambitious security professionals should consider creating a honeypot (false infiltration opportunity) to see if it is infiltrated – and if so, in what manner. If you use Suricata, a free open-source intrusion detection solution, you can obtain a list of botnet recognition signatures for it. Also, you should always monitor for any attempts to connect to known C&C servers.

**Static vs behavioural detection**

There are two different ways to detect botnets – static analysis and behavioural analysis. Static analyses are simple, quick and resource friendly. Behavioural analyses are much more detailed; however, they are also more resource-intensive.

Static analysis is your first line of defence because it looks for a highly specific match to something like a malware signature, specific executable or C&C connection address. Unfortunately, it doesn’t always work. Botnet managers (‘herders’) are getting more and more refined in the way they evade such simple clues, using counters such as file

**Best practice**

1. Deploy both host- and network-based botnet detection tools; neither can identify every occurrence every time on its own.
2. Ensure your host-based IDS or an anti-viral tool is capable of finding the common endpoint symptoms of botnet infection and is often updated with the most recent C&C server information.
3. Implement a honeypot (or several) if you are protecting reasonably important data, have lots of brand equity in your company’s name, or are a particularly juicy target for a lawsuit by a victim of a botnet-based attack starting from your network.
4. Use static analysis at a minimum, but companies should focus botnet detection on the more effective behavioural analysis if at all possible.
5. Talk to in-house and external specialists regarding P2P botnet detection techniques.
6. Ensure that the rules for your behavioural, network-based botnet detection systems consider less common systems.

A botnet for sale on an underground cybercrime website.
polymorphism to change the executables in unexpected ways, URL obfuscation to disguise the targets of DDoS attacks, server proxies, and even suddenly changing the IP addresses of their own C&C servers. Therefore, static analysis is just not enough for botnet detection.

For this reason, behavioural analysis, or analysis that looks for anomalous activity that may be indicative of an infection, is an essential approach to spotting botnets. For example, the timing of attacks can give a lot away – a C&C server usually distributes blanket orders for bots to carry out specific behaviours, creating a lot of network activity at a given point in time.

The normal amount of time for an endpoint to connect to different outbound servers tends to be low for bots as there isn’t a human driving that network activity. There are also more failed connection attempts for this reason and those connection attempts are more likely to involve IP addresses instead of server names. Also, port-scanning the local network for new intrusion opportunities is classic behaviour of a bot. All of these actions can be found by applying SIEM or network IDS rules that expand a company’s botnet detection abilities.

One development in botnets that further complicates detection methods is the rise of a peer-to-peer (P2P) management architecture. This works in a decentralised way, such that there is no central C&C – bot commands are instead issued by peers. These types of botnets are harder to spot, although infected bots will act similarly to the way they would in a traditional botnet because the bot herder still has the same objective.

While attack campaigns in the past might have passed up ‘lower priority’ systems and devices that do not store payment or other sensitive information, we are seeing botnets designed specifically to go after them. We can attribute this to today’s lower compute cost as well as the pervasiveness of devices deployed with vendor default passwords, such as network devices, wireless access points and even CCTV cameras.

**Botnet tools**

Luckily, as botnets have evolved, so have the tools to find and destroy them. Focused open-source solutions like Snort and more comprehensive, integrated security intelligence tools can help you, by:

- Establishing when network activity is unconventional in predefined ways.
- Finding its network origin.
- Analysing its nature and impact.
- Directly quarantine limit, or eradicate local bots.

Looking to the future, botnet detection solutions are getting smarter in a variety of ways, some tech-centric (such as implementing machine learning for botnet pattern recognition), some human-centric, and some that combine the two. For example, the Open Threat Exchange (OTX) harnesses crowd-sourced wisdom and combines this with the technology of an integrated SIEM, log management, threat detection and vulnerability management. This gives users a powerful platform to detect and defeat botnets and other modern threats, because when the IT security community works together by exchanging information, we have an opportunity to stay one step ahead of potential attackers.

**About the author**

Garrett Gross is senior technical product manager at AlienVault. He has held many positions in IT over the years and leverages that expertise to help fine tune how AlienVault markets to the security community.

**References**


---

**Anonymity networks and the fragile cyber ecosystem**

Hamish Haughey, Gregory Epiphaniou and Haider M Al-Khateeb

It is well known that government agencies have had the capability to eavesdrop on public switched telephone networks for many decades. However, with the growing use of the Internet and the increasing technical capabilities of agencies to conduct mass surveillance, an individual’s right to privacy is of far greater concern in recent years. The ethical issues surrounding privacy, anonymity and mass-surveillance are complicated, with compelling arguments for and against, due in part to the fact that privacy and anonymity are desired by criminals and terrorists, not just individuals who care about their privacy.
Considering ongoing political debates concerning the abuse of personal information, international espionage, law enforcement and acts of terrorism, it is clear to see that this is a complex landscape for discussion. There is currently no agreed consensus on anonymity or privacy and it is challenging to find one when the reality is that freely voicing one’s opinions is punishable by law in one country and not in another.

Technologies such as The Onion Router (TOR) employ the concept of virtual circuits to establish a connection between hosts with strong cryptographic primitives to assure confidentiality while also, enabling entities to protect their identity and preserve privacy without exposing information to an adversary. There are however a range of attacks available to adversaries that can degrade that anonymity.

Privacy enhancing technologies

TOR is not the only solution to this problem, and a vast number of Privacy Enhancing Technologies (PETs) have been developed over the years. Anonymity in cyberspace has always been a domain of conflicting interests as regards freedom of speech, but at the same time gives enough power to adversaries to mask their illegal activities. Certain services on the hidden web – such as dynamic content, unlinked pages, private websites, limited access content and scripted pages – offer a higher degree of anonymity preservation, as well as much wider access to information. This can be potentially invaluable information to companies, organisations and governments, but also to organised cybercrime.

On the other hand, the technologies are of concern to parties who wish to know what communications are taking place, such as: law enforcement who wish to stop illegal drug trafficking or child abuse rings; governments who wish to protect their nation from terrorism; organisations that wish to identify the source of network attacks; commercial businesses that collect user data for analytical purposes in order to sell data to other businesses or enhance targeted advertising campaigns; espionage and counter espionage agencies; and oppressive regimes that would like to prevent dissident communications. This results in a form of anonymity arms race, with researchers openly disclosing vulnerabilities to anonymity systems in order to improve the technology and make users more aware, together with agencies finding vulnerabilities that they can leverage to gain an advantage by degrading their target’s anonymity while maintaining their own.7

One could argue that, given the reasons above, the protection of the vulnerable and of national security have to be the most compelling arguments, and that governments and law enforcement have the strongest case for breaking anonymity or privacy. However, history has shown that such power is open to questionable, excessive or liberal usage.8,9

“TOR relies on volunteers who allow their servers to be used as TOR relays, which participate by serving as routers between the users and their destination”

Three technologies in particular have gained in popularity over the years: TOR; the Invisible Internet Project (i2p); and Freenet.10-12 Each one attempts to address the issue of anonymity in a slightly different way, primarily due to the different threat models they employ to quantify and describe anonymity online.

TOR

TOR is a low-latency, wired, free route, unicast, cycle-based anonymous network that provides sender-receiver and community anonymity by concealing the source, destination and contents of a message with a series of encryption layers.13 TOR relies on volunteers who allow their servers to be used as TOR relays, which participate by serving as routers between the users and their destination. The system uses perfect forward secrecy in allowing the user to build a circuit on the TOR network in which no relay is aware of both the source and the destination. A TOR client typically chooses three relays from a centrally maintained directory of nodes based on their performance capabilities in terms of bandwidth. The final relay is called an exit, and the first relay an entry.

The encryption process itself begins with the client encrypting the message together with the IP address of the destination server, using the public key of the exit relay to create the first layer of encryption. This encrypted message is in turn encrypted together with the IP address of the exit node using the public key of the middle relay, adding the second layer of encryption. The second encrypted message is once again encrypted together with the IP address of the middle relay using the entry relay’s public key, thus creating the third layer of encryption. This message is then sent to the entry relay, which will begin by unwrapping the top layer of encryption that it has the key for and forwarding the remaining message to the next relay.

Active traffic analysis is often employed by adversaries in which they inject traffic into legitimate communication channels in an attempt to enumerate and analyse traffic by initiating rogue routers. These common and easy to deploy attacks may seriously undermine anonymity in anonymous communications and proactive network defence techniques. Attacks in this category usually involve metadata for a given network connection, such as timing information and source and destination of messages, to degrade anonymity.

i2p

Compared to TOR, i2p is a high-latency network and is primarily designed to be self contained with hosted services being available only to users of the system. It does allow anonymous public Internet browsing through specific proxies that some volunteers choose to provide, but this is a secondary feature. This is in contrast to TOR, which is primarily for enabling anonymous Internet browsing with provisioning of hidden services as a secondary function. Unlike TOR, which relies on volunteers to provide their servers to act as relays, in using i2p the user participates by...
becoming a router, thus contributing to the system’s peer-to-peer functionality.

“A tunnel is unidirectional and will either send or receive traffic between users or from a user to a service. As such, a user is required to set up two tunnels for a two-way communication”

The i2p solution implements tunnels, which are similar to TOR circuits in that the encryption is layered and each router is only aware of the previous and next router. A tunnel is unidirectional and will either send or receive traffic between users or from a user to a service. As such, a user is required to set up two tunnels for a two-way communication – one tunnel for outgoing messages and one for incoming.

The system maintains a database of gateways for users’ inbound tunnels, based on Kademlia, a distributed hash table designed for peer-to-peer systems. This allows a sender to initiate communication from their outbound tunnel to a recipient’s inbound tunnel. This distributed database is known as netDB and is automatically shared by participating routers that are recognised by the system as having high-bandwidth capabilities. High-bandwidth routers that participate in the netDB distribution are known as floodfill routers. The netDB contains two types of data: leaseset data, which relates to destination services; and routerinfo data, which relates to router information such as its IP address and TCP port.

Leaseset data is only published in the netDB for inbound tunnel gateways and contains all of the currently available tunnels that can be used to reach a specific destination, together with the public key used to encrypt data to be sent.

In contrast to onion routing, i2p implements a similar system called ‘garlic routing’. This method allows packets to be packaged together in ‘bulbs’ or ‘clove’ within one garlic message that is sent over a tunnel, although typically only one clove is sent. On occasion, two other cloves can be sent, including acknowledgement information, and information for a return path to prevent the destination having to query the netDB. Furthermore, all i2p traffic is encrypted end-to-end by default.

A perceived benefit of i2p over TOR is that a client does not need to create a whole new circuit for every new communication, but can publish a small number of inbound tunnels that can be reused by multiple connections simultaneously. The i2p site states that as each tunnel only deals with traffic in one direction, an adversary would have to compromise twice as many communication paths in order to associate traffic compared to TOR.

Typically, i2p services are internal only and identified cryptographically, and as such DNS cannot be used to locate them. Services can be identified using a number of mechanisms including directly by their SHA256 cryptographic identity, a Base 32 encoded version of that identity, or with shorter friendly names that are stored in a local hosts file. There are also a number of lookup services known as jump sites that can be used to search for known i2p services and locate them.

Freenet

Freenet is a peer-to-peer distributed anonymous file-sharing network that allows users to anonymously publish and access files and websites (called freesi- tes) in so-called Opennet and Darknet modes. The system is also extended via a number of applications to allow anonymous social networking, discussion forums and email. All participants in Freenet contribute some of their
bandwidth and storage to the distributed peer-to-peer system, and allow the files of other users to be stored encrypted on their resources, and in turn accessed by other users. Identifying the location of files that have been stored in the system is by unique identifying keys that must be known to whoever needs to access those files. Storage is not permanent and files that are not searched for or accessed for extended periods will reach end of life and be removed from the system. Files are encrypted in transit and at rest.

“All discoveries of vulnerabilities that are disclosed to the community ultimately assist developers to work on a stronger system and inform users of the risks”

Routing and distribution of files in Freenet are both highly dynamic, with clusters of nodes naturally forming to enable files to be accessed more efficiently by being distributed among more nodes. Each node is supposed to be aware only of its own nearest neighbours and requests for files are forwarded on by those nodes if they do not hold a copy, until the request reaches a node that does.

**Attack vs defence**

We undertook a survey of 21 recent peer-reviewed research papers into attack methods on TOR, i2p and Freenet. It was found that there is an apparent tipping of the balance towards research into the TOR network compared to other PET solutions, which is perhaps unsurprising due to its more widespread adoption and publicity. There is also a heavier distribution of publicly available papers that propose novel attack mechanisms compared to defences, with 17 out of the 21 papers reviewed describing attacks, and 11 of those not providing detailed descriptions of a proposed defence. This does not necessarily infer that researchers have any kind of bias with respect to anonymity, because all discoveries of vulnerabilities that are disclosed to the community ultimately assist developers to work on a stronger system and inform users of the risks.

The attacks described all require that the attacker has a particular position with respect to the target. These positions include a destination website that a user might visit, a single node in a peer-to-peer system, a single node or relay in a network, multiple nodes or relays, a network position local to the source or destination, or high-level access to an Autonomous System (AS) or Internet Exchange (IX). Any particular attack typically focuses on one aspect of the system, whether the target is the user, a service host or the system itself.

**Attack difficulty**

Different attack types come with varying degrees of difficulty for the attacker. For instance, inserting a Floodfill node into the i2p network provides access to statistical information about the network. It is then possible to correlate geographical location of nodes with the number of applications by type and infer geographical regions responsible for a type of activity.15 By contrast, this type of attack is considerably easier than forcing a client to use a malicious relay, or having access to an AS- or IX-level router. However, the information that can be gathered by such simple attacks is limited. For instance, Timpanaro et al (2014) found that only two Russian cities were responsible for over a third of all worldwide traffic associated with i2psnark, an anonymous file sharing application that is included with the i2p software. This information is useful for providing a commentary on the general usage of i2p, but not for identifying individual users. However, it is acknowledged that this constitutes an overall reduction in anonymity of the system as a whole.

Access to relays is typically fairly easy to achieve, as the PET systems are designed to allow anyone to participate.16-24 For instance, an attacker could create a number of virtual machines with ease given today’s readily available cloud computing services and inject them into the anonymity network. The more malicious relays that are inserted by the attacker, the greater the chance that circuits will start being built with entry and exit relays that belong to the attacker, although this obviously comes with a cost, which is considerably increased due to the number of TOR relays in existence. Zhang et al (2011) discusses how, with this prerequisite, it is possible for an exit relay to insert a signal into the traffic flow that can be detected by the entry relay. This was achieved by manipulating TOR application code to finely control when a write event is called, resulting in the output buffer being flushed. Manipulating whether one or three cells are sent in an IP packet allows a binary signal to be created. If the signal sent by the exit relay is matched by the entry, then the user has been identified together with the website they have been visiting. The difficult part of this class of attack for an adversary is getting the target to use entry or exit relays that the adversary controls, or compromising relays that belong to others.

“The more malicious relays that are inserted by the attacker, the greater the chance that circuits will start being built with entry and exit relays that belong to the attacker”

The most powerful attacks are those that have a prerequisite of AS- or IX-level access, as these greatly increase the chances of observing servers that are facilitating PET participation.25-30 In particular, it has been found that the likelihood of an adversary with IX-level access serving ASs that appear on both sides of a circuit is far higher than previously thought, due to traffic between the user and entry relay, and between the exit relay and destination, passing through a number of routers on multiple ASs and IXs.

Liu et al (2013) conducted a real-world study to understand the distribution of IXs on TOR entry and exit relays, and quantify the likelihood of the same IX or AS appearing on both ends of a circuit. The study performed traceroutes from TOR clients to a list of entry nodes, and also from remote destination servers back
to TOR exit nodes. Every identified hop on the route was associated with an AS and an IX using IP-to-AS number lookup tools and an IX mapping project. A single AS was found to have a 17% chance of routing traffic for both ends of a circuit, and a single IX had a 73% chance. This would allow completely passive analysis of traffic patterns in order to correlate website interaction by a user, if the adversary were to have access to routers with this level of Internet routing.

“It is possible to enumerate the network topology in Freenet by enabling an undocumented debug mode or by reviewing the routing tables created and shared between neighbouring nodes”

This kind of attack has a high difficulty due to the access level required, but also because of the amount of processing power necessary to analyse the traffic and correlate traffic patterns. The concern is that the organisations that do have the required access could choose to make use of that power, share it with other organisations, or that their security controls might not be sufficient to prevent an attacker from compromising their systems and performing their own analysis. Given recent revelations that government agencies have been known to intercept Cisco routing equipment in transit, and the possibility of insider attacks in order to update the firmware and allow back door access, it is not unreasonable to consider that this attack method is already being used either with or without the knowledge of the AS or IX operator.31,32 The authors of Liu et al (2013) put forward novel modifications to the TOR node selection algorithm that would reduce this scenario by taking into account specific ASs and IXs on the path to relays.

Attack targets

Fifteen out of the 21 reviewed papers discussed attacks on the network, which seems to be the biggest area of concern for developers of PETs. For example, it is possible to enumerate the network topology in Freenet by enabling an undocumented debug mode or by reviewing the routing tables created and shared between neighbouring nodes up to two hops away.22 Any single node can enumerate the location of up to 402 neighbours, so with collaborating nodes it is trivial to describe the network topology.

This made it possible to enumerate large portions of the Freenet network in order to identify target nodes. A simulation with 200 virtual machines proved that it was possible to pollute the target’s routing table and insert a malicious node. This position can then be used to identify user activity.

There is a known effect of denial of service (DoS) in reducing anonymity. The main proposition is that by increasing traffic or denying service to TOR relays, the attacker can force their own relays to be more likely to appear in constructed circuits for the purposes of performing traffic association attacks.33 Another approach that is possible with sufficient control is to block the access to TOR relays for all clients. Winten and Lindskog (2012) propose that, based on observed behaviour inside the Chinese Great Firewall, the authorities are able to actively fingerprint the TLS handshake between clients and entry relays and dynamically block this activity. They also found that only one of the TOR directory servers was accessible from within China, which may be for their own ease of management, or could suggest that the Chinese authorities may have some level of control over a TOR directory server.

“By analysing TCP traffic at both ends of a circuit, where the identity of one end is already known, it is possible to associate the traffic and de-anonymise the user”

A common theme in many studies appears to be that PETs are vulnerable to attacks that require the adversary to position themselves at multiple specific locations on the communication path in order to perform traffic analysis and identify the source or destination.34,35 Given the difficulty but effectiveness of this attack, it is unsurprising that it receives a lot of attention. The premise is that by analysing TCP traffic at both ends of a circuit, where the identity of one end is already known, it is possible to associate the traffic and de-anonymise the user or identify the service that they are visiting. One example of this has already been discussed in the creation of a signal that hides within the TCP traffic. The other variation on this association attack is by counting the number and timing of TCP packets at both ends of a circuit. Computer algorithms have been developed that can perform this task, although papers tend to focus on simulations involving very specific traffic sent from a known source to a known destination. The difficulty and resources required to perform this kind of analysis is expected to greatly increase on a real TOR relay that is handling actual traffic, and even more so in the case of AS- or IX-level attacks where the Internet traffic of thousands of concurrent connections may need to be filtered.

A small subset of attacks attempt to reduce anonymity by retrieving information about nodes, either directly or indirectly. Tin et al (2013) is one such example, which demonstrated that by serving malicious JavaScript within a website, it is possible to cause bursts of traffic through the TOR circuit in use by the affected visitors. These bursts of high traffic may be sufficient to reduce performance on the relays in use on the circuit, which in turn could be detected by a described probing mechanism. This probe consists of a client, probe server and an exit relay, which are used together to create very specific circuits designed to contain only one real TOR relay. This is facilitated by leaving the exit node switched off when not in use and advertising low bandwidth capabilities so that it is unlikely to be used by any real circuits. By initiating connections to other real TOR relays using the controlled exit node, it is then possible to measure their current performance in terms of bandwidth and latency with respect to the probing system. Baseline measurements would need to be taken for all TOR relays in the directory and used to compare with
any measurable effect of the bursting, which would be set to happen with precise durations and intervals.

It was not discussed how practical it is to actually initiate connections to all existing TOR relays from a single system, or whether multiple probing systems would need to be set up, and the proposed defence against this attack would be subject to the same considerations. This attack is also likely to become less effective as more PET users become aware of the dangers of JavaScript. For instance, the Tor Browser package has JavaScript disabled by default and the user must go and manually enable it.

**User error**

There are many ways of covertly tracking a user's behaviour on the Internet, such as browser cookies, etc. And there are necessary precautions to avoid revealing their identity as suggested by the providers of PET systems. However, some attacks may exploit user error, or problematic behaviour in their habits that increase their exposure.

“Internet routing is asymmetric in its nature and visibility of only one direction of this traffic flow is required to analyse traffic”

An interesting point has been made in that configuration error is a large risk for inexperienced users of the i2p system. For example, it is possible to create an i2p service served over the i2p network, but which is also available publicly on the same server. Similarly, it is possible using i2p to create a ‘hidden’ service that simply points at a publicly facing service. These are examples of configuration mistakes made by users who do not fully understand the implications, and similar user configuration errors can occur on TOR hidden services, or PETs in general.

**BGP interception and asymmetric analysis**

One particular study in recent years gained significant attention due to the implications of high-level adversaries engaging in massive scale anonymity degradation. The study in question proposed a novel attack method, which they called Raptor (Routing Attacks on Privacy in TOR).

The study built on previous work by Vanbever et al (2014) in an attempt to quantify the effectiveness of AS-level adversaries in de-anonymising users of TOR, and described three basic assumptions that the researchers go on to show can be exploited individually or in combination. First, Internet routing is asymmetric in its nature – ie, the path that a packet takes is not necessarily the same as that of its reply – and visibility of only one direction of this traffic flow is required to analyse traffic. Therefore, the attack surface and likelihood of exposure to an adversary performing a passive traffic association attack are greatly increased. Second, BGP ‘churn’ over time due to network changes or transient issues such as router failures increases the chances that a regular user of TOR will cross paths with a particular AS or IX, facilitating passive traffic analysis. Finally, it is possible for an AS operator to make false BGP announcements in order to have traffic intended for another AS route through their own routers in an active attack, which positions themselves on a target circuit.

The active attack comes in two versions, hijack and intercept. The drawback of the hijack is that the intercepted traffic never reaches its intended destination and so the client experience will be interrupted, potentially raising the alarm, or causing a new circuit to be built. An interception on the other hand allows the traffic to continue to its destination and the client session remains active. There are three possible scenarios for the interception attack. First, the adversary already monitors a website or exit node and targets a set of known TOR entry relays. Second, the adversary already monitors the client to entry relay side and targets known TOR exit relays. Or third, the adversary targets both entry and exit traffic in order to perform general monitoring, data gathering and correlation.

It should be noted that while the impact of a hijack to a user is obvious, there is no mention of the impact on user experience during an interception in terms of latency or packet loss while the new routes propagate and converge.

**Combination attacks**

Something that does not seem to have gained much research attention is the possibility that, with sufficient resources, a powerful adversary could combine several documented attacks to further degrade anonymity. Some attacks are likely to be more compatible with others and the end result may be an increased success rate, or decreased number of resources required.

For example, if an AS- or IX-level adversary could enumerate the list of TOR nodes that it does not have visibility of and carry out a DoS or circuit-clogging attack – as described in Tin et al (2013) – against those relays, as well as passively intercept entry and exit relay traffic with the advantage of asymmetric routing, then this could increase exposure for traffic analysis without having to perform a BGP interception attack.

**Legal and ethical considerations**

The research highlights many examples of legitimate uses for PETs, but also a great number of illegal use cases. As such there is mention of the providers of TOR relays having to deal with legal and administrative challenges due to their systems being identified as engaging in illegal activity.

“The concept of ‘accountable anonymity’ is difficult to reconcile as the terms are commonly perceived as mutually exclusive”

The fact that the definition of what is legal varies from country to country highlights an opportunity for the development of a transparent and configurable system.
to identify and classify malicious activity, and block it accordingly. In the case of TOR, it is noted that this could only be applied to exit relay traffic, as entry and intermediate relays are unaware of the intended destination or the communication. There is already research into the area of allowing anonymity but preventing malicious activity. However, the implementation of such solutions will always be subject to close scrutiny due to the overriding concerns regarding privacy, trust in the government, and varying laws and regulations by country. Unfortunately, the concept of ‘accountable anonymity’ is difficult to reconcile as the terms are commonly perceived as mutually exclusive. However, there is some conceptual commentary in favour of the plausibility of such a system.

There are views of the opinion that there is no such thing as good censorship. To address such concerns while also assisting providers of TOR relays who are worried about the repercussions of providing their resources, there may be a way for TOR operators to publish the level of monitoring that they perform. A client could in that case choose whether they wish to use TOR circuits that perform no monitoring, or choose relays that block particular types of activity. For instance, a member of the public may be quite happy to use a TOR exit relay that blocks known malware, botnet activity, and requests to sites that are known to be illegal in some territories (or immoral in their own views), in the knowledge that they can still anonymously access sites that are banned in their own country, or to escape other types of tracking that service providers may engage in. Conversely, strong believers in absolute privacy would still have the option of selecting relays that engage in absolutely no monitoring or blocking, on the understanding that they are sharing their source IP address with less scrupulous users, and that these connections would be of greater interest to the authorities, which in turn are expected to have greater de-anonymising capabilities.

A system of this kind may increase attacker capability as they could publish their malicious relay as a particular type (e.g., highly permissive), and thus be more likely to be included in a circuit used by particular users. Strong privacy advocates are unlikely to endorse such a system for this reason, however an interesting observation in this debate is that terrorists and criminals in recent years have been observed to ‘hide in plain sight’, which could remove some of the argument for government access to encryption systems.

Identifying new vulnerabilities in software creates a risk to the users of that software, and for such widely-used anonymity systems this comes with real risk to the lives of some of those users. Researchers may wish to consider the responsible disclosure of these vulnerabilities to the software developers in a similar way to that which is generally accepted in the wider software industry.

**Conclusion**

The complexity and decentralised management in anonymous networks can lead to information leakage that can be potentially catastrophic when it comes to privacy preservation online. In specific anonymous networks the adversaries can alter the way packets are encapsulated and encryption is applied to gain information about the synchronisation counters distributed across these nodes and impair the encryption/decryption process in a multi-hop environment. Different threat modelling employed in each implementation underpins the probability of anonymity exposure rate based on these actions. Integrity attacks also manifest, especially in cases where integrity verification does not take place while traffic is encrypted between different nodes, making degradation attacks a green field with respect to existing literature.

Research into attacking and defending PETs will no doubt result in further iterations of improvements to the technologies, unless any particular solution becomes fundamentally broken, at which point users will likely move to the next solution. This arms race is expected to continue until a wide agreement can be reached on anonymity and privacy, which may still be a very long way into the future.

**About the authors**

Hamish Haughey has a background in web hosting systems and supporting infrastructure and currently works as an information security analyst. He is studying for an MSc in cyber-security with the University of Northumbria through QA Training and has research interests in network security, anonymous systems and ethical hacking. He has a strong risk-based approach to security in real situations and actively assists with audits to establish gaps and then agree appropriate remediation with the organisation, as well as updating written policies when required.

Dr Gregory Epiphaniou is a senior lecturer in cyber-security in the Department of Computer Science and Technology at the University of Bedfordshire. He has worked as cyber-security consultant and trainer for QA Training with several industry partners in information security domains. He has delivered a wide range of industrial and bespoke certifications including but not limited to CISSP, CISM, CEH and BCS. He has also been a trainer and developer for the GCHQ Summer School in cyber-security.

Dr Haider M al-Khateeb is a lecturer in the Department of Computer Science and Technology, University of Bedfordshire and a Research Fellow in applicable computing and an active member of the National Centre for Cyberstalking Research (NCCR). He has a first-class BSc (Hons) in Computer Science, and PhD in Computer Security. He has cross-disciplinary skills in computing but is mainly interested in cyber-security and digital forensics. His current research is in areas including memory forensics, authentication methods, Internet of Things security and cyberstalking. Haider is also a Fellow of the Higher Education Academy (FHEA), UK.

**References**


4. ‘Who uses TOR’. The Tor Project. www.torproject.org/about/torusers.html.en


34. Kwon, A. ‘Circuit fingerprinting attacks: passive deanonymization of TOR'.
hidden services’. In USENIX Security Symposium, 2015, 24th Proceedings.
36. ‘Want TOR to really work?’. The Tor Project. www.torproject.org/download/download.html.en#Warning.
43. ‘Isn’t censorship sometimes necessary?’. Freenet. https://freenetproject.org/about.html#is_censorship_necessary.

A month new, a new data breach

Paul German, Certes Networks

The wave of data breaches washing over enterprises around the world has surfaced an essential truth about IT security – typical enterprise security architectures are vastly outmoded in terms of how modern applications are shared, managed and accessed by users. And with each month bringing a new data breach, the implications for both compromised employees and the wider infrastructure are significant.

Security officers are losing sight of the boundaries of their enterprises as both applications and users become equally mobile. Mobility and cloud strategies are increasing productivity, driving down costs and enabling businesses to become increasing agile; spinning up on-demand resources, or moving applications to cloud environments as and when required is fast becoming the norm.

In the inevitable high-profile fall-out of the Office of Personnel Management (OPM) breach, the debate has raged about the lack of encryption – despite the fact that encryption alone could not have prevented a breach on this scale because there was no effective and secure segmentation of users or data.¹ With an initial breach of 4.2 million personal records and an additional 15 million records that may have been compromised, many organisations are now attempting to make haste to avoid another such monumental breach of personal information. With this in mind, we must be careful not to repeat old mistakes: the OPM breach really reveals that it is time to think differently about security and embrace cryptographic user-to-application segmentation.

Theories abound

Every major data breach – and the OPM data breach was a classic – prompts a huge array of theories regarding what could and should have been done to prevent it. When millions of personal records about government employees go missing, the investigations are both intense and high profile. From the extensively reported hearing by the House Committee on Oversight and Government Reform to calls for the OPM’s senior management to resign, this has been a breach that has played out in the public eye. And the question is how many more data breaches will it take for organisations and agencies to wise up and implement new cyber-security tactics and tools?
The general conclusion has been that the biggest issue was not the failure to block the initial breach but a lack of controls, time to detection and other safeguards that should have prevented intruders from obtaining any useful information. In reality, the OPM failed to take appropriate actions and implement modern cyber-security solutions. With the OPM’s – hopefully former – strategy, once a single user or corporate application was compromised, everything became at risk. Networks are only as strong as their weakest component – whether it be leaked login credentials or a rigid, legacy environment based on firewalls.

“When an intruder has the credentials of a user on the network, then data can be accessed even if it’s encrypted, just as the users on the network have to access data”

The fact that the data stolen in this massive breach was not protected by data masking, redaction and encryption is something of a red herring. What the OPM breach really highlights is the continued problem of traditional network-based segmentation – namely the ability to compromise a single user’s identity to gain access to a mass of cross-organisational information.

Yet in an era of continued evolution of the threat landscape combined with an increasing diversity and complexity of the underlying IT architecture, just how can a Chief Information Security Officer (CISO) impose greater control and achieve that essential user specific level of application and data control?

Security best practice

There are some aspects of security best practice that are now a given. A defence-in-depth approach that combines multiple layers of prevention and detection technologies, combined with procedural controls and policies is essential; user identification and access control is a standard tool to provide central administration and control; and intuitive intrusion detection tools are becoming increasingly key to identify when breaches occur before they have had time to gain vast swathes of data – although this latter issue is certainly one with which organisations continue to wrestle.

Other areas of security best practice remain opaque. And one of the biggest issues that continues to challenge the CISO is the need to segment sensitive and non-sensitive applications, or to segment networks into manageable areas that not only restrict access but also ensure that, should unauthorised access occur, critical applications and data are not compromised.

“A cryptographic relationship creates a clean and unbreakable link between each user and permitted data and applications. With this approach, an organisation can ensure that in the event of a breach the intruder cannot reach out beyond those defined limits”

One fact, however, is clear: simply throwing encryption into the mix is not the answer. As OPM spokespeople have insisted, even if the information had been encrypted, it might not have been enough to stop attackers from getting usable data from this intrusion. According to the OPM, when an intruder has the credentials of a user on the network, then data can be accessed even if it’s encrypted, just as the users on the network have to access data – which is what occurred in this case.

LANS, WANS and other environments are not encrypted and are very easy for hackers to traverse. This means that hackers are free to move laterally through the IT infrastructure once they have compromised a single user’s credentials. In other words, they can ‘hop’ from a fairly innocuous application, like a CRM system, into a more sensitive application containing credit card data, patient information or other more valuable targets. That lack of internal segmentation, the freedom for hackers to move laterally, is how many hacking victims were exploited.

If, however, the OPM had had effective segmentation in place, this breach could never have reached this massive scale because the intruder could only have accessed that data and applications to which that user had been permitted, ensuring controlled access. Lateral movement from the compromised application into the more sensitive applications would have been prevented, effectively containing the breach and limiting its impact through segregation and compartmentalisation.

Software-defined security

So how can that be achieved? The key is to leverage the power of encryption in a highly focused and targeted way to create a cryptographic flow between each user and each application. Building on the identity and access control technology widely deployed, a cryptographic relationship creates a clean and unbreakable link between each user and permitted data and applications. With this approach, an organisation can ensure that in the event of a breach the intruder cannot reach out beyond those defined limits and privileges to access other restricted information.

...Continued on page 20
One of the most compelling aspects of this model is that it removes the constraints specific to the infrastructure and embraces a new, software defined security approach. Applications and data are located across a hugely diverse infrastructure — from LANs to WANs, private to public clouds, mobile networks, the Internet and other environments. Segmentation techniques utilised in each part of this infrastructure are equally diverse and fragmented, with VLANs, IPsec, TLS, SSL, ACLs and a range of other tools all playing a role in segmenting traffic. This ‘segmentation fragmentation’ and the difficulty with configuring and managing it from end-to-end is the primary reason that effective segmentation is so rarely deployed in practice.

But with each specific ‘user to application’ cryptographic relationship, the infrastructure becomes irrelevant. The issue is, what applications/data should each user be permitted to access and how should they be permitted to access these applications? The answer to this question should then guide segmentation implementation that is oriented around users and applications, not the infrastructure.

Most importantly, this evolved approach to access control and application protection can now be driven by business rules and requirements, as opposed to being limited by what the infrastructure can deliver.

“The intruder cannot use a single compromised user identity to gain free access across the board and hop laterally from one application to another containing more sensitive data”

Taking this approach, the privilege escalation that occurred in the OPM breach simply cannot occur. Rather than relying on traditional network segmentation to control access, with this cryptographic relationship between user and permitted applications, if the user is compromised the intruder gets access to this permitted information – but no further. The intruder cannot use a single compromised user identity to gain free access across the board and hop laterally from one application to another containing more sensitive data.

Conclusion

There is a huge kneejerk reaction to this OPM breach, with demands that encryption is enforced across the US public sector to safeguard this critical data. But the risk is that organisations will make this entire process too complicated – and still fail to achieve the level of security required. There is a massive difference between encryption to the point of entry and using encryption to manage the relationship between a user, the devices that can be used, and the permitted applications wherever they reside. Encryption alone is not the answer. Instead, the solution lies in strong encryption married to identity and access management controls, aligned with applications and user access rights as determined by business rules.

Organisations need to start considering security in a different way – and it is creating that user to application specific cryptographic relationship that will be the key to, finally, preventing these huge – and continuous – breaches in vital data security.

About the author

Paul German is VP EMEA for Certes Networks. He has more than 18 years of experience and was most recently VP/GM EMEA for Sipera Systems, an IT security solutions company sold to Avaya in 2011. In addition, German has broad experience having held key roles with Cisco, Siemens Network Systems and Lehman Brothers. He holds a Bachelor of Science honours degree from Brunel University, London.

Resource


9 April 2016
BSides Oklahoma
Tulsa, OK, US
http://bit.ly/1O3pYZt

15–16 April 2016
BSides Canberra
Canberra, Australia

2–4 May 2016
North America CACS
New Orleans, US
http://bit.ly/1Nqhu2c

23–27 May 2016
HITBSecConf
Amsterdam, Netherlands
http://conference.hitb.org/

28–29 May 2016
LayerOne
Los Angeles, US
www.layerone.org

30 May 2016
International Workshop on Traffic Measurements for Cyber-security
Xi’an, China
http://wtmc.info/

1–3 June 2016
CyCon: International Conference on Cyber Conflict
Tallinn, Estonia
http://ccdcoe.org/cycon/

1–3 June 2016
Symposium on Electronic Crime Research
Toronto, Canada
http://apwg.org/apwg-events/ecrime2016/cfp

19–22 June 2016
Applied Cryptography and Network Security
London, UK
http://acns2016.sccs.surrey.ac.uk/