ELSEVIER

network SECURITY

ISSN 1353-4858 June 2021

www.networksecuritynewsletter.com

Featured in this issue:

Applying the principles of zero-trust architecture to protect sensitive and critical data

Data is a company's most valuable asset, and energy companies, particularly electric, oil and gas, remain at risk of hacking attempts due to their major social and economic importance.

Zero-trust principles can help to ensure that IT systems are protected, mitigating the risk to company operations and sensitive and critical data. However, as David Greenwood of ISN Solutions explains, this model often requires careful planning to ensure productivity and that access to data needed for daily work is maintained.

Full story on page 7…

Grid cyber security: secure by design, continuous threat monitoring, effective incident response and board oversight

Grids are possibly the highest expression of engineering serving society. We have lots of infrastructure, spanning generations of engineering and technology, still in service.

This presents a problem as there are many who see these grids as vulnerable and high-value targets. Dale Geach of Siemens UK examines where the weaknesses lie and how they have already been exploited. As grid operators seek the huge benefits of modernising grid operations, it's important that they adopt a secure-by-design network with continuous threat detection and watertight response plans.

Full story on page 9…

Attack graph reachability: concept, analysis, challenges and issues

An attack graph (AG) is an abstraction that represents the paths by which an attacker could break a security policy, leveraging interdependencies among discovered vulnerabilities.

However, current AG implementations are inefficient on large-scale networks. The increase of the number of hosts in networks causes an increase in the time it takes to generate the AG, especially the calculation time and the complexity of determining reachability. In this article, the authors examine techniques that calculate the reachability using a matrix or a hyper-graph.

Full story on page 13…

US authorities recover most of Colonial Pipeline ransom

Shortly after the creation of a new Ransomware and Digital Extortion Task Force, the US Department of Justice (DOJ) announced that it had recovered most of the ransom payment made by Colonial Pipeline – although questions remain over how it managed to do that.

Continued on page 2...

Contents

Editorial Office: Elsevier Ltd
The Boulevard, Langford Lane, Kidlington,
Oxford, OX5 1GB, United Kingdom
Tel: +44 1865 843239
Web: www.networksecuritynewsletter.com

Publishing Director: Sarah Jenkins
Editor: Steve Mansfield-Devine
E-mail: smd@contrarisk.com

Columnists: Andrew Cooke, Airbus Security; Karen Renaud; Dave Spence, Context Information Security; Colin Tankard, Digital Pathways

Production Support Manager: Lin Lucas
E-mail: l.lucas@elsevier.com

Subscription Information
An annual subscription to Network Security includes 12 issues and online access for up to 5 users.
Subscriptions run for 12 months, from the date payment is received.

More information: www.elsevier.com/journals/institutional/network-security/1353-4858

12987
Digitally Produced by
Mayfield Press (Oxford) Limited

...Continued from front page

The task force will act as a central co-ordinating body, and all federal prosecutors across the US have been instructed to work with it on ransomware and other extortion cases – such as where attackers are blackmailing victims over stolen data. It will also track attacks and payments. This is part of the US Government's strategy in which cyber attacks are being treated as seriously as terrorism. The hope is that the potential consequences will make attacks less attractive and lucrative for cyber criminals.

The recovery of the ransom paid by Colonial Pipeline, after its operations were shut down by the DarkSide ransomware, is being seen as a major success for the new strategy.

The ransom payment was 75 bitcoins – worth, at that time, around $4.4m (although the value of Bitcoin has dropped rapidly since then). The DOJ said that it was, "able to track multiple transfers of bitcoin and identify that approximately 63.7 bitcoins, representing the proceeds of the victim's ransom payment, had been transferred to a specific address, for which the FBI has the 'private key', or the rough equivalent of a password."

This statement raises as many questions as it answers. How did the FBI come to be in possession of a private key for a Bitcoin account used in the attack? And why is the amount recovered less than that paid?

The answers to both questions could be that the FBI has been successful in taking over the server of a DarkSide affiliate. DarkSide itself operates as a ransomware-as-a-service outfit. It doesn't mount the attacks itself, but provides the infrastructure, malware and tools and takes a cut of the proceeds garnered by its affiliates. The missing 11.3BTC is around 15%, which looks about right for DarkSide's commission. It's therefore possible that the FBI has managed to compromise the affiliate involved in the attack.

When DarkSide recently announced that it was closing down its operations, part of its statement claimed that access to its servers had been shut off and that, "funds from the payment server (belonging to us and our clients) were withdrawn to an unknown account". It's possible that this statement was true and that what DarkSide was suffering was the FBI at work. Due to the drop in value of Bitcoin, the recovered funds are worth around $2.3m.

"Following the money remains one of the most basic, yet powerful tools we have," said Deputy Attorney General Lisa Monaco for the DOJ. "Ransom payments are the fuel that propels the digital extortion engine … the United States will use all available tools to make these attacks more costly and less profitable for criminal enterprises. We will continue to target the entire ransomware ecosystem to disrupt and deter these attacks."

The operation to recover the funds involved the Special Prosecutions Section and Asset Forfeiture Unit of the US Attorney's Office for the Northern District of California, the DOJ Criminal Division's Money Laundering and Asset Recovery Section and Computer Crime and Intellectual Property Section, and the National Security Division's Counterintelligence and Export Control Section – all co-ordinated by the new task force.

Sam Curry, chief security officer at Cybereason, commented that these developments, "have put threat actors on notice, and for the ransomware writers and other malware authors – now the gloves are off. However, this sends a clear message to the criminals: you are not immune to repercussions."

The DOJ statement is here: https://bit.ly/3cpEDV4.

Phishers impersonate USAID

The same group responsible for the recent SolarWinds attack has been busy again, this time taking control of an email account belonging to the US Agency for International Development (USAID) and exploiting it to send out phishing emails.

The compromised email account belonged to USAID's Constant Contact service. As such, the spear-phishing emails sent from it would have appeared entirely legitimate. Several thousand emails were sent to around 3,000 accounts at 150 organisations. Most of the targets were

Threatwatch

Kubernetes malware

Kubernetes clusters are being targeted via malware that affects Windows containers, according to Palo Alto Networks' Unit 42. Dubbed Siloscape, the initial infection vector is via servers – including web servers – and databases with known vulnerabilities. The payload, CloudMalware.exe, tries to perform remote code execution (RCE) on containers and exploits a number of container escape techniques to obtain SeTcbPrivilege privileges. "Siloscape mimics CExecSvc.exe privileges by impersonating its main thread and then calls NtSetInformationSymbolicLink on a newly created symbolic link to break out of the container," explains Unit 42. "More specifically, it links its local containerised X drive to the host's C drive." The malware also uses Tor to connect with command and control servers. There's more information here: https://bit.ly/2T2Lcpy.

PayloadBIN by Evil Corp

A new strain of ransomware, known as PayloadBIN, is actually the work of the notorious Evil Corp group. The naming of the malware may be an attempt to evade US Government sanctions and blacklisting by other cyber criminals. At the end of May, the data leak site used by the Babuk – which had announced its plans to quit ransomware to focus on data theft and extortion – was renamed 'payload bin'. But it seems the new ransomware has nothing to do with Babuk. Evil Corp has hit problems after being sanctioned by the US Treasury Department. This led to ransomware-negotiation companies refusing to handle payments to the gang. It seems the naming of this malware is a ruse to get around that problem. There's more information here: https://bit.ly/3g0Ku5o.

Exploiting VMware

A piece of Python-based, cross-platform malware has been tweaked to take advantage of recently discovered vulnerabilities in VMware vCenter servers. Dubbed FreakOut by its discoverers, Check Point, it was first seen in January and is also known as Necro and N3Cr0m0rPh. Being built around Python scripts, it's capable of running on almost any OS. It exploits app vulnerabilities and brute-forced SSH passwords to enrol infected machines in a botnet, communicating with its command and control servers over IRC. According to Cisco Talos researchers, a new version has updated capabilities for attacking VMWare vSphere, SCO OpenServer, Vesta Control Panel and SMB-based exploits. There's more information here: https://bit.ly/3gdCNaI.

SkinnyBoy from Fancy Bear

The nation-state attack group variously known as Fancy Bear, APT28, Sofacy and other names is using a new piece of malware to target military and government institutions. According to Cluster25, the SkinnyBoy malware is being deployed as part of an attack campaign against foreign affairs ministries, embassies, the defence industry and the military sector in a number of European countries and the US. The malware is a second-stage payload, the first stage being spear-phishing. SkinnyBoy gathers information about victims before downloading further payloads from the command and control server. It's spread via a Microsoft Word document carrying a macro that extracts a DLL file which in turn acts as the malware downloader. There's more information here: https://bit.ly/3wZ9aRl.

RAT by email

A spam campaign, exploiting compromised email accounts, is being used to spread a remote access trojan, dubbed StrRAT by Microsoft. The malware is being used to steal information, log keystrokes and take control of systems but, in some cases, presents itself to victims as ransomware – presumably as a distraction tactic while it goes about exfiltrating data. It appends the extension '.crimson' to files, but doesn't actually encrypt them. The Java-based malware also has the capability to download additional payloads. It's spread via emails that use classic social engineering techniques, such as purportedly being about payments. There's more information here: https://bit.ly/34TQaHI.

government agencies, think tanks, consultants and non-governmental organisations. Most were in the US, but organisations in as many as 24 countries were targeted.

The Microsoft Threat Intelligence Centre (MSTIC) was the first to attribute the attack to the Russian nation state attackers APT29 (aka Cozy Bear or Nobelium). This group is thought to be a branch of the SVR intelligence agency. The Russian Government has denied any involvement in the attack.

"Upon a recipient clicking on a spear-phishing email's hyperlink, the victim computer was directed to download malware from a sub-domain of theyard-service[.]com," the US Department of Justice (DOJ) said. "Using that initial foothold, the actors then downloaded the Cobalt Strike tool to maintain persistent presence and possibly deploy additional tools or malware to the victim's network."

The Cobalt Strike malware also reached out to the domain worldhome-outlet[.]com to receive further payloads. The emails purported to be alerts – some, for example, claimed that Donald Trump had published new documents on election fraud. If victims clicked on a link in the email, an ISO disk image file would be downloaded on their PCs. Clicking on this file would mount the ISO and files contained inside it (LNK and RTF) would execute a DLL that in turn downloaded Cobalt Strike.

"If the device targeted was an Apple iOS device, the user was redirected to another server under Nobelium control, where the since-patched zero-day exploit for CVE-2021-1879 was served," Microsoft said. "The successful deployment of these payloads enables Nobelium to achieve persistent access to compromised systems. Then, the successful execution of these malicious payloads could enable Nobelium to conduct action-on objectives, such as lateral movement, data exfiltration and delivery of additional malware."

The four new pieces of malware consisted of: an HTML attachment named 'EnvyScout', which attempts to steal the NTLM credentials of Windows accounts and drops the malicious ISO; 'BoomBox', which downloads encrypted payloads, gathers data about the local Windows domain and sends it, encrypted, to the command and control (C2) server; a malware loader dubbed 'NativeZone'; and a shellcode downloader and launcher named 'VaporRage', which handles much of the communications with the C2 server.

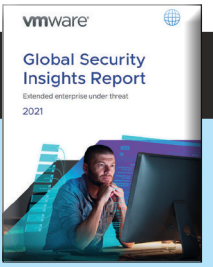
The USAID attack may be just part of a broader campaign. "FireEye has been tracking multiple waves of related spear-phishing emails that have been sent since March 2021," said John Hultquist, VP of analysis at Mandiant Threat Intelligence. "In addition to the USAID content, they have leveraged a variety of lures, including diplomatic notes and invitations from embassies."

The DOJ has now seized control of the two domains used in the attack, although this is unlikely to help organisations that have been infected with malware payloads.

Microsoft has more information on the attack here: https://bit.ly/3pvKtcF.

Report Analysis

VMware Carbon Black: Global Security Insights Report 2021



There has been a common refrain among cyber security practitioners – seen, not least, in the pages of this publication – that the pandemic has forced radical changes on organisations. And not all of them have been up to the challenge.

However, other changes were already in progress before the virus got its grip on the world. Many organisations were in the throes of digital transformation and others were increasingly moving infrastructure, data and applications to the cloud, in one form or another.

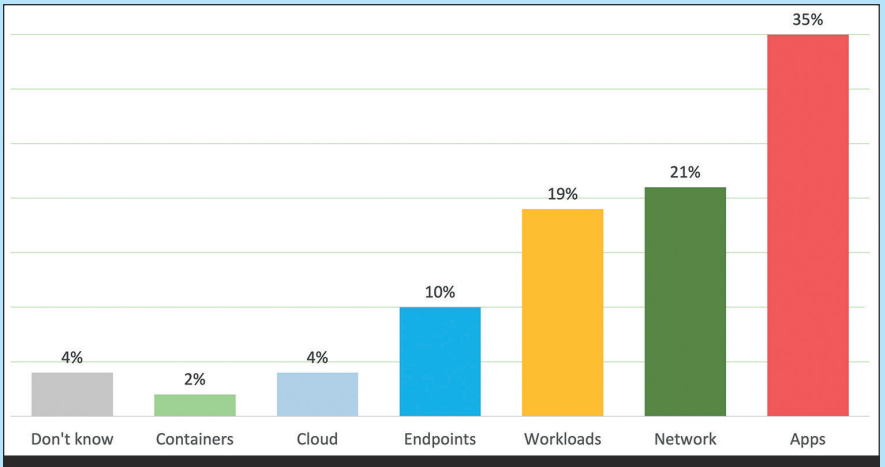
Neither business nor technology stands still. Firms need to adapt constantly to changing market conditions and the opportunities offered by emerging technologies, with the cloud being among the most prominent. Alas, as we know only too well, improvements and transformations in security often lag well behind, seen as an afterthought. But with Covid-19 having shone the spotlight on the potential pitfalls opened up by the forced adoption of practices such as working from home, has this been a wake-up call for organisations to shake up their attitudes to security? Well, the opportunity is there.

VMware's polling of more than 3,500 CIOs, CTOs and CISOs found that nearly four-fifths of organisations had experienced some form of cyber attack directly due to staff working from home. But the company attributes this problem to organisations attempting to stretch their existing security technologies, policies and procedures to fit the new normal. This approach has proven to be a poor fit for what amounts to a radically different infrastructure.

The overall proportion of organisations hit with a data breach is similar to the home working issue – in this case, just over four-fifths. And of those, 82% were 'material' breaches – ie, significant and involving the loss of valuable or sensitive data. And yet, organisations seem to be in denial. Only 56% of security professionals admitted to being worried about a material breach in the coming year and only 41% have updated their security policies and approaches in order to mitigate the risk.

Pretty much every organisation has moved some part of its operations to the cloud and this has increased the attack surface as well as the number of vulnerabilities that are poorly understood (if at all) by security practitioners. Getting on for two-thirds (61%) of security professionals say they understand that they need to view security differently in this kind of environment, although exactly what they're doing about it is another matter.

Actually, one thing it's clear they are doing is putting their faith in the cloud for security as well as business transformation. Pretty much all of those polled (98%) either use or plan to use a cloud-first security strategy. And that's fine, but that too opens some weak spots. Let's take the example of anti-malware products. These all now use cloud-based heuristic or behavioural analysis to counter the threat of zero-day malware or patch lag. But someone working at home may have intermittent access to the Internet. There's an opportunity here for malware that pauses before doing its dirty work.



What organisations believe are the most vulnerable breach points. Source: VMware.

Ransomware is one particularly pernicious form of malware that has been known to bide its time before springing into action. And VMware's survey found that it is the most common cause of a breach (14%), alongside third-party software – the sort of thing a home worker might install on a work laptop at home. Generally, attacks have become more sophisticated, says the VMware report, with cloud-based attacks being the most common.

The report also finds that applications and workloads are what are keeping CISOs awake at night. Nearly two-thirds (63%) reckon they need better visibility over data and apps to be able to thwart attacks. This is complicated by the cloud because data is moving so frequently between users and off-premise storage. And, of course, many of the applications are now off-premise. It's hard to feel that you're in full control when your precious apps and data are on someone else's computer.

Forty three per cent of security professionals said that they intend to build more security into their infrastructure and apps and reduce the number of point solutions. That's a good move, but one can't help wishing the percentage was significantly higher. Certainly, point solutions, their lack of integration and the visibility problems they create are problematic. With the move to the cloud, the effectiveness and importance of on-premise solutions – the traditional firewalls, intrusion detection systems et al – have been diminished in the eyes of many. But it's crucial not to lose sight of the fact that important data still exists within the perimeter and on endpoints, especially when it's being used.

"The race to adopt cloud technology since the start of the pandemic has created a once-in-a-generation chance for business leaders to rethink their approach to cyber security," says Rick McElroy, principal cyber security strategist, VMware. "Legacy security systems are no longer sufficient. Organisations need protection that extends beyond endpoints to workloads to better secure data and applications. As attacker sophistication and security threats become more prevalent, we must empower defenders to detect and stop attacks, as well as implement security stacks built for a cloud-first world."

The report is available here: www.carbonblack.com/resources/globalsecurity-insights-report-2021/.

In brief

Australian backdoor

The Australian Federal Police (AFP) and the FBI in the US have worked together on a sting operation in which a phony 'secure' communications app was promoted to, and adopted by, members of organised crime groups. Operation Ironside began around three years ago. An app named ANOM was developed by the FBI and loaded onto specially modified phones that could not make calls or send emails outside the app. Calls could be made only from one ANOM-equipped phone to another. These phones were then sold to known criminals and quickly became popular on underground markets, where it was believed they were entirely secure, although they were actually being monitored by a number of law enforcement agencies in multiple countries. As a result, in overnight raids, 4,000 AFP officers in Australia made more than 200 arrests, issued over 500 warrants, seized AU$45m in funds and 3.7 tonnes of drugs and allegedly prevented the murder of a family of five. It's expected that the FBI and Europol will soon release details of similar raids.

VMware flaw under attack

A critical remote code execution (RCE) vulnerability in VMware's vCenter product (CVE-2021-21985) is being actively exploited by attackers. VMware has released an update, but many servers remain unpatched and vulnerable. The bug is considered critical, with a CVSS score of 9.8, and, according to VMware: "A malicious actor with network access to port 443 may exploit this issue to execute commands with unrestricted privileges on the underlying operating system that hosts vCenter Server". In fact, attackers could use it to take over an entire network. Proof-of-concept exploit code has been released by researchers and threat intelligence companies report active scanning by attackers for this flaw. VMware has an advisory here: https://bit.ly/3x6mUK5.

MQTT bugs

Flaws in three popular message brokers could result in denial of service (DoS) attacks against Internet of Things (IoT) devices. The Message Queuing Telemetry Transport (MQTT) message brokers are designed to receive, store and forward messages between devices and apps. If these messaging hubs are taken down, 'smart' devices will simply stop working. These could include domestic devices such as smart bulbs and door locks as well as industrial solutions. The three pieces of software – RabbitMQ, EMQ X and VerneMQ – can be tricked into filling available memory, at which point they will fail. All three have been patched and users are encouraged to update as soon as possible.

Fortinet flaw exploited

The FBI has been prompted to issue a warning about a flaw in a Fortinet product after it was exploited to compromise the website of a US municipal government. The attack took advantage of an SSL VPN vulnerability (CVE-2018-13379) in Fortinet appliances running FortiOS. The FBI and the US Cyber security & Infrastructure Security Agency (CISA) warned about this problem and others affecting the same devices, back in April. But it seems that many users – which are mainly government agencies and major organisations – have been slow to patch. And the FBI believes that numerous attacks are underway. "Access gained by the APT actors can be leveraged to conduct data exfiltration, data encryption or other malicious activity," said the FBI. "The APT actors are actively targeting a broad range of victims across multiple sectors, indicating the activity is focused on exploiting vulnerabilities rather than targeted at specific sectors." The FBI warning is here: https://bit.ly/3wacyZK.

Transport network compromised

North America's largest transport network, New York's Metropolitan Transportation Authority (MTA), was penetrated by attackers working for China's Government, according to a report in the *New York Times*. Citing an MTA internal report, the article says that the hackers did not gain access to systems that controlled trains and that rider safety was never put at risk. The MTA report also insisted that no customer data was compromised. However, the attackers managed to achieve persistence on the systems for several days, gained access to three of MTA's 18 computer systems and may have exfiltrated sensitive information. And it's the third time the MTA's systems have been breached. The *New York Times* report is here: https://nyti.ms/3w5Vhk7.

RDP abused

The remote desktop protocol (RDP) was exploited in 90% of attacks investigated by Sophos in the past year, the firm says, with the breaches leading to ransomware infections in 81% of the cases. In a new report, 'Active Adversary Playbook 2021', the firm said that RDP is frequently abused as the first step in gaining a foothold on targeted systems. It is also exploited in 69% of cases to move laterally across networks. Gaining access this way effectively renders protections such as VPNs and multi-factor authentication useless, claims Sophos. Its report is here: https://bit.ly/3g0UnzW.

Credential-stuffing surge

Akamai says it saw no fewer than 193 billion credential-stuffing attacks in 2020. With so many user databases having been leaked – not least thanks to the huge number of people now online – cyber criminals have access to large lists of credentials. They are using these to attempt penetration of all kinds of systems, including social media platforms and enterprises. In its '2021 State of the Internet/Security' report, Akamai said the financial sector was the most heavily targeted with this kind of attack, with 3.4 billion credential-stuffing attempts, a 45% increase over the previous year. Of the nearly 6.3 billion web application attacks in 2020, more than 736 million were aimed at financial services organisations – an increase of 62% from 2019. The most common form of attack against websites was local file inclusion (LFI), which accounted for 52% of the total, followed by SQL injection (33%) and cross-site scripting (9%). The report is available here: https://bit.ly/353XNeI.

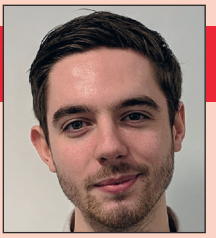
Crypto-currency raider linked to North Korea

An operation focused on compromising crypto-currency exchanges in order to steal funds is now being linked to North Korea's nation-state hacking group, Lazarus. The CryptoCore operation, first identified by ClearSky last year, has stolen hundreds of millions of dollars-worth of crypto-currency from exchanges in the US, Israel, Europe and Japan over the past three years. The group uses spear phishing as its initial attack vector. ClearSky thought the group was based in Eastern European countries – most likely Ukraine, Russia or Romania. However, additional research by other cyber security companies and organisations – most notably F-Secure, Japan's Computer Emergency Response Team (JPCERT/CC), and NTT Security – have switched the focus to North Korea. The communist state is known for using cyber attacks as a way of raising revenue for the country's treasury. ClearSky has said it agrees with this analysis and has issued a new report, available here: https://bit.ly/34XrZIL.

Massive password list

A new password list is circulating on a hacker forum. The 100GB text file is alleged to contain 82 billion passwords aggregated from data breaches over the years, although an analysis by CyberNews put the number closer to 8.4 billion. All the passwords have 6-20 characters and have had white space removed. The list has been given the name RockYou2021 – a reference to the RockYou data breach in 2009 which yielded 32 million passwords in plain text. The original RockYou list is still used today by both cyber criminals and security practitioners for brute force credential attacks. A later wordlist – the Compilation of Many Breaches (COMB) – is also popular, as it contains 3.2 billion passwords. It's likely that this has been amalgamated into the RockYou2021 list. CyberNews is providing a password checker where you can see if your passwords are in the list. It's here: https://bit.ly/3gmuWYz.

Threat Intelligence



Analysing three new complex phishing tactics

Tom McVey, Menlo Security

Phishing is one of the most-used terms in the cyber security dictionary. It is defined as the fraudulent practice of purporting to be a reputable or known entity to trick individuals into infecting their devices or revealing sensitive information.

The phishing spectrum is a broad one. While, arguably, the most widespread technique is generic email phishing, attempts and techniques have expanded greatly in recent years to include targeted spear-phishing, SMS (smishing), voice call (vishing) and website phishing.

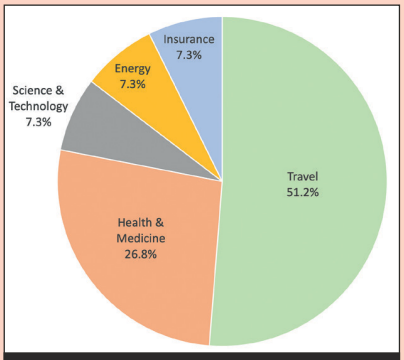
But what do cyber attackers have to gain?

From our email and bank accounts to accessing our doctors' records online, the ever-expanding number of digital services that continue to be deployed in the aim of making our lives easier all have one thing in common: to access them, we need to provide a defined set of credentials.

Be it our email addresses, user IDs, passwords or pin codes, the value of obtaining these credentials is immense. In doing so, perpetrators may successfully execute identity fraud, data breaches, ransomware and more – threats that all have the potential to lead to extremely damaging outcomes.

Indeed, generic scams can be easy to spot. The challenge today, however, is that many attacks are expanding, not only in terms of volume, but equally in the way of complexity.

In March 2021, Menlo Labs observed a distinct rise in credential phishing attacks that targeted a variety of different channels where many attackers created sophisticated fake login pages and forms impersonating commonly used corporate services.



Office 365 phishing campaign targets.

The bulk of these focused on Outlook and Office 365 login pages – somewhat unsurprising given Gartner's outlook showing the Microsoft productivity suite to have an 87.5% market share in 2018 (https://cnb.cx/3fPRWyE). Our observations showed that players in the travel industry accounted for 51.2% of all Office 365 phishing campaign targets for the month, owing to the niche targeting of airline duty-free shop login credentials. The remainder beyond this was shared between the health and medicine industry (26.8%), science and technology (7.3%), energy (7.3%), and insurance (7.3%).

While the Microsoft Office 365 suite might be the most natural phishing vector, with an estimated 145 million users as of April 2021, it is not only one. There was also a distinct uptick in phishing pages impersonating popular cloud services like Microsoft Azure, OneDrive, Firebase, Box and Dropbox, and, most recently, Evernote.

The expanding volume of attacks across a widening spectrum of platforms is just one part of the challenge, however. Attackers are likewise becoming increasingly intelligent in the ways they create and deploy phishing scams, with many modern techniques capable of bypassing traditional detection solutions.

The use of data URLs and encoding to mask content is a prime example. Here, the JavaScript code that logs submitted credentials in a remote URL is hidden, and all custom cascading style sheets (CSS) and images are embedded on the page. As a result, entire phishing pages can be rendered in a single load with no additional resource requests relating to the JavaScript, CSS or image, allowing them to evade detection solutions reliant on the content-type header.

Dynamic content generation, particularly relating to Microsoft Office 365, is another technique that has advanced, with threat actors now able to append a user's email address on the URL. Here, a phishing page path can be dynamically generated to show an automatically filled user's email address within a fake login box. Signature-based security systems rely on filename and/or file path patterns, yet this use of dynamic content generation uses .php files to ensure random generation that bypasses detection. Further, APIs such as Clearbit can be used to dynamically load brand logos – a technique that drastically enhances spear-phishing tactics.

A third such tactic we observed came from the use of local HTML and PDF decoy files to load phishing content. Our study witnessed the impersonation of Daum – a popular web service provider in South Korea – where visiting the phishing landing page would trigger the download of a decoy HTML file. Once the local HTML file is opened, a phishing form is loaded with the filled-in username, thanks to the email having been appended to the URL as a parameter.

With decoy systems able to run content locally on the client machine without retrieving information from a server, they are once again able to slip through many outdated detection nets.

Yes, we have seen an uptick in the use of decoy phishing, dynamic content generation and data URLs and encoding, but these are just three small droplets among what is an incredibly sophisticated and rising tide of credentials phishing tactics.

What's the solution? Unfortunately, there is no single silver bullet, but user awareness is the most important piece of the puzzle to address.

With 95% of all cyber attacks stemming from human error, all organisations should work to educate their employees to promote better practices and stamp out the opportunity for phishing attacks to succeed.

Applying the principles of zero-trust architecture to protect sensitive and critical data

David Greenwood

David Greenwood, ISN Solutions

In today's digital landscape, data is a company's most valuable asset, and energy companies, particularly electric, oil and gas, remain at risk of hacking attempts due to their major social and economic importance. Using key zero-trust principles can help to ensure that IT systems are protected, mitigating the risk to company operations and sensitive and critical data.

Over the course of the pandemic, businesses of all genres and sizes within the oil and gas sector have faced increased risk of a network breach. Reported cases from UK businesses as a whole rose by 20% in 2020, while Hiscox's 'Cyber Readiness Report' revealed that the energy sector bore the highest burden for financial losses as the result of a breach.[1] This, coupled with the continuation of widespread remote working, has accelerated change in IT infrastructure design, including increased implementation of zero-trust security models.

Never trust, always verify

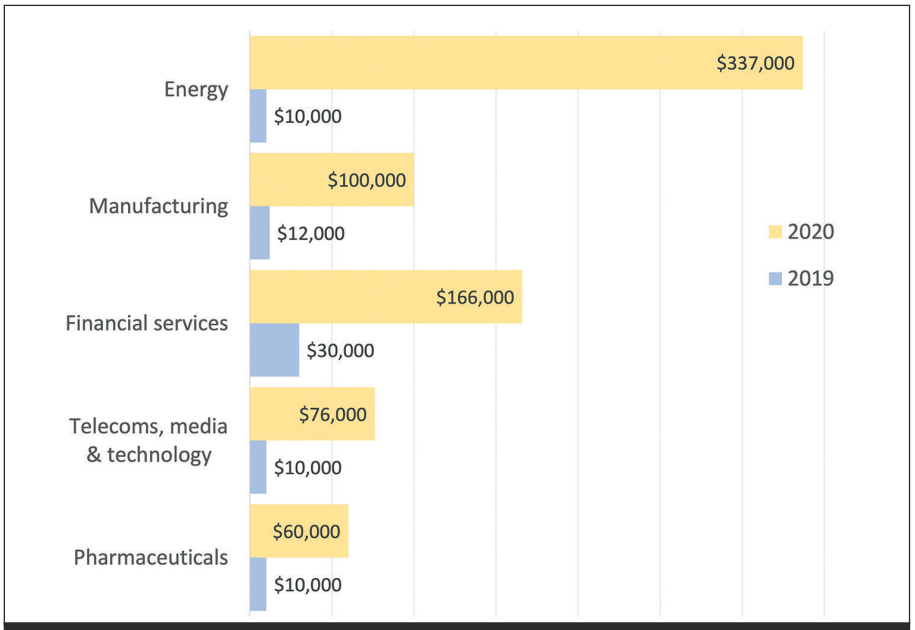
The underlying ethos to zero-trust security models is the assumption that any attempt to access the company network is a potential breach. In comparison, outdated 'trust but verify' security models assume that the user should be granted access, but ask for verification (usually single factor authentication such as a username and password) 'just in case'.

'Trust but verify' security models pose potentially costly risks, as threats such as trojans can capture login credentials, even if data is hosted in the cloud, and are no longer a valid defence against modern cyberthreats.

The zero-trust security model takes verification multiple steps further. The model incorporates stringent security protocols, with multi-factor authentication as a minimum, as well as inspecting and logging all traffic. Access requests originating on a local area network (LAN) are treated with the same level of suspicion as if they had come from a wide area network (WAN), which in IT security terms is analogous to the Wild West. This is because the need to defend against threats from inside companies is being increasingly recognised.

Hackers, meanwhile, are turning to bribery to access systems and disrupt operations and, even as long ago as 2012, in the case of oil and gas company EnerVest, a disgruntled employee was able to sabotage company systems, which resulted in extensive disruption to business operations for well over a month.[2] In today's ever more aggressive cyber environment, the threats are even greater.

Security models that analyse additional factors beyond user credentials, such as the user's location, device and access habits, and are able to spot anomalies, can more reliably ascertain whether users are who they claim to be or whether there is a breach. This enables quicker response to potential threats, and the quicker a response to a breach, the easier it is to limit the damage. This can also help companies to react to insider



Sectors bearing the brunt of cyber attacks – median losses. Source: Hiscox.

threats if the model is designed to flag unusual user activity or login times.

Additionally, in zero-trust security models, user access to data is controlled by what is termed as 'just in time' (JIT) and 'just enough access' (JEA) principles. These ensure that employees can access the data they need to stay productive, but other sources of data and areas of the network are restricted, to limit the scope of damage from a successful hacking attempt or malware or ransomware infection, thus preventing infection spreading across the rest of the entire network and all devices.

Verification that works

Usually, one of the first steps taken by companies to implement a zero-trust security model is to enable multi-factor authentication (MFA). But it should be ensured that the process is accepted by users and any impacts in terms of costs in employee time and productivity are avoided.

As an example, using biometric identifiers (such as facial recognition or fingerprint scan) as part of the verification process has become far more viable now that new models of laptops and mobile phones often include features like facial recognition, or are capable of supporting it. And combining biometric identifiers with login credentials can be quicker and more user-friendly than combining credentials with a one-time password (OTP) sent to another device. However, not all employees may be comfortable using biometric identifiers.

If identification is linked to a device or object, such as a mobile phone or network key, it may be worth considering how much employees commute or travel for work. As physical objects can be lost or stolen, more movement amplifies the risk and, for employees who work abroad and/or offshore, issuing replacements may be difficult and costly.

MFA can be combined with single sign-on (SSO) to simplify username and password management for users and administrators. This can save employees having to remember multiple passwords. It also results in savings on time costs and avoids productivity hits for companies that outsource IT support, thanks to a reduction in the number of password reset requests. But the biggest benefit to enabling SSO is to help reduce the risk of data theft from an inside source.

Inside threat

Cyber security industry analyst Forrester forecast in its 'Predictions 2021: Cybersecurity' report that one third of security breaches in 2021 will be caused by internal incidents, both accidental and malicious.[3]

As it becomes ever easier to move data to an external drive or to a cloud location, the risk of data loss and theft has increased. In addition, during the pandemic, as many employees continue to work from home, there is less scrutiny to mitigate data theft, and even less chance of companies knowing about it.

And it's not just data theft that is a point of concern for the energy industry. Breaches to control systems or safety-critical systems and changes to software could result in devastating – potentially even life-threatening – consequences, in the worst-case scenario. And if the breach comes from an inside source, this could be even more dangerous as the activity may not be noticed as unusual for some time or until a problem is discovered.

Implementing SSO may contribute to helping to limit damage in these instances, as it offers administrators the option to only remove or freeze one set of credentials in the event that an account needs to be immediately locked down. This also negates the risk of some account credentials being overlooked, either in the event that a hacker could have gained access to other credentials or part of credentials to a multitude of systems and applications, or if an employee has simply left the company and all accounts need to be closed.

The drawback to SSO, however, is that there may be some instances of access to sensitive data that requires another layer of protection. In these cases, OTPs could be sent to a verified device.

Company culture

The principle of 'never trust, always verify' shouldn't just be applied to IT systems design, but should be fully integrated into company culture. Some scams and attempts at data theft, such as business email compromise (BEC) scams or malware posing as system or browser updates, rely on social engineering tactics rather than forcing entry to the company network. These can be particularly difficult to defend against as they are triggered simply by human error.

As theft of information is one of the primary motives for attacks against energy companies, BEC represents a potentially serious threat. Hackers may try to obtain company plans on mergers, acquisitions or bidding strategies to sell on to a competitor, for example.

"Some scams and attempts at data theft, such as business email compromise (BEC) scams or malware posing as systems or browser updates, rely on social engineering tactics rather than forcing entry to the company network"

In this scam, hackers spoof an email domain to very closely match the email address of a company's CEO or senior management, and distribute an email asking for sensitive information to company employees. The scam relies on employees being too busy or stressed to properly examine the sender's email address or register the request as unusual, and so the potential of these scams succeeding has grown due to the stress of the pandemic. Falling victim to a BEC attack can have wide-ranging consequences: as well as sensitive company information being passed to competitors, there is also the potential for cancellation of business deals and loss of revenue, reputation and customers.

Secure all locations

The golden rule of IT security is that there should be no single point of failure, which includes ensuring that no unvetted, unsecured devices can access the network. This is vital to the oil and gas sector in particular as it continues to undergo digital transformation and equip employees in offshore and remote locations with small, portable devices to access critical company data.

The problem with introducing these devices to company networks is that it's often forgotten that they can create an entry point for hackers. The most obvious threat is ransomware, which has the potential to spread quickly from one infected device across an entire network. It can take weeks until companies are able to resume operations as normal after a ransomware attack, and the cost of a ransom is likely to be high, especially now that ransomware operators are deploying two-stage attacks, demanding a second ransom with the threat of publishing sensitive data online.

The potential cost of ransomware attacks today goes beyond severe financial loss, including damage to existing customer relationships and severe difficulty attracting new business, especially if it's discovered that customer data has been published online.

While enabling multi-factor authentication on VPN services and having endpoint security solutions can help to prevent ransomware operators from accessing the network, the key defence against ransomware is to be prepared for it. This means proper backing up of data in multiple locations, which can help companies to avoid paying a ransom and to resume business operations more quickly, and ensuring that there are no unsecured endpoints that leave an easy way in for hackers.

Careful planning

Switching to a zero-trust security model often requires careful planning to ensure productivity and that access to data needed for daily work is maintained. Additionally, one of the main issues that can arise with this type of security model is the need for ongoing administration. While, like any aspect of IT security, it's crucial to avoid systems falling into obsolescence, the level of administration required to maintain a zero-trust system can be high for large companies with hundreds of employees. As people leave jobs and change role within companies, this may require new permissions.

If need be, companies can partner with an IT MSP specialising in network resilience and security to advise on, or assist with, the implementation and required ongoing maintenance of new security architectures and protocols.

About the author

David Greenwood is an IT specialist and CEO of IT services and support company ISN Solutions (www.isnsolutions.co.uk), where he has led the company in servicing the energy sector, working in offshore, remote and challenging environments for the past 21 years. Previously, he built a career as an information technology specialist and consultant, working with upstream oil and gas companies.

References

1. 'Hiscox Cyber Readiness Report 2020'. Hiscox, Oct 2020. Accessed Jun 2021. www.hiscoxgroup.com/sites/group/files/documents/2020-06/Hiscox-Cyber-Readiness-Report-2020.pdf.
2. 'Vengeful EnerVest Operating network engineer pleads guilty to intentionally damaging computer system'. DataBreaches.net, 29 Jan 2014. Accessed Jun 2021. www.databreaches.net/vengeful-enervest-operating-network-engineer-pleads-guilty-to-intentionally-damaging-computer-system/.
3. 'Predictions 2021: Cybersecurity'. Forrester, 26 Oct 2020. Accessed Jun 2021. www.forrester.com/fn/1fxTKia7eYFcDCZrieHlwX.

Grid cyber security: secure by design, continuous threat monitoring, effective incident response and board oversight

Dale Geach

Dale Geach, Siemens UK

Grids are possibly the highest expression of engineering serving society. From the aqueducts created in the sixth century BCE to the modern era birth of gas utilities in the 1820s, followed by electricity generation and distribution grids in the 1880s and the computer information-sharing age kicking off in the 1960s, we have been creating distribution infrastructure.

By their very nature these grids are vulnerable – miles of pipes and cables span the globe, uncountable valves and switches control flows; the flick of a

lever or switch, the severing of lines and blocking of pipes can shut down or shut off parts or maybe all of a network.

This century has seen the creation of remote control on a massive and still growing scale. It has also seen the evolution of the cyber attack. This has raised the level of attack from the physical assault on infrastructure to the viral invasion of the systems that control the grids. The past decade has also seen the new phenomenon of the reversal in the historic movement towards central generation and distribution control to the advent of microgrids, local production of renewable energy and grid-edge processing of information. The outcome is that we have lots of infrastructure, spanning generations of engineering and technology, still in service adapting to concurrently evolving sophistication in the methods used to attack our grids.

Cyber attackers

Who wants to attack our grids and why? The prime movers are nation states, as actors or sponsors, followed by terrorists, wishing to overthrow nation states, and then criminals seeking financial gain. Nation states top the bill because they have the resources – financial, time and manpower – to plan and execute significant and high-impact attacks. Their intent is generally to disrupt supply, and thereby disrupt social or economic activity; destroy capability or capacity and thereby postpone the development of assets which may give competitive or military advantage; or maybe just to see what is possible, for the furtherance of military or economic strategies for future defence or aggression.

As per the publicly available reports, the Stuxnet and BlackEnergy cyber attacks, respectively, were presumed to have been executed or sponsored by nation-state actors.[1,2] The BlackEnergy attack caused widespread power outages in 2015 and the attackers notably were in the power systems for at least six months, undetected, before manifesting, raising the question of other undetected 'time-bombs' waiting to appear or that maybe failed to manifest.

In 2017, the NotPetya destructive cyber attack masqueraded as ransomware, but its purpose was principally to disrupt. In the same year, the WannaCry ransomware attack disrupted services/operations globally across multiple sectors, including health and utilities, hitting more than 300,000 computers in 150 nations, causing billions of dollars of damage.

System weaknesses

The examples above highlight different aspects of the cyber attacks on grids. The Stuxnet attack identified one piece of equipment and found an entry point through which a command could be sent that made centrifuges, essential to the development and enrichment of nuclear material, spin so fast that they self-destructed.

"What gives rise to the potential for these types of attack is the significant increase in attack surface created by the introduction of Ethernet connectivity to operational technology (OT), allowing components to communicate with each other and ultimately to IT systems and the Internet"

BlackEnergy, the first publicly acknowledged cyber attack to result in power outages, was more complex – a multi-layered attack that started with identifying individual staff and targeting them with very personalised spear-phishing emails carrying malicious Excel documents with macros to infect the IT system. Patience eventually gained illegal entry into the company's IT system, where the attackers harvested credentials and information to move from IT to the OT side of three energy businesses, the ICS/SCADA system, where passwords were changed to prevent anyone from stopping the attack. The attack then targeted devices at substations with malicious firmware to make the devices inoperable. Finally telephone systems were used to generate thousands of calls to the energy company call centres in a denial-of-service attack to prevent customers reporting outages.

NotPetya, a malicious data encryption tool inserted into a legitimate piece of software used by most of Ukraine's financial and government institutions, spread rapidly via trusted networks, and displayed a ransom note asking for payment in Bitcoins. As the ransom notes did not display an identification ID that would enable the attacker to know whose data to decrypt, there was no means for victims to recover data once it had been encrypted, making it more accurate to describe the attack intent as destructive.

This latter attack is part of a number of well understood and prevalent attacks on IT systems, which, intentionally or by chance, can affect utilities. Stuxnet and BlackEnergy marked a new departure as they were specifically targeted and crafted to attack OT systems within the grids. What gives rise to the potential for these types of attacks is the significant increase in attack surface created by the introduction of Ethernet connectivity to operational technology (OT), allowing components to communicate with each other and ultimately to IT systems and the Internet.

Exposed systems

It used to be the case, and in too many instances still is, that OT systems were not considered as exposed to cyber attacks because they were not connected to the Internet. Two faults arise in this thinking: as BlackEnergy showed, OT frequently reports to, or is monitored by, the IT system and so can be attacked via an IT breach. Grid operators are now seeking the huge benefits of modernising and digitalising grid operations. Leveraging industrial Ethernet and IP-based communication and international standards for substation automation, such as IEC 61850, are delivering benefits and enabling a range of new opportunities in support of:

- Real-time situational awareness of the grid state
- Load prediction and shift.
- Rapid fault isolation and recovery.

- Integration of renewables and distributed energy resources (DER).
- Achieving net-zero carbon emissions targets.
- Reducing employee health and safety risks.
- Enabling remote maintenance and monitoring capabilities.

And these are just to name a few. The decentralised model enabled by smart grid technologies provides greater flexibility in the event of disruptions as well as potentially reducing the time needed for recovery, as it might be possible to localise disruptions. However, these grid modernisations expose the grids to greater cyber risks due to increased connectivity and inherent vulnerabilities in legacy systems.

Old designs

Much of the grid's infrastructure was designed before the advent of cyber attacks and specifically before the advent of the higher levels of connectivity for OT. They were created for a non-connected world. So, for example, Modbus, one of the commonly used protocols in supervisory control and data acquisition (Scada) environments since 1979, for remote monitoring, control, and data acquisition, is inherently insecure because security was not a concern when it was designed and it does not connect to the Internet directly.

Legacy ICS/OT have inherent security issues, such as use of insecure protocols and the inability to enable secure passwords, which create greater cyber security challenges. Historically, the security of OT systems in a utility or manufacturing setting has been relegated to maintenance teams and physical security, which don't have the capabilities or expertise or real-time visibility into their ICS/OT networks to proactively identify vulnerabilities or detect malicious activities and prevent incidents.

Initial access to a targeted organisation's ICS/OT system is still usually gained via exposed IT systems, or by exploiting a supplier's network, and people remain the weakest link, mainly unintentionally but also with malicious intent by the individual or inducements or pressure from third parties.

Secure by design

It is very important to first understand the threat landscape and prioritise mitigations based on risks and impact to critical systems rather than trying to implement everything in one shot. Effective risk management depends on the maturity of three core capabilities – secure by design; continuous threat monitoring with effective incident response; and governance and board oversight.

Intelligence agencies and governments in the US, the UK, the EU and elsewhere have identified cyber security of CNI as one of the top threats to their national, social and economic security. The EU introduced the Network and Information Security (NIS) Directive and, in a move to protect CNI, President Donald Trump signed an executive order banning US grid operators from buying and installing electrical equipment manufactured outside the US.[3,4]

"Historically the security of OT systems in a utility or manufacturing setting has been relegated to maintenance teams and physical security"

As a best practice, ICS/OT networks should be segmented from business/office IT networks but improper segmentation, design flaws and inadequate security controls can allow an attack to pivot from IT to the ICS/OT network. However, in the ICS/OT world, we must be aware that increasingly segmenting an existing OT network may not always be feasible or practical. With OT, unlike IT, installing patches automatically is often not an option; many of these systems run 24/7 and reboots to install a security patch or new firmware version would unacceptably interrupt operations.

For many systems still in use, and likely to be for some time, patches are no longer available. They are often provided only for some 15 years, a long time in IT and a brief interlude in OT. For those running Windows XP and Server 2003, support has not been available for some time, support for Windows 7 ended in January 2020 and Niagara AX support is being phased out now. However, patching would not solve all problems anyway because many attacks do not involve a zero-day, instead exploiting legitimate, standard product design features and inherent design flaws. Most current industrial protocols and products lack basic authentication features and integrity verification. The goal should be to take a balanced approach by limiting the attack surface and using secure configuration and hardening, minimising exposure and patching wherever feasible and required.

Living with vulnerabilities

To a certain extent, we must accept that OT/ICS systems were not originally designed with cyber security in mind, so we must live with inherent vulnerabilities in legacy systems for some considerable time to come. Cyberthreats are a new reality in today's hyperconnected world.

For effective cyber security measures, international standards and best practices such as IEC 62443, the NIS Directive, NERC CIP, NIST CSF, Mitre ATT&CK framework for ICS, and so on may be very helpful. However, the situation is often confused by standards that contradict each other or conflict. Governments and international authorities need closer cooperation and communication to ensure that the guidance issued to grid operators is both consistent and readily available to all who need them.

In a realistic world, there is nothing even close to 100% protection, so organisations need to be pragmatic in their approach. Identification of attacks in prospect and in progress, coupled to an ability to stop the attack are part of the answer. Resilience to recover from the attack is the other part of the response.

Threat monitoring

For safety and operational reasons, ICS uptime requirements are extremely high

and any intrusive security technologies (such as intrusion prevention systems) that could accidentally impact operations are unlikely to be acceptable. Use of passive and detection technologies such as AI, machine learning (ML) and big data analytics can be game changers by improving the detection of hidden threats.

AI techniques can develop data models, monitor patterns and trends and identify anomalies using baseline data models. Behavioural pattern monitoring can improve threat intelligence and prediction, enabling faster attack detection and response by making it possible to process and analyse vast quantities of data – with parsing, filtering and visualisation done in near real time.

Big data analytics also provides a critical step towards improving cyber security capabilities and managing cyber risk more efficiently and effectively. For example, in the Ukraine power grid cyber attacks, AI/ML could have been used to detect zero-day attacks, using anomalies, abnormal user behaviours, abnormal network traffic and connections, as well as spotting abnormal hours and user-to-device connections to predict potential attacks. Identifying and detecting malicious activities within massive amounts of data is possible only by using AI/ML.

Incident response

People are at the heart of cyber security, the risk landscape is continuously changing so strategy and plans must evolve and be adjusted to keep pace. Cyberthreats are inevitable and no matter how robust an organisation's security posture is, it is still susceptible to zero-day and sophisticated attacks. Therefore, in addition to having robust security design and risk-management practices, organisations must prepare for emergency situations and always have a tested incident response plan in place and a team of experts trained and ready to handle incidents.

Cyber security measures are more than technologies and processes. The first action is to identify the nature of the incident and be clear whether it is a cyber attack. There are lots of things that can go wrong without any outside agency. Many response plans still have a strong bias towards loss of data – they ensure that the plan has a proper consideration of an attack on the OT side. Organisations need to recognise that isolating or limiting the incident may require shutting down critical processes, so they must ensure that access is always available to a number of appropriate authorisers for such decisions. And they need to be absolutely clear about the skills that may be needed to deal with an incident and that external support is available in depth if the skills or resources are not available in-house, particularly in respect of older equipment.

Board oversight

An effective cyber security programme requires board and executive management support and leadership, as well as a security culture across the organisation, built upon proactive risk management and ability to recover in a predictable manner from a cyber attack. The chief information security officer (CISO) or chief risk officer (CRO) should be reporting to and periodically updating their boards about cyber risks and their preparedness. Cyber capabilities should be realistically assessed, and the organisation's security posture compared to its risk appetite and industry peers.

As well as looking to their own security, grid operators need to positively vet their supply chains to ensure that equipment is compliant but also that staff are fully trained in awareness and prevention protocols. This must particularly apply to external suppliers who have online access into grid systems for remote installation, maintenance or upgrading.

Conclusion

As we transform our traditional grids to modern or smart grids, we will have to deal with hybrid OT/ICT environments containing both the most advanced and variably aged legacy components. People remain the greatest vulnerability, so train, retrain and reinforce.

Network segmentation may not be easy and patching of legacy systems may not be realistic, so we must live with inherent security vulnerabilities. To counter cyberthreats to grid networks and CNI in general, organisations across all critical sectors, and governments need to partner to secure their CNI without creating an environment of fear and panic. Boards must be held accountable for the security of their critical infrastructure and regulatory agencies must empower and incentivise organisations to continuously enhance cyber resilience.

Only a secure-by-design network with continuous threat detection, a tested emergency response plan in place and an enterprise risk management programme with board oversight can defeat adversaries and defend national, social and economic security.

About the author

Dale Geach is technology and innovation manager at Siemens. He has worked in the energy sector for more than 26 years, holding a breadth of roles in design, testing and delivery of digital technology for critical national infrastructure (CNI), more specifically power grids. As head of technology and innovation within Siemens Smart Infrastructure, Geach is responsible for portfolio development, which serves not only the current needs of industry, but also focuses on the technology and service needs of a resilient future grid.

References

1. 'Stuxnet'. Wikipedia. Accessed Jun 2021. https://en.wikipedia.org/wiki/Stuxnet.
2. 'BlackEnergy'. Wikipedia. Accessed Jun 2021. https://en.wikipedia.org/wiki/BlackEnergy.
3. 'NIS Directive'. ENISA. Accessed Jun 2021. www.enisa.europa.eu/topics/nis-directive.
4. Miller, Maggie. 'Trump issues executive order to protect power grid from attack'. The Hill, 1 May 2020. Accessed Jun 2021. https://thehill.com/policy/cybersecurity/495711-trump-issues-executive-order-to-protect-us-power-grid-from-attack.

Attack graph reachability: concept, analysis, challenges and issues

Zaid J Al-Araji, Sharifah Sakinah Syad Ahmed, Raihana Syahirah Abdullah, Ammar Awad Mutlag, Hayder Adil Abdul Raheem, Siti Rohanah Hasan Basri

An attack graph (AG) is an abstraction that represents the paths by which an attacker could break a security policy, leveraging interdependencies among discovered vulnerabilities. However, current AG implementations are inefficient on large-scale networks. The increase of the number of hosts in networks causes an increase in the time it takes to generate the AG, especially the calculation time and the complexity of determining reachability.

The complexity of the network has reduced the reachability rate; in other words, it reduces the accuracy of the reachability calculation. In this article, we will provide an overview of the AG phases, and will explain in detail the reachability phase with the factors that affect the reachability calculation and the limitation of the reachability phases.

Networks have grown rapidly in terms of both complexity and size. However, attacks have increased as well, prompting the need for cyber defence analysis. Therefore, any security analysis should consider the cause-effect of the relationships between existing vulnerabilities to secure the network. Attack graphs have been proposed to determine the relationship and interaction between the vulnerabilities of an exploitable network.[1]

'Generating an attack graph is useful in merging vulnerabilities to present the attack routes that lead to a target inside the network'

The attack graph is a security model that shows the chains of existing vulnerabilities and exploits that could exist in various forms in a network.[2] Generating an attack graph is useful in merging vulnerabilities to present the attack routes that lead to a target inside the network. Security professionals should concentrate on vulnerabilities or errors that pose the highest risks via evaluating the path of attack that could be exploited.[3]

Attack graph generation is usually divided into three main phases. The first phase is attack graph reachability, which mostly considers the calculation of accessibility conditions between the network hosts. The second stage is the attack graph modelling phase, which involves how to map the independent attack templates and attack graph structures. And the last stage involves graph core building; after all paths of the attack are determined, there are many paths that might be pruned to generate the attack graph. The attack graph construction has many issues; in this article, we will focus on the reachability analysis issues.

Major challenge

Scalability remains a major challenge in creating the attack graph. In large networks, the whole graph might be expensive to traverse in both storage space and time. When the number of host devices, the complexity of network topologies and number of vulnerabilities increase linearly, the number of graph nodes and edges will increase exponentially. Considering that calculating resources are affected by the processor performance and storage space, it is essential to design algorithms for the construction of attack graphs that are adaptable to variable networks.[4]

Improving the attack graph reachability analysis is one of the ways to solve the scalability.

'Attack graphs represent one method of analysing information about a network and its vulnerabilities. Many researchers have proposed them as a way to recognise critical vulnerabilities in the network, create adversary models, evaluate network protection and recommend improvements'

The reachability analysis stage mostly explores the reachability conditions in the goal network, which defines whether two given devices can reach each other. Increasing the efficiency and the accuracy of the reachability analysis will reduce the time of the reachability calculation and reduce the complexity of the network, which will in turn reduce the complexity of the scalability issue.

In this article, we will explain the analysis phase steps and conditions of reachability needed to increase the reachability rate. We will describe the limitations and challenges of previous research and the open issues to be tackled by future research.

Attack graph

The idea of an attack graph was proposed by Phillips and Swiler.[5] The attack graph

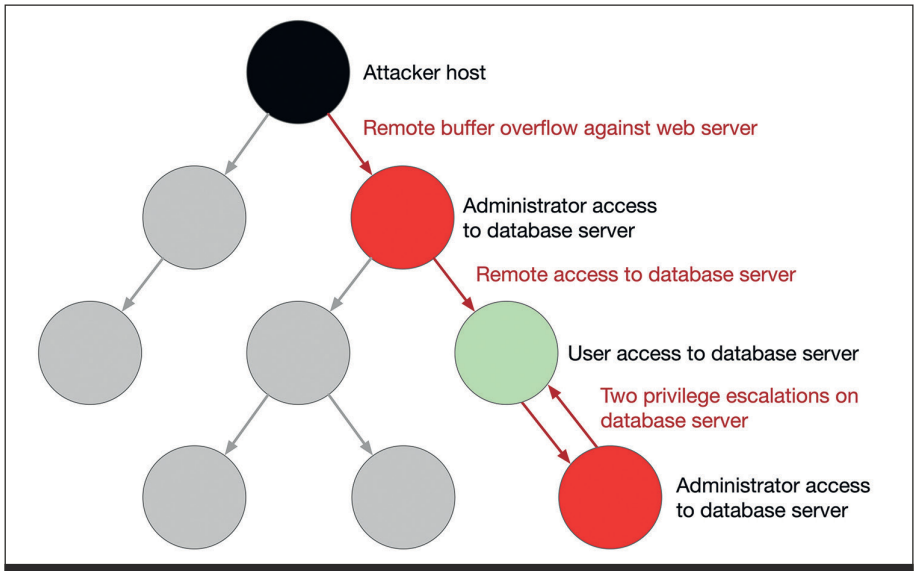


Figure 1: Attack graph example.

is a model used to describe all the potential paths in which the attacker might compromise a security policy by exploiting interdependent vulnerabilities. Beside security policies, the minimum information needed for building the attack graph contains host vulnerabilities and the connectivity of hosts.[6]

The attack graph shows the chains of the existing vulnerability exploits in the network that could be in different forms.[7] Attack graphs represent one method of analysing information about a network and its vulnerabilities. Many researchers have proposed them as a way to recognise critical vulnerabilities in the network, create adversary models, evaluate network protection and recommend improvements to enhance security. Since the first representation of an attack graph, the concept has remained the same. Conceptually, it is possible to think of an attack graph as a directed node-link network. Each node represents a network status that an attacker has managed to achieve. Links relate to behaviour, allowing the attacker to reach a given state. Figure 1 illustrates this particular representation.[8]

The attack graph can be generated (constructed) in three stages, as shown in Figure 2. The first stage is attack graph reachability; in this stage, the reachability for each node must be defined to determine the path between the nodes. Reachability analysis mostly takes into account the conditions of reachability between the network devices. These conditions can be determined by reachability contents such as a firewall, access control and so on.[9]

The second stage is attack graph modelling, which takes into consideration how to map independent attack templates and the structure of the attack graph. Usually, in this phase the elements and logic of one or more attacks are expressed via an attack template that describes the achieved/required capability of an attacker. The attack template may contain high-level, abstract adversaries and threat models or low-level vulnerabilities and exploit models. The threat model could be built via determining the relationships between exploit and vulnerability models.

The last phase is the attack graph core building stage; this refers to the heart of the algorithm to generate the attack graph. In this stage, the attack paths are defined, and many paths could be pruned to generate the final attack graph. An attack graph core-building technique could be considered from two different viewpoints. The first involves the attack path determination methods described above. The second perspective is attack path pruning.

After generating the attack graph, it can be used for many purposes in a positive or negative matter. In this article, we are generally interested in increasing the network security level. The major use for the attack graph is the computation of network security metrics, near-real-time security analysis, counter-measure recommendations, and network design generation.

Reachability analysis

Reachability in graph theory refers to the ability within a graph to get from one vertex to another. Unfortunately, reachability management is usually a mess because the network is typically too large and has high complexity.[10]

The phase of reachability analysis mostly explores the conditions of accessibility in the network, which defines whether two given devices could reach

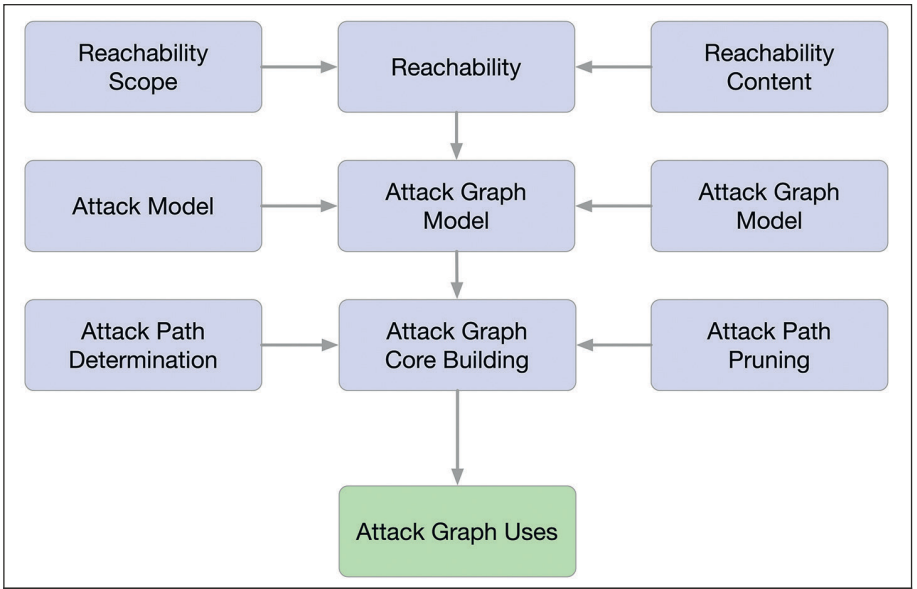


Figure 2: Attack graph generation.

one another. It may also specify more particular information, like whether two host applications might reach each other, which protocol might be used to communicate among the two devices, and so on.

The reachability conditions between the network devices are usually defined using the matrix of reachability, with the columns and rows indicating the network devices. Each cell in the matrix is represented by a Boolean number. The Boolean number represents the accessibility (reachability) between two hosts. The conditions refer to network information such as access control, firewalls, etc.

Calculating reachability is complex, but information about accessibility must be obtained and used by an attack graph system employing real network data. Reachability calculation uses information about the topology of the network, filtering tools, application relationships, trust relationships and intrusion prevention system (IPS) modelling to find all paths among source devices and destination ports. The rule sets of all filtering devices on the network should be imported and modelled.

Reachability is difficult to analyse manually, especially when hundreds or thousands of filtering functions exist in firewalls and Network Address Translation (NAT) rules. Visual representations of reachability greatly simplify the task of a system administrator in trying to understand the security of large network's security.[11]

The reachability analysis phase is divided into two parts – reachability scope and reachability content. In the following sections, we will explain in detail each part.

Reachability scope

Reachability scope includes the methods to calculate the reachability for the network. It can also be seen as clarifying the way to calculate the reachability of the network by calculating parts of it individually or the whole network at the same time. It is divided into two parts – the entire network and atomic domains, as in Figure 3. Determining the reachability of the whole network means calculating the reachability for every part of it at the same time. Most researchers use the reachability of the entire network as an input in attack graph core building. Therefore, the calculation of the reachability using the whole network takes time, especially when the network is large and has many firewalls. Also, if any update takes place in any host, the whole network reachability needs to be recalculated, which wastes time, as the reachability of even unchanged hosts must be calculated again.

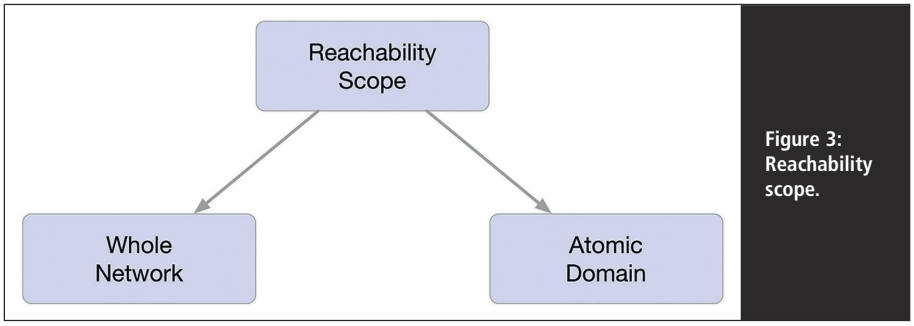


Figure 3: Reachability scope.

The atomic domain approach means that the reachability of each node or group of nodes is calculated individually. Therefore, atomic domain reachability could produce inherent support for the development of an attack graph core-building algorithm. The atomic domain could reduce the calculation time, especially when any update happens, because with this approach you need only update the hosts that have been changed.

The reachability content can be classified using many criteria, such as firewall rules modelling, as shown in Figure 4.

Firewalls

Firewalls are standard mechanisms for protecting networks but, like any other security mechanism, they become useless when incorrectly configured.[12]

Firewalls are usually categorised as host firewalls or network firewalls. Network firewalls filter traffic flowing between more than one network and run on the hardware of the networks. Host firewalls run on the host computers and control network traffic in and out of those devices.

Usually, the reachability of the firewall model uses tuples in the form: *[source IP –> target IP:portnum/protocol]*. These sets are seen as conditional binary decision diagrams (BDDs). The BDD is an effective means of expressing a Boolean equation such as $x \vee (z \vee y)$.

The firewall model uses rules, rule groups and chains. A rule matches a subset of accessibility and acts upon it. These operations include authorising and rejecting traffic. The general requirements will be discussed later.

Each rule is a part of a rule group. A rule group absorbs a set of reachability as input and generates three output sets: *<A, D, R>*. The A, D, R notation, adapted from FIREMAN, refers to the collection of allowed traffic (A), denied

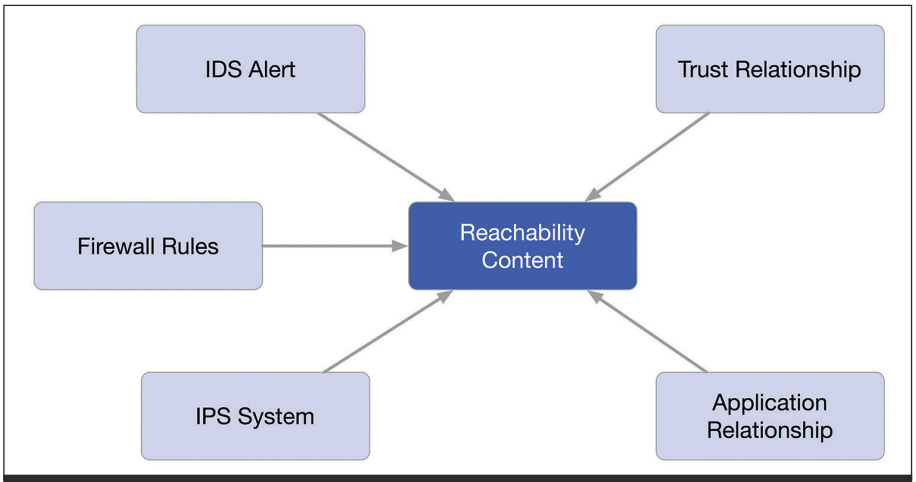


Figure 4: Reachability content criteria.

traffic (D) and traffic that was not acted on via a permit-or-deny rule (R).

A chain determines the next traffic step for the rule group's three output sets. To adjudicate traffic that passes through it, every interface in the network is allocated an inbound and outbound chain.

The firewall has a direct effect on reachability, especially on large networks. Usually, the network has many firewalls to protect it from any unauthorised access. While calculating the attack graph reachability, each firewall must be calculated, as must each rule inside the firewall to determine whether access to each host is allowed. But in large networks, calculating the rules of each firewall has drawbacks. The time taken to calculate the attack graph reachability will be high, and it will have an effect on the attack graph generation time.

IPS systems

An IPS is a common approach for defending networks. An IPS prevents attacks from getting inside the network by comparing traffic behaviour picked up by sensors to records of known bad behaviour using pattern recognition techniques. When attacks are identified, the IPS logs and blocks the offending data.[13] In other words, an IPS uses signatures to detect activities in network traffic and by hosts, performing outbound and inbound packet detection, and is able to block activities before access to the network is achieved and any damage is done. An IPS is an upgrade from an intrusion detection system (IDS) because it not only has the ability to detect intrusions but can also take action against the intrusion and possibly malicious network behaviour.[14]

In the attack graph, in the case of mapping an attack vector to vulnerabilities, the IPS will be treated the same as a firewall. The IPS filter will be expanded to allow blocking of traffic depending on vulnerabilities and privileges, adding vulnerabilities and privileges to the normal tuple of protocol, destination port, destination IP and source IP. But, as we mentioned earlier, calculating IPS reachability has some issues in large networks, especially when the network has many such systems, and this will increase the calculation time. Also, it's possible that some of the more complicated rules inside the IPS filtering model have to be removed.

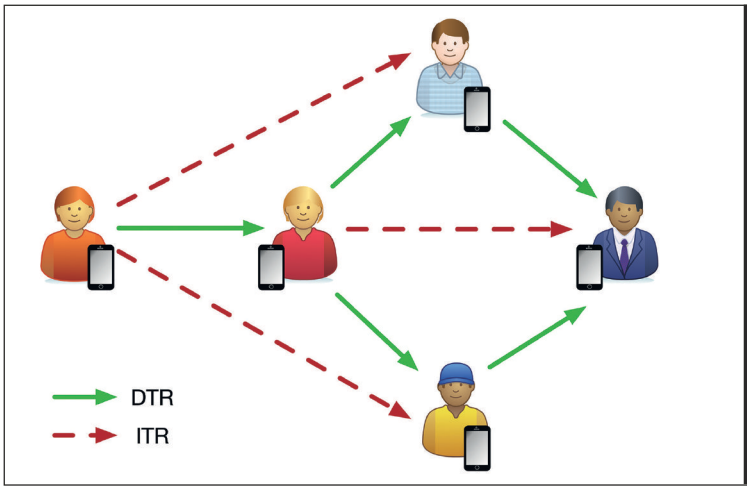


Figure 5: Trust relationship establishment.

Trust relationship

A logical link can be formed between directory domains so that users' and devices' rights and privileges in one domain are shared with the other. As an example, it may allow users to sign in once and have access to all related services without having to be authenticated again.

When building a relationship between entities, trust is important. Different domains use different techniques of modelling and calculation to test trust. Probability and statistics are popular techniques employed in dynamic networks where the topology is rapidly changing to evaluate modelling and trust calculations.[15,16]

Trust is a crucial security concept. It indicates expected behaviour and the belief that a particular entity will produce a desired or specific result, and will function predictably under certain circumstances. In the attack graph, calculating the trust relationship is important in order to know the privileges that can be obtained for each node.

Trust relationships are usually built in two ways – a direct trust relationship (DTR) and an indirect trust relationship (ITR), as shown in Figure 5.[17] *A* DTR is built without the intervention of third parties – for example, the trust *A* has in *B (T(A → B))*. Indirect trust relationships are built with the aid of third parties – for example, the trust relationship A has in C and D through B. This last one is also known as the derived trust relationship. The third party could be a person or a system.

Application relationship

The 'relationship among applications' and the 'relationship among components' elements are defined via operation mechanisms, network topology and service requirements.[18] The relationship could be complicated – for example, an aviation enterprise network dispute programme. It could also be simple, for instance, traditional series-parallel relationships, the connection in computer networks between a client and a server, the fault detection architecture in a network and the connection in a transport network between the main road and a feeder road.

In the calculation of reachability information, account is taken of the usage relationship between network applications. In the generation of the reachability conditions, it is necessary to integrate the application relationships in the target network as comprehensively as possible. The reason is that it's believed that if the attacker has access to the device, he could exploit any software vulnerabilities installed on that server. This modelling is inaccurate but it can be enhanced by matching the results of existing vulnerability scanners with the filter rules used in the network.

In fact, many researchers model the conditions of reachability between network hosts by disregarding or abstracting the conditions of reachability between the software installed on the networks.

Even those that account for the relationships between software applications represent only the accessibility conditions between the devices. This modelling gives rise to the creation of paths of attack that could not be satisfied in the real world by an attacker.

IDS alert

Warning correlation techniques for intrusion detection systems aid in determining if an isolated alarm is part of an ongoing multi-step intrusion into the network. It also helps in reconstructing the scenario of the attack. Most techniques of alert correlation use previous knowledge about strategies of attack or alert dependencies. Alerts are aggregated with similar attributes from other techniques (such as alerts with same destination addresses) or statistical patterns.[19]

As with IPS solutions, an IDS alert can be treated as a firewall; the firewall can deal with traffic coming from outside the network while the IDS can deal with the danger that comes from inside the network – in other words, it can deal with the hosts that already have the authority to access the target.

IDS sensors identify malicious behaviour and create alerts via a reporting component. The sensor is an independent process, which could be network-based (NIDS) or host-based (HIDS) – eg, Snort or Samhain.[20] An IDS can provide three pieces of information – these concern malicious behaviour, information about vulnerabilities and system information. This information will give us a deeper understanding about the network and can also be used in determining reachability to generate the attack graph.

Noel, Robertson and Jajodia were the first to use the attack graph in minimising the effect of false alarms by correlating isolated intrusion alerts as part of multi-step attack paths.[21] The method of warning correlation is focused on the smallest path in the attack graph between exploits. Additionally, any IDS alert that doesn't appear in the attacker's potential future activity (as seen in the attack graph) could easily be counted as false.

The IDS alert has some issues, the major one being fake IP addresses. The information from the IP packets is read via the IDS, but the network address can still be spoofed. If an attacker is using a fake address, it makes the threat more difficult to detect and assess.

Reachability calculation

Reachability is the first and main phase in building an attack graph. The reachability determines the ability for each node in the network to access another host. All reachability contents that have been discussed in this article are used to calculate the reachability conditions.

The attack graph generation assumes that each host in the network has an IP address and is connected to one or more interfaces. These interfaces have zero or more ports that are open and accept connections from other hosts. Each port has a port and protocol number. Each host and port could have zero or more instances of vulnerabilities, unique defects or configuration choices that an attacker can exploit.

There are many ways to calculate reachability depending on conditions. The reachability conditions can be calculated using a matrix. The rows and columns in the matrix represent the main reachability content and the other cells represent the conditions between the hosts. [22] We use a straightforward method to calculate the reachability, where the row represent the source interfaces on the host and the column represents the target ports on the destination interfaces. Ingols, Chu, Lippmann, Webster and Boyer used an *I–K* matrix in which *I* represented the number of the interfaces and *K* represented the number of server ports.

Zhao, Wang, Zhang and Zheng calculated the reachability using a link matrix (LM).[23] This is used to describe the interconnections between network hosts. The rows and the columns represent the hosts inside the network while the cells inside the matrix are represented with Boolean numbers where 1 means these hosts are reachable and 0 means they are not. The authors also used an attack rule that describes the attack preconditions, difficulties and effects. But this work calculates the reachability manually and has a problem with large networks because it will take considerable time, especially if the network has many firewalls and IPS/IDS systems.

Another way to calculate the reachability is using a hypergraph. Karypis and Kumar proposed a hypergraph partitioning approach based on a greedy k-way refining algorithm. The main purpose of a reachability partitioning hypergraph is to achieve load balancing in aspects of hyper-vertex weights as well as decreasing the number of hyper-edges throughout the search agents.

Kaynar and Sivrikaya calculated the reachability using a hyper-graph as an initial task for every distributed search agent to generate the full attack graph. The authors represented the reachability conditions between the software applications. The filtering rules on the firewalls, routers' access control lists and security policies applied between software applications are among the factors that define the conditions of reachability. All of these factors are taken into account in reachability hypergraph generation, in which a hyper-vertex indicates a software application. A hyper-edge implies a collection of target and source software applications so that, if specific conditions are met, the source applications can directly access the target applications. These conditions are stored in the hyper-edge, thus allowing direct reachability between the software applications.

There is an advantage to using a hyper-graph to calculate the reachability. In terms of storage space, a hypergraph is more efficient than a reachability matrix or a graph itself. But the calculation of the hyper-graph is complex.

Challenges and open issues

With so many vulnerabilities these days, the need for security is a major issue for security administrators. Direct exploitation of vulnerabilities can provide access to different resources in the target device. However, as the complexity of networks has increased, attackers might follow indirect or direct reachability paths to control the target device. Indirect reachability is hard to detect

because it needs vulnerability information and complex analysis. Hence the attack graph comes into play.

An attack graph is a combination of edges and nodes where a node represents exploited vulnerabilities on a device and edges represent reachability from one device to another. Reachability analysis mainly examines the conditions of reachability within the target network which, from a simplified perspective, decides whether two hosts are allowed to access one another.

Actually, by abstracting or disregarding the reachability conditions between the installed software on the hosts, most of the previous works mentioned in this article model reachability conditions between the target network hosts. In this section, we will highlight some of the challenges for future work in attack graph reachability analysis.

- Using too few reachability elements may lead to missed information about the next step of the attack, as discussed by Ingols, Lippmann and Piwowarski, in which calculations for reachability were made only for inline firewalls usually used in network borders. The previous system's performance rapidly deteriorated with many personal firewalls because every personal firewall was modelled individually as an inline firewall.
- As mentioned in Kaynar, the main issue in the reachability conditions calculation is how to consolidate as closely as possible the results of the reachability graph or matrix. Another challenge has been mentioned: how to improve the application relationship when calculating reachability because accounting for the relationships between software applications reflects just the conditions of reachability between the hosts. Obtaining and using the IPS locations and signatures to calculate the reachability conditions also gives more accurate results compared to just using the filtering and access control rules and trust relationships. This approach could eliminate the paths of attack that employ exploits blocked by the IPS signatures. But IPS solutions have proven to be efficient protection in the past; new complex attacks require more correlation between alerts in order to understand the purpose of the attacks before intervening and stopping them.
- There's a challenge in calculating the reachability automatically. Many researchers calculate the reachability manually. The manual calculation needs time, especially on large networks.
- Even where researchers have generated the attack graph and calculated the reachability automatically, the conditions for generating the attack path map mean it is difficult to achieve automatic extraction, and the generation algorithm is too complicated to be applied to large-scale complex networks. [24]

Conclusions

An attack graph models possible paths that a potential attacker can use to intrude into a target network. The generation of an attack graph suffers from the complexity of state space. We noticed that many previous works try to solve this problem by enhancing the reachability calculation. Calculating the reachability can be done using a matrix or a hyper-graph. These two ways each present reachability conditions in the network, like firewall rules etc.

However, previous research has many limitations. Some of the researchers, for example, didn't include all the conditions while calculating reachability, while others do the calculation manually. The conditions for generating the attack path map are difficult to achieve automatically and generating the attack graph automatically is too complex. These limitations can affect the time of generating the attack graph in large networks.

About the authors

Zaid Jasim Al-Araji is a PhDt candidate at the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Malaysia. He received his BSc in computer science from the University of Mosul, Iraq in 2011. He obtained his MSc in computer science from Universiti Teknikal Malaysia Melaka in 2017. His current research interests include security.

Dr Sharifah Sakinah is currently an associate professor in the Department of Intelligent Computing & Analytics (ICA), Faculty of Information & Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM). She received her bachelors and masters degrees in applied mathematics in the School of Mathematics at the University of Science, Malaysia. Following this, she received her PhD from the University Of Alberta, Canada in 2012 in intelligent systems.

Dr Raihana Syahirah Abdullah is a senior lecturer in the Department of Computer Systems and Communication at the Faculty of Information Technology and Communication, Universiti Teknikal Malaysia Melaka (UTeM). She is a member of the information security, digital forensic and computer networking (INSFORNET) research group. She completed her bachelor degree in computer networking, her masters in computer science and her PhD in network security at (UTeM). Her research interests include computer and network security, botnets and malware analysis.

Ammar Awad Mutlag is a PhDt candidate at the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Malaysia. He received his BSc in computer science from the College of Information Technology, Imam Sadiq University, Iraq in 2009. He obtained his MSc in information technology from Andhra University, India in 2015. His current research interests include artificial intelligence, and biomedical computing.

Hayder Adil Abdul Raheem is a PhDt candidate at the Faculty of Technology Management and Technopreneurship, Universiti Teknikal Malaysia Melaka. He received his BSc in information technology management from the Technical College of Management, Iraq in 2009. He obtained his MSc from the Universiti Teknikal Malaysia Melaka. His current research interests are in information technology, industry 4.0, innovation and high-tech.

Siti Rohanah Hasan Basri, SPd MPd, completed her masters in educational technology from the University of Pelita Harapan, Indonesia and gained a TEFL/TESOL certificate from Asian College Teachers, Bangkok, Thailand.

References

1. Karypis, G; Kumar, V. 'Multilevel k-way Hypergraph Partitioning: Initial Partitioning Phase'. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 1999. Accessed Jun 2021. https://www.hindawi.com/journals/vlsi/2000/019436/.
2. Kaynar, K. 'A taxonomy for attack graph generation and usage in network security'. Journal of Information Security and Applications, vol.29, pp.27-56, 2016.
3. Aksu, MU; Dilek, MH; Tatlı, E; Bicakci, K; Ozbayoglu, M. 'Automated generation of attack graphs using NVD'. In proceedings of the Eighth ACM Conference on Data and Application Security and Privacy, pp.135-142, 2018.
4. Li, M; Hawrylak, P; Hale, J. 'Concurrency strategies for attack graph generation'. 2nd International Conference on Data Intelligence and Security (ICDIS), pp.174-179, 2019.
5. Phillips, C; Swiler, LP. 'A graph-based system for network-vulnerability analysis'. In Proceedings of the 1998 Workshop on New Security Paradigms, pp.71-79, 1998.
6. Ramos, A; Lazar, M; Filho, RH; Rodrigues, JJPC. 'Model-based quantitative network security metrics: a survey'. IEEE Communications Surveys & Tutorials, vol.19, no.4, pp.2704-2734, 2017.
7. Hamid, T. 'Attack graph approach to dynamic network vulnerability analysis and countermeasures'. University of Bedfordshire, 2014.
8. Williams, L; Lippmann, R; Ingols, K. 'GARNET: A graphical attack graph and reachability network evaluation tool'. In International Workshop on Visualization for Computer Security, vol.5210 LNCS, pp.44-59, 2008.
9. Kaynar, K; Sivrikaya, F. 'Distributed attack graph generation'. IEEE Transactions on Dependable and Secure Computing, vol.13, no.5, pp.519-532, 2016.
10. Khakpour, AR. 'Network reachability: quantification, verification, troubleshooting, and optimization'. Michigan State University, Computer Science and Engineering, 2012.
11. Ingols, K; Chu, M; Lippmann, R; Webster, S; Boyer, S. 'Modeling modern network attacks and countermeasures using attack graphs'. In 2009 Annual Computer Security Applications Conference, pp.117-126, 2009.
12. Bodei, C; Degano, P; Galletta, L; Focardi, R; Tempesta, M; Veronese, L. 'Language-independent synthesis of firewall policies'. In 2018 IEEE European Symposium on Security and Privacy (EuroS&P), pp.92-106, 2018.
13. Rengaraju, P; Ramanan, VR; Lung, CH. 'Detection and prevention of DoS attacks in software-defined cloud networks'. In 2017 IEEE Conference on Dependable and Secure Computing, pp.217-223, 2017.
14. Jamar, R; Sogani, A; Mudgal, S; Bhadra, Y; Churi, P. 'E-shield: Detection and prevention of website attacks'. 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), vol.2018, Jan, pp.706-710, 2017.
15. Blaze, M; Feigenbaum, J; Lacy, J. 'Decentralized trust management'. Proceedings 1996 IEEE Symposium on Security and Privacy, June, pp.164-173, 1996.
16. Lamba, A. 'A through analysis on protecting cyber threats and attacks on CPS embedded subsystems'. SSRN Electron. J, vol.1, no.3, pp.48-55, 2014.
17. Almenárez, F; Marín, A; Díaz, D; Cortés, A; Campo, C; García-Rubio, C. 'Trust management for multimedia P2P applications in autonomic networking'. Ad Hoc Networks, vol.9, no.4, pp.687-697, 2011.
18. Li, R; Kang, R; Huang, N; Chen, W; Chen, Y. 'A practical approach for network application reliability assessment'. Eksploatacja i, vol.44, no.4, pp.17-27, 2009.
19. Barik, MS; Sengupta, A; Mazumdar, C. 'Attack graph generation and analysis techniques'. Defence Science Journal, vol.66, no.6, pp.559-567, 2016.
20. Roschke, S; Cheng, F; Meinel, C. 'Using vulnerability information and attack graphs for Intrusion Detection'. In Sixth International Conference on Information Assurance and Security, no.8, pp.104-109, 2010.
21. Noel, S; Robertson, E; Jajodia, S. 'Correlating intrusion events and building attack scenarios through attack graph distances'. In 20th Annual Computer Security Applications Conference, pp.350-359, 2004.
22. Ingols, K; Lippmann, R; Piwowarski, K. 'Practical attack graph generation for network defense'. In 22nd Annual Computer Security Applications Conference (ACSAC'06), pp.121-130, 2006.
23. Zhao, Y; Wang, Z; Zhang, X; Zheng, J. 'An improved algorithm for generation of attack graph based on virtual performance node'. In 2009 International Conference on Multimedia Information Networking and Security, vol.2, no.2, pp.466-469, 2009.
24. Ma, JC; Wang, YJ; Sun, JY; Chen, S. 'A minimum cost of network hardening model based on attack graphs'. Procedia Engineering, vol.15, pp.3227-3233, 2011.

The Firewall

Quantifying cyber risk

Colin Tankard, Digital Pathways

There is no question that all types of organisations are exposed to cyber risk. It is an inevitable part of doing business in the digital world.

In a recent Enterprise Strategy Group (ESG) survey, the market researcher found that 82% of organisations believe that cyber risk has increased over the past two years and that 69% of business and technology leaders believe cyber security is primarily a technology area, with little or no connection to the business. (The report is available here: https://bit.ly/3wNEh1U).

For many, cyber risk is seen as complex – an issue to be discussed in technical terms or after a security breach, when decisions are taken in 'knee jerk' fashion.

These circumstances highlight a fundamental challenge for today's security leaders in bringing data security to the same level as other business initiatives. To achieve this, cyber risk needs to be presented in a form that can be measured in financial terms, ultimately helping non-technical stakeholders, such as the board, to understand how cyber risk translates into business risk.

"CISOs must work within the technical and business realms to make informed, data-driven decisions"

It is clear that today's chief information security officers (CISOs) must work within the technical and business realms to make informed, data-driven decisions in order to secure the necessary budget and protect a company's interests. To do this effectively, a cyber risk quantification framework is needed that allows them to report to non-technical stakeholders in a language they understand, aligned with other initiatives that receive funding.

By quantifying cyber risk financially, CISOs can analyse cyber risk in the same way the organisation looks at all other types of risk that impact on financial targets. This process puts the fluid nature of cyber risk into a comprehensible business context. It will help stakeholders understand a company's potential financial exposure, due to various risk factors and impact scenarios.

Having these data-driven insights, decision-makers can allocate resources and prioritise remediation plans, based on how much a company stands to lose financially if it does not address a particular gap in its data security strategy.

From here, risk modelling can be expanded to incorporate other data points, such as third-party risk management. Monitoring of suppliers to mitigate cyber risk by measuring their security ratings against recognised cyber security metrics enables a company to form an acceptable risk register of suppliers and partners. This may lead to a vendor needing to improve its cyber protection or assess its position with its own vendors to continue to do business with the organisation.

A final part of understanding the cyber risk can be linked to social media or dark web chatter. Often this can indicate a data breach or leaked data that may be in the process of being used to attack an organisation. (It may also indicate a successful breach that the company doesn't know about yet.) Having this insight enables a company to quantify the risk and to immediately take measures to block the attack, if the risk is deemed high.

Such a holistic view of cyber risks will lead to an overall strengthening of a company's cyber security posture. Ensuring that data security purchases have a return on investment will elevate all data security controls to a point that they are seen as a business benefit rather than a cost.

EVENTS CALENDAR

Due to the Covid-19 pandemic, many conferences are being cancelled, postponed or converted into virtual events. The events listed here were still planned to proceed at the time of publication.

13–15 July 2021
Infosecurity Europe
Virtual conference
www.infosecurityeurope.com

31 July – 5 August 2021
Black Hat USA
Las Vegas, US
www.blackhat.com

5–8 August 2021
DEF CON 29
Las Vegas, US
https://defcon.org

11–14 August 2021
International Conference on Information Security and Cryptology
Qindao, China
https://cst.qd.sdu.edu.cn/inscrypt_2021/

23–27 August 2021
HITB Singapore
Virtual conference.
https://conference.hitb.org/hitbsecconf2021sin/

24 August 2021
Chicago Virtual Cyber Security Summit
Virtual conference.
https://cybersecuritysummit.com/summit/chicago21/

2 September 2021
Enterprise Security & Risk Management
Virtual conference.
https://whitehallmedia.co.uk/esrmamericassep2021/

6–8 September 2021
Copenhagen Cybercrime Conference
Virtual conference.
https://bit.ly/3nJ6wM7