Unicode flaw leaves most software open to malicious code injection

Researchers at Cambridge University have discovered a trick that could allow someone to insert malicious functionality into source code in a way that’s unlikely to be picked up by either automated or manual code reviews. This could open the way to serious software supply chain attacks, especially with open source repositories.

A paper by Nicholas Boucher and Ross Anderson details a more serious evolution of the infamous right-to-left override (RLO) attack, which has previously been used to obscure the names of executables in emails. The problem – dubbed Trojan Source by the authors – lies in Unicode’s ability to display text in a left-to-right or right-to-left order, depending on the language used. Although normally set as a default for a given document, program or script, this function can be overridden using Unicode’s bidirectional (BiDi) algorithm to allow mixed-script use – for example, a document containing both Latin alphabets and Hebrew or Arabic.

A malicious actor could exploit this feature to insert malware into comments or strings in source code in such a way that it would not be obvious to anyone reading the code, but it would be treated by a compiler as valid instructions and would therefore be compiled into the executable.

“While both comments and strings will have syntax-specific semantics indicating their start and end, these bounds are not respected by BiDi overrides,” say the researchers in the paper. “Therefore, by placing BiDi override characters exclusively within comments and strings, we can smuggle them into source code in a manner that most compilers will accept.”

Although they say they have not seen any examples of this technique being used in the wild, the researchers believe the threat is so great that they carefully coordinated disclosure of the problem with 19 organisations, many of which are releasing fixes in compilers, interpreters, code editors and repositories.

Programming languages affected include C, C++, C#, JavaScript, Java, Rust, Go and Python. The researchers found that the problem affects Windows, macOS and Linux, and they determined that the problem affects many of the most popular code editors, including VS Code, Atom, Sublime Text, Notepad, Xcode, vim and emacs, as well as web-based services such as GitHub and Bitbucket. Malicious code embedded in comments and strings easily survives being cut and pasted, which means that users of code-sharing and support sites like Stack Overflow are not immune.

“Some attacks provided strange highlighting in a subset of editors, which may suffice to alert developers that an encoding issue is present,” says the paper. “However, all syntax highlighting nuances were editor-specific, and other attacks did not show abnormal highlighting in the same settings.”

The BiDi issue is being tracked as CVE-2021-42574. There is another related issue that exploits homoglyphs – visually similar characters – that is being tracked as CVE-2021-42694.

“We’ve seen a variety of novel attacks on software supply chains in 2021 and this is another example of how the trust placed in development processes can be

Continued on page 2...


The annual Crime Survey for England and Wales (CSEW), carried out by the Office for National Statistics (ONS) has shown a rapid rise in what the organisation classifies as computer misuse crimes. Yet prosecutions under the Computer Misuse Act are falling.

The CSEW is a telephone-based survey which, the ONS claims, is more representative of crime figures than those gathered by police forces because it includes unreported incidents. Extrapolating from its figures, the survey estimates that there were 1.8 million computer misuse crimes in the year ending June 2021. That’s similar to the figures recorded in 2017. However, the earlier survey included teenagers, whereas the most recent figures are limited to adults. The new data also shows a significant increase over the figures in more recent years, including an 85% jump from 2019.

The real number of crimes is difficult to estimate. Many people don’t know that they have been victims and others are reluctant to admit it – either to the police or people carrying out surveys. The sharp rise this year is believed to be connected to the high level of data breaches – not least those that result from ransomware attacks. The ONS reports a 161% increase in “Unauthorised access to personal information (including hacking)” offences. It goes on to say: “This included victims’ details being compromised via large-scale data breaches, and victims’ email or social media accounts being compromised.”

The National Fraud Intelligence Bureau (NFIB) has also recorded a 31% increase in the number of ‘Hacking – personal’ offences being report to Action Fraud, which is tasked with handling most tech-related crime.

Recent figures show a 20% drop in prosecutions under the Computer Misuse Act. There have also been ongoing complaints about the inaction of Action Fraud.

One slightly odd figure that the ONS mentions is that its ‘Nature of fraud and computer misuse in England and Wales: year ending March 2019’ report showed that 68% of people involved in data breaches claim not to have been affected at all by the incident. It’s debatable whether most people are actually in a position to...
A new malware loader is being used as part of a highly active spam campaign. Dubbed SquirrelWaffle, it exploits Qakbot malware and the Cobalt Strike security tool hidden inside malicious Microsoft Office documents. According to researchers at Cisco Talos, the campaign started in mid-September. SquirrelWaffle emails typically contain hyperlinks to malicious Zip archives hosted on attacker-controlled web servers. Although the campaign has many similarities to Emotet, it hasn’t yet reached that malware’s scale of operations. However, the researchers warn that this situation could change and that SquirrelWaffle could be the basis for an extremely active and large-scale spam operation. There’s more information here: https://bit.ly/3kkbOqg.

BrakTooth PoC
The group of vulnerabilities recently discovered in Bluetooth implementations, dubbed BrakTooth, are now exploitable thanks to the publication of proof-of-concept (PoC) code. The US Cybersecurity & Infrastructure Security Agency (CISA) warns that: “An attacker could exploit BrakTooth vulnerabilities to cause a range of effects from denial of service to arbitrary code execution.” The PoC code, published on GitHub, is based around a specific Bluetooth development kit, but it’s likely that it will form the basis for more exploits affecting a broader range of hardware platforms. CISA is urging manufacturers, vendors and developers to review the code and apply updates or workarounds to their solutions. There’s more information here: https://bit.ly/3BVNyax.

Password manager exploited
Threat actors are targeting a known flaw in the Zoho ManageEngine ADSelfService Plus password manager to attack organisations in a number of sectors, including technology, defense, healthcare, energy and education. The vulnerability (CVE-2021-40539) is a critical authentication bypass that allows unauthenticated remote code execution (RCE). The flaw was patched in September, but it has been under active attack since August and many organisations are still vulnerable, according to researchers at Palo Alto Network’s Unit 42. A breach allows attackers free access to a victim’s Active Directory and cloud accounts. There’s more information here: https://bit.ly/3EXki52.

Linux kernel bug
A flaw in the Linux kernel (CVE-2021-45267) could result in clustered machines being compromised. The bug is in the Transparent Inter Process Communication (TIPC) module which is used for communication within Linux clusters. According to researchers at SentinelLabs, TIPC can be used as a socket and can be configured on an interface as an unprivileged user. A flaw in message size validation can result in a heap overflow that an attacker can exploit to mount an attack. The flaw exists in kernel versions 5.10 to 5.15, although the TIPC module is not enabled by default. There’s more information here: https://bit.ly/300bkH.

Cisco flaws
Cisco has warned of two vulnerabilities in its Catalyst PON series of optical switches that have the maximum severity rating of 10. CVE-2021-34795 relates to something the firm calls an “unintentional debugging credential”, although it hasn’t yet provided full details. It’s likely that it’s a backdoor created for engineers during development and testing that was left in place when the products rolled out. The hidden credential provides root-level access. CVE-2021-40113 can be exploited by an unauthenticated remote attacker to perform a command injection attack on the equipment’s web-based management portal, thanks to insufficient validation of user-supplied input. There’s more information here: https://bit.ly/2Yrnzdr.

US bans spyware
The US Government has issued sanctions against four companies – in Israel, Russia and Singapore – that are accused of selling spyware and hacking tools to governments and nation-state threat groups.

The Bureau of Industry and Security (BIS), part of the Department of Commerce, has added the companies to the Entity List, which prohibits the export, re-export or transfer of the firms’ products or services.

Two Israeli companies are among the newly sanctioned organisations.

“Investigative information has shown that the Israeli companies NSO Group and Candiru developed and supplied spyware to foreign governments that used this tool to maliciously target government officials, journalists, businesspeople, activists, academics, and embassy workers,” said the BIS.

NSO Group has been embroiled for some time in controversy over the use of its Pegasus spyware, which has been deployed against, among others, European heads of state.

In addition, Positive Technologies in Russia and Computer Security Initiative Consultancy in Singapore have been added to the list for what the BIS called, “their engagement in activities counter to US national security”. It added: “These entities traffic in cyber exploits used to gain access to information systems, threatening the privacy and security of individuals and organisations worldwide.”

Positive Technologies had already been sanctioned, in April 2021, for having allegedly provided assistance to Russia’s FSB intelligence agency in cyber attacks against US targets.

In theory, US companies wanting to trade in the sanctioned firms’ products can still do so, but will have to apply for a licence. This licence has a ‘presumption of denial’, meaning that it’s almost certain no-one will get one. The US Commerce Department explained: “BIS considers that transactions of any nature with listed entities carry a ‘red flag’ and recommends that US companies proceed with caution with respect to such transactions.”

The move has been welcomed by many, including Amnesty International, which with French advocacy group Forbidden Stories recently published a report detailing what it claimed were examples of NSO’s Pegasus spyware being used to violate human rights and to target government officials and members of civil society.

“With this move, the US Government has acknowledged what Amnesty and other activists have been saying for years: NSO Group’s spyware is a tool of repression, which has been used around the world to violate human rights,” said Danna Ingleton, deputy director of Amnesty Tech, in a statement.
Report Analysis

Lookout: 2021 Energy Industry Threat Report

The energy industry is becoming a focus of attention for everyone concerned with cyber security – including the businesses themselves, governments, infosec specialists and – naturally – threat actors. And that last group is all-encompassing too, embracing both cyber criminals out to make a fast buck and nation-state entities looking to wage war by other means.

For attackers, the energy business has many attractive properties. It forms a major segment of a nation’s critical national infrastructure (CNI), not least because bad things can happen when energy services go down. In the electricity generation and distribution sector, for instance, a disruption can put water supplies, transportation systems, hospitals and many other dependent services in critical, life-threatening situations. That’s why electricity firms rank continuity of service as their top priority. Information security gets pushed down the list – ironically, as poor cyber security can so easily lead to the failure of operational technology (OT) and business disruption.

“The energy industry is closely linked to the safety and well-being of society,” says the report. “Being responsible for an important part of our global infrastructure, these organisations sit at the centre of everything from food supplies, education to healthcare and economic growth. For this very reason, the industry is also at the epicentre of cyber attacks: 17.2% of all mobile cyber attacks globally target energy organisations, making the industry the biggest target for hacktivists, cyber criminals and nation-state sponsored attackers.”

It’s easy to see why energy firms would be a target for nation-state hackers. These days, it’s probably easier to take down a powerplant with malware than it is with a bomb – and far more deniable. And most energy firms report being constantly scanned by entities that have good reason to believe are acting on behalf of governments. However, a government needs a good reason to actually attack, and is unlikely to do so outside of an ongoing political conflict. Such attacks have happened – we’ve seen energy firms hit hard with wiper malware, for instance. And cyber espionage is not infrequent. However, the majority of nation-state activity so far has tended to be restricted to reconnaissance and probing for weaknesses to use another time.

With cyber criminals, it’s a different matter. The criticality of many energy services, linked with the deep pockets many of the enterprises enjoy, makes these companies choice targets for ransomware operators and those engaged in extortion attacks based on distributed denial of service (DDoS).

Attackers will use any opening to worm their way into corporate networks. Increasingly, this means taking advantage of the many weak spots in mobile platforms.

Mobile devices offer two vectors that malicious actors are using to their advantage. One is phishing, because it’s harder to employ good defensive behaviours – such as hovering a cursor over a link – on mobile devices than on desktop platforms. And people are often more trusting about messages received on their phones. The other vector is malware, and this is an area where Android has its own special problems.

With the pandemic and the rise of mobile working, the use of personal mobile devices has also burgeoned. The number of unmanaged mobile devices in the industry has risen by 41% in the past year, the report says. There was also a 44% increase in the number of mobile devices connecting to the networks of energy firms, so we can probably deduce that a high percentage of those are personal phones and tablets that aren’t equipped with corporate defences.

With these devices inside the corporate network – behind the firewall and intrusion detection systems – they represent a very significant threat. It’s no surprise then that in the energy business, one in seven (13.6%) employees was subject to at least one phishing attempt in the past year. That’s a higher rate, says Lookout, than all other industries combined. This rate increased to 20% in the first half of 2021, indicating a worsening of the situation.

Then there’s the malware. “Malware delivery – or tricking employees into installing malicious apps on their device – is lucrative cybercrime,” says Lookout. “While ransomware is only one example of the many types of malware, companies reported 2,474 incidents to the FBI, costing $29.1m in losses during 2020.”

The Android ecosystem is so fractured that around 56% of users are running versions that are not receiving security patches. According to Lookout’s estimates, this makes them vulnerable to nearly 300 exploitable flaws. When you add in trojanised apps, that’s a large threat surface. And it doesn’t end there.

“Nearly 95% of mobile app threats facing the industry are either riskware or a vulnerability,” explains the report. “Riskware are legitimate programs that pose potential risks due to a software incompatibility, security vulnerability or compliance violation. Riskware is not the same as a vulnerability, which is a defect in software code that can be exploited by an attacker.”

So what are the bad guys after? In two-thirds of attacks, the malicious actors want to steal credentials. Of course, that’s not the ultimate goal. The access provided by the credentials could allow them to engage in cyber espionage, or drop malware for the purposes of destruction or ransom.

The report is available here: www.lookout.com/threats/energy-industry-threat-report.
Rise in DDoS attacks
Although most attention is on ransomware these days, distributed denial of service (DDoS) attacks have continued to increase in both volume and sophistication, according to Kaspersky. In the third quarter of 2021, the number of DDoS attacks grew by 24% while the number of so-called ‘smart attacks’ grew by 31%, compared to the same period in 2020. Kaspersky defines smart DDoS attacks as those that are more sophisticated and often targeted, and they can be used not just to disrupt services but also to make certain resources inaccessible or to steal money. The firm’s report says that: “Some of the most-notable targets were tools to fight the pandemic, government organisations, game developers and well-known cyber security publications.” Some of the most-notable, large-scale DDoS attacks over the past quarter involved a new, powerful botnet called Meris, which is capable of sending out a massive number of requests per second. This botnet was seen in attacks against two of the best-known cyber security publications – the Krebs on Security website and InfoSecurity magazine. Other notable DDoS trends in Q3 included a series of politically motivated attacks in Europe and Asia, as well as attacks against game developers. “Over the past couple of years, we’ve seen the crypto-mining and DDoS attack groups competing for resources, since many of the same botnets used for DDoS attacks can be used for crypto-mining,” said Alexander Gutnikov, a security expert at Kaspersky. “While we were previously seeing a decline in DDoS attacks as crypto-currency gained in value, we’re now witnessing a redistribution of resources. DDoS resources are in demand and attacks are profitable. We expect to see the number of DDoS attacks continue to increase.” There’s more information here: https://bit.ly/31AQm0d.

DarkSide bounty
While the DarkSide ransomware group claims to have closed down, its members are still being actively sought by law enforcement in the US. The US Department of State has now placed a bounty on the heads of the cyber criminals, offering a reward of up to $10m for information leading to the location, arrest and/or conviction of owners, operators and affiliates of the DarkSide group. DarkSide disappeared after drawing too much attention when one of its affiliates attacked Colonial Pipeline and disrupted fuel supplies in the US. It has since re-emerged as BlackMatter – which itself now claims to have shut down due to “pressure from the authorities” and the disappearance of some members. You can submit tips at: https://tips.fbi.gov, or via WhatsApp, Telegram or Signal.

Ukraine identifies hackers
The Security Service of Ukraine (SSU) has released details of threat actors it alleges are part of the Armageddon (aka Gamaredon) group that has carried out more than 5,000 attacks against 1,500 or so public authorities and critical infrastructure in the country. According to the SSU’s Cyber Security Department: “They are officers of the ‘Crimean’ FSB and traitors who defected to the enemy during the occupation of the peninsula in 2014.” The five men are named as: Sklianko Oleksandr Mykolaiovich, Chernykh Mykola Serhiiovych, Starchenkov Anton Oleksandrovych, Miroshnichenko Oleksandr Valeriovych and Sushchenko Oleh Oleksandrovych. The SSU goes on to claim that the Armageddon group is a special unit within Russia’s security service, the FSB, created to target Ukraine. Its activities are coordinated by the FSB’s 18th Centre (Information Security Centre) based in Moscow. Five members of the group have been notified that they are under suspicion of treason, and the SSU says: “Investigations and forensic examinations are underway to bring the FSB employees to justice for the following crimes: espionage; unauthorised interference in the work of computers, automated systems, etc; creation of malicious software or hardware for use, distribution or sale.”

Top hardware laws
Mitre and the US Cyber security & Infrastructure Security Agency (CISA) have produced a list of the 12 commonest hardware flaws, much like OWASP’s Top Ten for web applications. The list is based on the Common Weakness Enumeration (CWE) database maintained by Mitre, and its publication is intended to provide guidance for designers and programmers involved in product development. At the head of the list is improper isolation of shared resources on system on a chip (SoC) devices, followed by on-chip debug and test interfaces with improper access controls. However, the Hardware CWE Special Interest Group (SIG) says that it should not really be viewed as a hierarchical list. All of the flaws are important – as are a further five that didn’t quite make it on to the list. Developers need to address all of the issues if their products are to be safe. There’s more information here: https://bit.ly/3qwvrbC.

China alleges data theft
Although often cast as the perpetrator of cyber-crimes, China’s Government has released details of major attacks on the country’s infrastructure that it claims were carried out by foreign entities. Tied to coincide with the seventh anniversary of the country’s anti-espionage law, China’s Ministry of State Security said that the three attacks involved airline data stolen by an overseas intelligence agency, shipping data gathered by a consulting firm that handed it on to a foreign spy agency, and the deployment of ‘weather devices’ to transfer sensitive meteorological data abroad. The ministry did not name which foreign countries were involved, nor did it clarify whether the attacks were connected.

NPM backdoored
Two highly popular packages in the NPM JavaScript repository, with combined weekly downloads of 22 million, have been found to contain backdoors. The coa library is a parser for command-line options, and rc is a configuration loader. Both contained malicious code that allows an attacker to gain access to a developer’s accounts via downloaded malware. In both cases, multiple versions of the code are affected, including the most recent. The malicious versions have been removed from the registry. However, many coders may still have copies on their local systems. Users of coa have been advised to downgrade to version 2.0.2 as soon as possible, and users of rc should downgrade to version 1.2.8. NPM has recently come in for a number of criticisms relating to its stewardship of the repository, which is open and has little in the way of code vetting or security measures.

Ransomware profits
Ransomware earned its operators and affiliates around $390m in the first half of 2021, according to the US Government’s Financial Crimes Enforcement Network (FinCEN). And total ransomware-related financial activity – that is, funds being moved around on crypto-currency blockchains – may have reached $5.2bn. The figures come from a Financial Trend Analysis report that examines Suspicious Activity Reports (SARs). There were 458 such transactions in the first half of 2021, plus another 177 older reports that have since been determined to be suspicious. Much of the financial activity was the result of attempts to launder funds. The report goes into details about 68 ransomware variants, with REvil, Conti, DarkSide, Avaddon and Phobos being the most common. The median ransom demand was $148,000. There’s more information here: https://bit.ly/3H51DGj.

Cyber skills shortage
The UK’s cyber skills shortage is getting worse, according to a report by recruitment firm Harvey Nash. Information security is now the most sought-after tech skill, and nearly half (43%) of the organisations surveyed by Harvey Nash said they had vacancies in this sector. The shortfall between skilled professionals and jobs available has worsened by a third in the past year, the report says. There’s more information here: www.harveynashgroup.com/dlt. Meanwhile, analyst firm Forrester is predicting a massive ‘brain drain’ as skilled professionals leave the infotech industry due to burn-out. There’s more information here: https://bit.ly/3mV8S4.
Evaluating your security for remote working

Tom McVey, Menlo Security

The rise and consolidation of remote and hybrid working models is arguably one of the most significant fallouts that have emerged from the Covid-19 pandemic. Where the daily operations of the vast majority of organisations globally were previously bound to offices, companies today are realising the distinct benefits of cultivating more-flexible working environments that are favourable to all parties.

It’s hard to deny the merits of remote and hybrid working. But equally there are several challenges to consider – not least those connected with security.

A recent survey of more than 500 IT decision-makers in the US and the UK found that while 83% of organisations are confident in their ability to control access to applications for remote users, three quarters are currently re-evaluating their security strategy in the wake of new ways of working and the growth in cloud application use (https://bit.ly/3vJ5yVI).

This review process is critically important. The differing security protocols that work for on-premises setups and remote cloud-based architectures are worlds apart, and organisations must update and adapt to protect their people and assets from cybercrime. However, it is just as critical that these reviews lead to the right kind of outcomes.

According to the survey’s findings, 75% of organisations continue to rely on virtual private networks (VPNs) for controlling remote access. The fact that many businesses continue to rely upon traditional and inherently insecure security protocols such as VPNs is a significant problem. Yet with many conducting security reviews, this does provide a real opportunity to improve their security postures.

The research shows that little over a third (36%) of organisations are taking a zero-trust approach as part of their remote access strategy, despite this being a highly effective means of drastically improving an organisation’s overall security posture in one swift transition. Zero-trust policies ensure that users are only provided with access to those applications and resources they truly need to do their job effectively. While traditional security models tend to assume that everything within an organisation’s network should be trusted, zero trust flips this on its head.

A key reason why some of the most notorious cyber attacks of recent times have been so damaging, including SolarWinds, is down to the ability of hackers to move laterally within a network, accessing and exfiltrating data and elevating privileges without any meaningful resistance. Zero trust drastically reduces the chances of this happening, shifting away from legacy ‘castle and moat’ security policies and taking a new approach that is rooted in the principle of continual verification. It recognises that trust is a vulnerability, and so commands that all traffic – be it emails, websites, videos, documents or other files that originate from either inside or outside an organisation – must be verified.

Circling back to the survey, the vast majority of organisations agree that zero trust is a logical approach. Three-quarters of respondents believe that hybrid and remote workers accessing applications on unmanaged devices pose a significant threat to their organisation’s security. Meanwhile, 79% state that while they have a security strategy in place for remote access by third parties and contractors, there are growing concerns about the risks they present, with just over half (53%) planning to reduce or limit third-party/contractor access to systems and resources.

Indeed, this is a sensible course of action. Controlling user access to private applications is more important than ever, and zero trust is a crucial way in which this can be achieved. It is the perfect starting point for transforming security to deal with modern threats, entailing the continuous authentication of all available data points, limiting user access to specific applications, and reducing risks by assuming that a breach is always imminent.

Indeed, it can be daunting in terms of knowing where to start in implementing a comprehensive zero-trust architecture. But working with an expert or security provider can make this an easy, seamless process, securing access to applications from all devices to minimise IT and security workloads. Providers may also have access to incredibly useful tools such as isolation – a technology that can help to achieve zero trust in its truest sense.

Through isolation, the browsing process is moving from the desktop to the cloud, creating something of a digital ‘air gap’ between the Internet and the endpoint. Here, all content is safely rendered to ensure that complete peace of mind is maintained throughout all daily tasks. All email and web traffic moves through the isolation layer in a seamless, user-experience-focused manner, where the content is visible but never downloaded to the endpoint. As a result, isolation-based zero trust does not leave anything to chance.
After the pandemic: securing smart cities

Kevin Curran, Ulster University

Covid-19 has forced us all to rethink the way we live, work and socialise, with technology proving how vital it has become to cities and how have they responded to the pandemic. Indeed, innovations such as the Internet of Things (IoT), 5G and location-based services have been used to help minimise the risk of transmission, maintain social distancing measures and ensure the continuation of vital services.

According to Capgemini’s ‘Fast-forward to the future: defining and winning the post-Covid new normal’ 2020 report, the pandemic has ‘cemented technology’s role at the heart of transformation, driving new ways of interaction, sharing, engaging, and decision making’.3 Undoubtedly, these technologies will play a major role in the UK’s ‘Build Back Better’ recovery strategy as local councils and political leaders are now taking ‘inclusive measures to pair economic revival with environmental sustainability, urban mobility and energy efficiency’.

“What smart city technologies, when deployed correctly, can be implemented at scale, allowing growing populations to be serviced more far more easily”

Governments around the world will now be looking to deploy new technologies and innovations at scale within a city’s ecosystem as society begins to return to ‘normal’, whatever that may be. This includes a mixture of residential, industrial, commercial, retail and public sector bodies alongside greenways, parks and the public realm.

Ultimately, governments are aiming to transform the delivery of public services through a citizen-centric approach, resulting in greater efficiencies and more responsive services that can drive inclusive growth. For example, the City of London Corporation is currently reviewing its long-term strategy and planning policies – not only to account for flexible working practices but also push for the adoption of new smart city technology and renewable energy networks, including making 5G and broadband readily available for business across the centre of London.

What is a ‘smart city’?

A ‘smart city’ fundamentally relies on IoT to deliver all vital public services. These include addressing problems with clean water, air pollution, traffic and landfill waste. Sensor-enabled devices can help monitor the environmental impact of cities and collect details about sewers, air quality and garbage. Smart city technologies, when deployed correctly, can be implemented at scale, allowing growing populations to be serviced more far more easily.

While a true smart city is designed from the ground up, many cities are now integrating technologies that operate over IoT to improve public services. Some argue that the future growth of the planet’s population can only be sustained through scalable smart city technology. One area, of course, is traffic management, which can be improved using smart traffic lights, road-implanted sensors and even communications with future ‘smart-cars’.

Sensor-enabled IoT devices deployed in smart cities can also help to monitor the environmental impact of cities, collecting details about sewers, air quality, rubbish and energy consumption. Connected technologies can also be used to increase awareness and visibility into individual energy and resource use. IoT-enabled thermostats can make transparent decisions to turn heating on, based on fluctuating energy costs. Moreover, smart IoT water management sensors, in combination with data analytics programmes, can provide consumers with increased visibility into the amount of water they use. Devices such as smart meters that increase visibility into usage have been proven to save money, as well as conserve natural resources.

However, one must remember that smart eco cities are basically cities that fundamentally attempt to integrate technology to achieve efficiencies in a multitude of domains. Other examples of integrating technology into cities are smart lighting, which only turns on in conjunction with nearby traffic or pedestrians; rubbish bins that alert when they need to be emptied; water sprinklers that autonomously test the soil conditions and turn on watering as required; and smart meters which remove the need for humans to check.

Increased connectivity

In the future, we will also require smart cars to become more integrated with national intelligent transport infrastructures to ensure that vehicles can operate safely and efficiently. Satellite navigation and traffic signal control systems will ensure that vehicles know when to stop, slow down and speed up as well as identify hazards in good time. This communication will result in better traffic management and significantly reduce the number of accidents. Ultimately, the roads beneath us will communicate with smart cars, most likely through indestructible sensors embedded within the road.

In fact, with smart technology, it is now possible to access live data, allowing real-time reporting of a structure’s condition, enabling managers to remotely monitor and predict routine and emergency maintenance. A key sensor can be worth a thousand visits by an inspection engineer, as it can alert to different patterns of frequency
and life. For instance, consider basic city-level infrastructure in a location that is regularly affected by extreme weather. When it comes to bridges, crucial road links and mainland connections, upkeep needs to be carefully monitored in order to keep a city running smoothly and safely. Wireless IoT bridge sensors can keep track of all aspects of a bridge’s health, collecting data in areas such as vibration, pressure, humidity and temperature. This data can be used to predict early signs of damage and deterioration, as well as monitor overall traffic volume.

Not without risks

However, increased connectivity between vehicles and wider national infrastructure is not without its risks. Modern vehicles have evolved to contain a complex network of as many as 100 independent computers, or electronic control units (ECUs). ECUs perform a variety of functions such as measuring the oxygen present in exhaust fumes and adjusting the fuel or oxygen mixture, improving efficiency and reducing pollutants. Gradually these ECUs have become integrated into nearly every aspect of a vehicle’s functioning, including steering, cruise control, airbag deployment and braking.

As electronics and related code become more integrated into modern vehicles, we are reaching a point where they will require similar protection to that of smartphones, tablets and traditional computers. There is a real worry about hackers controlling vehicles in different scenarios, from downloading rogue apps, to disabling the vehicle’s ignition or potentially overriding braking systems. The universal controller area network (CAN) bus on vehicles makes such breaches possible. Important aspects such as the speed control, steering and brakes are all located on a separate vehicle network, but there is still interconnectivity between both vehicle network backbones so that a breach in one can cause havoc in the other. It is still proving to be a rather difficult system to breach, but as more and more exploits get shared on the Internet, there is much cause for concern. As vehicles become more integrated into wider networks, there could be serious consequences.

Critical services exposed

Smart cities are inevitable but the introduction of advanced technologies into the fabric of a city comes with varying risks. Relying on a central technological hub to control an infrastructure can allow hackers to target a city more easily than ever before and a smart city is only as secure as its weakest part – and, increasingly, we are finding that this is an IoT device. Any part of a smart city infrastructure could be compromised – for instance, the street light system could potentially be targeted in a denial-of-service attack, leading to widespread blackouts.

Training of those who install IoT devices – such as gas engineers and plumbers – was outlined in an earlier UK Government report on the Internet of Things, which focuses on security by design.4 In this, leaders have proposed for the first time in the UK that providers should have to undergo mandatory cyber security training to prevent smart devices from being exploited by criminals or state-sponsored attackers. This can be done by having security professionals work more closely with industry bodies to embed IoT training as standard.

Ransomware attacks

Ransomware presents a continuous challenge and attacks are growing more sophisticated by the day. In fact, recently it has led to serious disruptions to vital services – as we saw earlier this year, with the ransomware attack by the group DarkSide, on the fuel pipeline carrying 50% of fuel in North America. Consider how an attack would affect our wider traffic or energy infrastructure. If other IoT innovations are further integrated, there will be an endless number of endpoints which nefarious actors can use to their advantage – and a single loophole could be catastrophic.

Ryuk ransomware is possibly the best example, due to its widespread popularity. It is a very sophisticated ransomware threat that has been targeting hospitals, government institutions, businesses and other organisations for the past five years. The group behind the malware is known for using manual hacking techniques and open-source tools to move laterally through private networks and gain administrative access to as many systems as possible before initiating the file encryption. Some attackers have adopted a ‘radio silence’ technique, through a sophisticated monitoring of system processes, where malware knows when to stay silent or lie dormant. There are in fact some really impressive ‘stealth mode’ techniques adopted by malware to evade detection. Techniques include frequently checking AV results and changing versions and builds on all infected servers when any traces of detection appear, in addition to monitoring memory consumption to prevent common server administration utilities from detecting the ransomware processes.

Adept measures

It is essential to deploy connected devices with sufficient security policies such as firewalls and intrusion detection and prevention systems. It’s also important to ensure there is confidentiality of customers’ data with encryption, strong passwords and certificate-based authentication across all devices. Device management agents can highlight failed access attempts and attempted denial-of-service attacks. To ensure the city is as safe as possible, all non-IoT devices must also be patched and kept malware free.

As indicated in earlier government reports, in order to prevent smart devices from being exploited by criminals or state-sponsored attackers, security professionals should work closely with industry bodies to embed IoT training as standard for engineers who install any connected devices. In fact, in the future, these engineers will have to understand the inherent risks of any given IoT device if they are planning on applying it to a wider public network.

City planners will also need to consider the public’s own cyber security awareness. In a recent survey conducted by the World Economic Forum (WEF), safety and security, and privacy and trust likely pose the greatest levels of risk, especially in the consumer IoT domain.5

Users are often unaware of the responsibility IoT manufacturers and service providers bear in order to mitigate privacy
risks, and the regulations that they have to meet with, with regard to how personal data is collected. However, users often lack the awareness and experience needed to properly manage their own exposure to IoT security risks. If more and more devices are added to city infrastructures and linked to vital public networks, including consumer IoT products, the public needs to have an understanding of the risks involved, and how we all have a part to play to mitigate any potential threats or unnecessary risks.

Plentiful risks

While the benefits of smart cities and further connectivity are plain to see – from more-efficient distribution of resources, improved road traffic management and safety, and reducing our carbon footprint, the cyber security risks are plentiful.

Industry and governments are working hard to respond to potential privacy threats: however, the road ahead will be challenging. Mapping a secure online or digital environment of this magnitude, with a multitude of endpoints will be tricky, and no project of this size has ever been completed before. Hackers and other nefarious actors will test its limits.

About the author

Kevin Curran is a professor of cyber security at Ulster University and an IEEE senior member. As executive co-director of the Legal Innovation Centre and group leader for the Cyber Security and Web Technologies Research Group, he has made significant contributions to advancing the knowledge of computer networking and systems, evidenced by more than 800 published works. Regarded as one of the top cyber security experts within the UK, he regularly comments on the latest technological developments and cyberthreats, including the Internet of Things (IoT) and smart devices, crypto-currency, phishing-attacks and ransomware.

References

1. ‘Fast Forward to the Future’.

Smart plugs invite cyber criminals into the home

Richard Hughes, A&O IT Group

The Internet of Things (IoT) is one of the fastest-growing technology markets and has steadily picked up steam in recent years as technology improves and price points drop, with the global IoT market forecast to reach revenues of $1.1tr by 2024. Fields including manufacturing, healthcare and transport have benefited hugely from the ability to create linked networks of smart devices, facilitating a growing level of operational automation and visibility.

The consumer IoT market has also grown rapidly and saw a further boost during the pandemic. Research commissioned by the UK Government found that almost half of consumers had bought at least one connected device over the course of the pandemic, including smart watches, TVs and cameras. However, while the flourishing market means more choice and competitive prices for enterprises and consumers alike, it has also contributed to the large number of devices that skip out on security in favour of low costs.

The market has long been overflowing with devices that lack basic security capabilities such as data encryption, or designs that make it difficult for users to carry out standard activity such as changing default passwords or applying updates. As a result, multiple regions are seeking to introduce legislation that will improve the security of connected devices, with the UK’s Department for Digital, Culture, Media & Sport recently announcing plans for laws that will ban weak default passwords and make it mandatory for smart device manu-
We selected two models of smart plug for investigation – the Sonoff S26 and the Ener-J WiFi. Both of these can be easily found for less than £15 on major online retailers such as Amazon, eBay and AliExpress, and are fairly representative of lower-cost smart plugs. The main focus of the investigation was to determine how these weaknesses could be abused by installing (flashing) malicious firmware into the device, exploiting it as part of a supply chain attack.

Security failings pile up

Before either smart plug could be used, it had to be paired with a mobile phone app. For the Sonoff S26 this was the eWeLink app and for the Ener-J this was the ENERJ SMART.

We started off with the Sonoff S26, which broadcasts an SSID secured with a WPA2 pre-shared key (PSK) once placed into pairing mode. So far so good, but a quick Google revealed the PSK to be simply ‘12345678’.

Implementing such a weak default password is bad enough, but the error is compounded by it being readily available in the online user manual, when users aren’t actually required to know this information in order to work the device. Armed with this PSK, a threat actor could monitor or intercept any communications between the smart plug and paired mobile app.

The use of a packet sniffer revealed that the Sonoff S26 and the app exchanged unencrypted data, including the wifi credentials, which was passed on to the smart plug. These could be exploited to clone the plug or use its API key to interact with the cloud server. Any attacker would easily be able to scoop this information up and freely join the user’s network, at which point he could access and attack other devices connected to it. From here a threat actor would have free rein to exploit vulnerabilities in other devices, such as accessing sensitive data, monitoring traffic or taking control of smart devices.

Once we finished with the Sonoff, we went through the same process with the Ener-J device. This model proved to be a little more secure as message-level encryption was applied to prevent the credentials being easily intercepted. Nevertheless, there are more secure routes that both models could have followed, such as having a unique PSK for each device included in the packaging.

### Malicious firmware

Next, we moved onto our main objective of assessing how easily an attacker could alter the firmware on a device – something that requires manually tampering with the product. Firmware vulnerabilities represent a serious threat, with Microsoft recently finding that 80% of enterprises have suffered at least one firmware attack in the past two years.\(^4\)
Both models were quite easy to open and reassemble with no trace of any tampering so we were quickly able to identify the brains of the device and how to access and upload firmware. Both models allowed firmware to be downloaded or uploaded (flashed) via four connections between the chip and a USB-to-serial converter, a cheap and readily available tool.

The trick to flashing

Like any other kind of malware, malicious firmware can be crafted to carry out a wide variety of functions and aid the threat actor in an attack. In this case we decided to exploit the Sonoff’s lack of security around wifi data during the pairing process.

After a few hours of work, we were able to create some rough but functional malicious firmware that would pair with the mobile app, retrieve the Wifi credentials, and connect to the network as usual – and then call home to send us the SSID, PSK and location of the smart plug. While the plugs lacked GPS functionality, this last piece of information was achieved by scanning for the strongest wifi base station ID (BSSID).

This is an effective form of supply chain attack, as the compromised plug completely bypasses most standard security defences. What’s more, installing malicious firmware is simple and low-cost, requiring only a USB-to-serial adapter available for under £5.

Minimising the threat

There are several steps consumers can take to reduce the risk of a cheap smart device inviting an attacker into their home.

Before making a purchase, users should conduct a risk assessment and consider the damage a connected device could do if it were compromised and remotely controlled. This is particularly important for devices such as kettles and heaters that could start fires, as well as devices with visual and audio recording capabilities. This assessment should also extend to the risk posed to the network and the devices connected to it.

Once a device has been purchased, there are a number of other steps users can take to secure their networks. First, the device should be examined for any signs of physical tampering before it is connected to the network. Users should also consider separating the device from their main network, which can be achieved by connecting a second access point to the router.

Finally, the device’s connections can be monitored with a packet sniffer tool to confirm that they are valid. All other devices on the network should be fully patched and updated to minimise their risk of compromise by an intruder.

Taking more action

We would also urge smart device manufacturers to take more action to prevent their products being exploited. For example, hardware components should ideally require a cryptographically signed firmware

Preventative actions taken by organisations with regard to firmware threats. Source: Microsoft.

Answers to the question, ‘When purchasing a smart device, have you ever checked to see if the device has a default password that is not unique to it?’. Source: Ipsos Mori.

Both models were quite easy to open and reassemble with no trace of any tampering so we were quickly able to identify the brains of the device and how to access and upload firmware. Both models allowed firmware to be downloaded or uploaded (flashed) via four connections between the chip and a USB-to-serial converter, a cheap and readily available tool.

The trick to flashing

Like any other kind of malware, malicious firmware can be crafted to carry out a wide variety of functions and aid the threat actor in an attack. In this case we decided to exploit the Sonoff’s lack of security around wifi data during the pairing process.

After a few hours of work, we were able to create some rough but functional malicious firmware that would pair with the mobile app, retrieve the Wifi credentials, and connect to the network as usual – and then call home to send us the SSID, PSK and location of the smart plug. While the plugs lacked GPS functionality, this last piece of information was achieved by scanning for the strongest wifi base station ID (BSSID).

This is an effective form of supply chain attack, as the compromised plug completely bypasses most standard security defences. What’s more, installing malicious firmware is simple and low-cost, requiring only a USB-to-serial adapter available for under £5.

Minimising the threat

There are several steps consumers can take to reduce the risk of a cheap smart device inviting an attacker into their home.

Before making a purchase, users should conduct a risk assessment and consider the damage a connected device could do if it were compromised and remotely controlled. This is particularly important for devices such as kettles and heaters that could start fires, as well as devices with visual and audio recording capabilities. This assessment should also extend to the risk posed to the network and the devices connected to it.

Once a device has been purchased, there are a number of other steps users can take to secure their networks. First, the device should be examined for any signs of physical tampering before it is connected to the network. Users should also consider separating the device from their main network, which can be achieved by connecting a second access point to the router.

Finally, the device’s connections can be monitored with a packet sniffer tool to confirm that they are valid. All other devices on the network should be fully patched and updated to minimise their risk of compromise by an intruder.

Taking more action

We would also urge smart device manufacturers to take more action to prevent their products being exploited. For example, hardware components should ideally require a cryptographically signed firmware

Preventative actions taken by organisations with regard to firmware threats. Source: Microsoft.

Answers to the question, ‘When purchasing a smart device, have you ever checked to see if the device has a default password that is not unique to it?’. Source: Ipsos Mori.
Coming off the tracks: the cyberthreats facing rail operators

Alex Cowan, RazorSecure

The rail industry has always been committed to ensuring safe and reliable journeys for passengers. It is an industry made up of a host of different organisations that all have specific safety responsibilities and defined safety duties, as detailed in regulations and standards. However, as trains have gone through a digital transformation and adopted new connectivity and devices, the risk of a new kind of threat has emerged – that of a cyber security attack. So let’s take a look at how the risk has manifested and what railway operators, train manufacturers and other industry parties can do to address this.

The rail industry is somewhat unique in that many of the trains being used on the rail networks today still pre-date the digital age. The shelf life of rail rolling stock can be more than 25 years. During this time, trains may be bought and sold between companies or even leased from one party to another. New systems, including connected technology, will be added as it becomes relevant to the needs of the train service provider.

A typical modern digital train includes upwards of 100 digital systems. Each of these is potentially a vulnerability unless it is protected, and with a mid-size fleet of 100 trains there are tens of thousands of systems that require protection. The sheer number of systems, some of which might be difficult to update, can present a large attack surface.

These systems will include, but not be limited to passenger wifi systems, an operational management system, CCTV systems, passenger management systems, automated door controls, plus passenger information and entertainment systems. In each case these may have been installed at different times, to different standards. These might be connected via hubs or routers either on each single carriage or via a connectivity device for the entire


References

References
train, depending on the circumstances. Many of these systems may not have been designed to be connected to the outside world but have been so connected as an afterthought. Furthermore, these systems can be very difficult to physically access and may not have been patched or updated for many years. As has been demonstrated in several high-profile cyber security attacks in the past, older software that is harder to update is often the most vulnerable to attack.

All-encompassing security

The rail industry has an incredibly effective culture around safety – the rail network is a very safe place to work and commute. The growing importance of cyber security to the rail industry’s long-term commitment to safety is now driving the industry to consider that if a train is not secure then it can no longer say that it is safe.

“With a mid-size fleet of 100 trains there are tens of thousands of systems that require protection. The sheer number of systems, some of which might be difficult to update, can present a large attack surface”

Most information flow in rail networks involves communications from operational technology (OT) control systems to operational and comfort systems onboard the train. Enforcing separation between networks is a key requirement of an effective security strategy for the industry, though this is not an area typically addressed by the traditional cyber security industry, which primarily focuses on issues around data protection rather than operational systems.

Thanks in part to these issues, discussions in the past five years around cyber security on rail networks have evolved. The former concern of the industry was to ensure the protection of data – be it passenger information or payment card details – based in computer storage systems. Increasingly, in the past two years, these conversations have evolved to encompass operational security from cyber attack as well as protection of data. This change is being driven by two important factors:

• The new CENELEC TS50701 Technical Specification for railway cyber security will push rail companies on designing and implementing more-effective cyber security programmes to better manage cyber security as a safety risk in legacy and new-build trains. 1
• New standards and legislation for operational technology – including NIS Directive (Europe) IEC62443 (Global), NIST Cybersecurity Framework (US) and AS7770 (Australia) – are giving clearer guidance on the requirements for both rolling stock and signalling as related to protecting the industry from cyber attack. 2–4

Evolution of networks

The movement of trains (known as rolling stock within the rail industry) is implemented by drivers, who stop and start, accelerate and brake. However, unlike cars, where line-of-sight driving is possible, a train driver often cannot see far enough ahead to be able to actively decide when to stop, accelerate or brake. This information is provided by signalers – the people with an overall view of the network, who know where the trains are, where blockages are and what action to take to keep the network running smoothly. Historically the signalers shared information with the train drivers in the form of trackside signs and traffic lights that told the drivers when to slow down and when to expect the unexpected.

Over time, this process has been increasingly digitised. For the travelling public, a more connected and integrated network between the track and the train translates into a smoother running rail system with fewer disruptions. This in turn means it is possible to run more trains on the same volume of track. These are both important considerations for passengers and the industry, as they drive efficiency.

With increasing digitalisation, signal-ling information has been broadcast onboard a train (so the driver looks in his cab for information instead of out of the cab window). From a purely safety perspective, this makes sense, since it ensures that train drivers are less likely to miss trackside signals. In-cab signals can be supplemented with sounds and alerts. But it also facilitates increased and strengthened interactions and communications between the train and trackside. And this can lead to new vulnerabilities within systems that could be exploited by hackers.

Increased risk

Like most industries, the rail system is aiming to be more efficient by improving system automation and increasing the number of processes that are remotely managed. This means combining networks, many of which might have been custom built for this purpose. Even with the best intentions from network designs, it can result in unintentional vulnerabilities for hackers to exploit.

Historically the cyber security industry has focused on protecting data from viruses or attacks that are typically launched against machines, of which there are millions worldwide. Usually, the first attack is successfully identified and then coders work day and night to find a patch to remove the vulnerability.

“It facilitates increased and strengthened interactions and communications between the train and trackside. And this can lead to new vulnerabilities within systems to be exploited by hackers”

Within days this has been distributed as a security update and, apart from the computers that were first attacked, all other systems are secured against risk. This ‘fast response’ method of protection works well in an enterprise environment because there are millions of Windows PCs and Linux servers that exist around the world which typically have the same or similar functionality and vulnerabilities.
A rail environment is completely different. Each fleet will have a unique network design with a variety of systems that may not have been integrated previously. The fleet itself operates over a long period of time and is subject to configuration drift as maintenance occurs, systems are replaced and requirements change. This creates an environment that is unique for every rail network or operator globally.

The further challenge for the industry is that, at any one time, there is a significant number of people accessing or trying to access rail networks. A cyber attack has the potential to target safety-critical systems, such as those that control the train’s speed. An attack on a signalling system could cause untold disruption to train movements, leading to further safety concerns. The 24/7 nature of the rail networks means that there is no time to consider whether a system failure is just that or is due to a more malicious motive. In an industry steeped in safety-first, the natural (and proper) response is to focus on the lowest possible risk.

“Only by combining resources and having an overview of the entire system – both onboard rolling stock and trackside – can the industry fully address threats”

For these reasons, the rail industry requires a holistic combination of techniques to improve cyber security. The solution is not simply the deployment of a firewall or a single control. Cyber security for the rail industry requires real-time monitoring and management over the entire life cycle of a train.

Identify, protect, record

To ensure best practice in addressing the issues around cyber security, the rail industry needs to have a clear understanding of the numerous connected train and trackside systems that are in operation. This may sound like a simple task but, as previously outlined, trains can often have a lifespan of 25 years or more and will undergo many changes during that lifespan, with new connected systems added by each owner or operator.

As information systems are added, replaced and connected, vulnerabilities can appear. The key priority is therefore to identify all the connected systems on a network and understand the behaviour and traffic flows between them.

Next the industry must implement the proper cyber security monitoring systems to ensure that anything out of the ordinary is quickly identified and shut down in real time. This facilitates real-time monitoring systems that understand what a normal pattern of behaviour is and can quickly identify something that is unusual or unexpected. This should trigger an immediate response to ensure that the primary issue of passenger safety is assured, while the risk is investigated. When the risk has been identified and addressed, a continuous process of updating systems can help to ensure that systems ‘learn’ from what has happened previously and get better, over time, at identifying what is unusual.

Finally, these monitoring systems need to record all activity monitored for future reference. This effectively creates a cyber security ‘black box’ for the railway industry to record and identify potential cyber risks so these can be stopped more effectively in real time and networks can adapt to new threats faster.

Cover all bases

The rail network creates challenges for cyber professionals and safety officers that are difficult to address. Trains are built to last, but this can lead to issues as more systems are connected over time unless train manufacturers and rail companies can keep a full record of every computer system added to every train and carriage on every network. This issue, combined with the challenge of updating computer systems on trains that are often being used almost 24-7, creates a risk environment unique to the industry.

While connected devices and systems help to drive important efficiencies, they also leave wider landscapes for attack, now and in the future. Only by combining resources and having an overview of the entire system – both onboard rolling stock and trackside – can the industry fully address threats. Despite this being a monumental task, there are reasons for positivity. The improved awareness of the risk of cyber attack, paired with its potential impact on safety, suggests that the industry is focusing more closely on this issue.

“A cyber attack has the potential to target safety-critical systems, such as those that control the train’s speed. An attack on a signalling system could cause untold disruption to train movements, leading to further safety concerns”

The historic track record of rail in delivering a fast, effective and extremely safe environment for travel also means the industry is in a prime position to focus on risk and react to it quickly. By taking an approach where cyber risk is identified fast, addressed immediately and recorded for the future, the industry can lead the way and ensure that its leadership in safety management continues in the digital era.

About the author

After 15 years’ experience in the gaming industry, Alex Cowan founded rail cyber security specialist RazorSecure. Since 2015, RazorSecure has been using machine learning to protect rolling stock, signalling and infrastructure systems. Employing his expertise from gaming, Cowan built RazorSecure’s technology, which provides intrusion detection cyber security software offered as a ‘software as service’ (SaaS) to the rail industry.

References

Protecting Active Directory against modern threats

Guido Grillenmeier, Semperis

Active Directory (AD) is something of an Achilles heel in the security posture of many organisations. There is no getting away from the fact that AD is vitally important. Operating as a database and set of services connecting users with network resources, it underpins much of the ability for employees to work seamlessly and efficiently on a daily basis. The challenge, however, lies in the fact that these directories contain critical information about your environment – from how many users and computers there are, to who has which permissions – while also housing sensitive information such as job titles, phone numbers and passwords.

In this sense, AD is something of a treasure trove within the organisation. It contains the keys to the kingdom and a map that reveals where to find resources that contain value. And naturally, cyber attackers are increasingly finding the value in getting hold of these keys.

Compromised directory

Ransomware attacks often stem from AD. In compromising the directory, attackers pave a clear path from which they may access all other applications, making it easy for them to go after sensitive business data, extract and encrypt it, before holding it against the victim organisation until it pays a requested ransom.

There are clear concerns being voiced from key bodies such as the National Institute of Standards and Technology (NIST) about the increasing number of attacks going after AD, with cyber criminals continuing to deploy an ever-rising number of dangerous tactics and techniques. To deal with this, any sound cyber security strategy must incorporate detection in order to spot hackers gaining access to, moving around within, or administering a network. Yet this is easier said than done, and organisations still struggle in the way of detection.

According to Microsoft, Mandiant and Lockheed Martin, the median number of days an attacker sits within a network and goes undetected is anything from 146 to 229. This is largely down to the abilities of today’s attackers, who are becoming increasingly adept at operating stealthily. However, at the same time, organisations often simply don’t have the right tools, skillsets and/or capabilities to be able to detect these activities in the first place.

Log consolidation

Let’s consider how domain controller event log consolidation and security information and event management (SIEM) solutions operate. These are common detection methods used today.

Every network records traffic in the form of event logs, which essentially act as logbooks that capture various types of information: who logged on via what computer; who created a new folder; failed password attempts; and many more common actions.

A SIEM solution adds logic to this, centralising all event log data in a main repository that is then continually monitored for anomalies and potentially malicious activity. Its benefit is that any such behaviour can easily and concisely be communicated to IT and security professionals for them to investigate further.

Although it has been around for a while, the SIEM approach is still a critical part of any organisation’s security posture, helping to detect many AD-related threats. Yet there are now modern techniques and tactics that intruders use that hamstring the effectiveness of event logs. This includes deactivating SIEM’s audit capabilities, or utilising tools that remove the audit trail: in such cases, SIEM is blindsided.

That’s not to say that companies shouldn’t use SIEM. The crucial point is that security is a multi-faceted topic and you therefore need to secure yourself from multiple angles. SIEM is still valuable, but you need to be aware of what you can’t monitor with this one single solution and cover any additional gaps that might be exposed.

Mimikatz and DCShadow

So, what can’t SIEM cover? Various attacks have been seen in the wild that leave no discernible trail in SIEM – or at least any evidence of malicious activity.

One example of this is the use of the DCShadow feature of a commonly used hacker tool – Mimikatz. With this mechanism, the intruder modifies the configuration partition of Active Directory to register any member server as a rogue domain controller that was never promoted to a real DC but is essentially trusted by other DCs in the same way as a true domain controller. From this point, the attacker can then make unauthorised malicious changes, such as tweaking group memberships of domain admins, or adding the SID of the domain admins group to the SIDHistory attribute of a compromised normal user.

Notably, the rogue DC injects these changes directly into the replication stream of the production domain controllers, which will not trigger any event-log entry to be written to any log that could report them – a technique that ensures that traditional SIEM-based log collection is bypassed.

Group Policy

Group Policy changes are also often an evasive method used by attackers. In March, many Spanish government agencies were hit with the Ryuk ransomware.
In this particular attack, changes were made to a Group Policy object that propagated the installation of Ryuk to remote endpoints within the victim organisations.

The issue with Group Policy is that event logs don't include details that explain what changes are made. You can see that a change has been made, but not what was involved. As a result, the SIEM is unable to differentiate between a malicious change, such as Ryuk, versus an ordinary, routine, operational tweak, and so no alarm bells will be set ringing.

**ZeroLogon**

A third known example comes in the form of ZeroLogon – an attack type that is becoming increasingly prevalent. With ZeroLogon, a proof-of-concept exploit code was made public, allowing attackers with network access to domain controllers to send a series of Netlogon messages consisting of streams of zeros. This forces the domain controller computer’s password to be changed to an empty string, in turn providing the attacker with ownership of the domain controller.

Executing ZeroLogon therefore provides complete control of an organisation’s crown jewels. From here, attackers can perform an endless series of changes in AD, using it as a path to attack other systems in your infrastructure.

In the case of SIEM, a single password change is unlikely to be flagged as suspicious. While many organisations adopt a policy whereby user passwords are updated every month, one additional password change outside of this is highly unlikely to be considered suspicious, and an attacker is therefore able to perform changes while going under the radar.

**Best practices**

Be it ZeroLogon attacks, Group Policy attacks or DCShadow, each of these different attack techniques has been specifically designed to bypass traditional detection solutions such as domain controller event log consolidation and security information and event management solutions.

What’s more worrying is that there’s a high chance that additional threats facing AD are out there, waiting to be unearthed. Threat actors are incredibly smart, and they have time to concentrate on finding new bugs that are able to circumvent any form of detection which the general populace simply isn’t aware of. And while companies continue to pay out ransoms to hackers, the incentive for them to do so will remain.

With this in mind, there is always a chance of being attacked, and new techniques will continue to emerge. Indeed, the fact that hackers will always seem to be one step ahead as a result makes the question of possible solutions seem complex – how can you prepare for something that you’re not aware of?

But it shouldn’t be complex or daunting. By following best practices, organisations can better protect themselves and their businesses from many threats, both new and existing. So, what are these best practices?

“Any sound cyber security strategy must incorporate detection in order to spot hackers gaining access to, moving around within, or administering a network. Yet this is easier said than done”

First, it is important to always monitor for malicious changes within Active Directory. While SIEM has its role to play as something of a first line of defence, organisations should go beyond this to adopt additional solutions capable of reading and understanding the replication traffic on the domain controllers themselves. By adding these additional tools as second and third lines of defence, organisations will secure peace of mind, knowing that they are able to transparently see any AD-related change.

These attitudes should also extend to changes within the Group Policy too. As we discussed, the issue with Group Policy is that event logs don’t include details about what specific changes are made. However, there are solutions available that can define specific protected objects to be monitored for any alteration, such as changes in membership to domain admins. Any time those specific protected objects are modified, the solution is able to detect these modifications and notifies the relevant network security teams.

It is equally important to gain awareness of the tell-tale signs of DCShadow so that your teams may ensure they’re not being used on your network. By default, Mimikatz leaves identifiable artifacts behind. So, by consistently reviewing AD for these artifacts, you will be able to spot when and where malicious activity might be taking place. Once you find a trace of Mimikatz in your environment, you must act quickly as you’ll already be a victim of a DCShadow attack. To help you with this, there are once again solutions available on the market that can quickly and easily show you what changes were performed at the replication level, which can then be analysed and addressed.

**Ability to react**

This leads into a third key point – that it is just as important to be able to react to attacks as it is to be proactive in defending against them.

As has been mentioned already, attackers are continuing to unearth new vulnerabilities, and there is, therefore, always a chance that even the most protected organisations will be breached.

“There’s a high chance that additional threats facing AD are out there, waiting to be unearthed. Threat actors are incredibly smart, and they have time to concentrate on finding new bugs”

For this reason, having a recovery plan in place is vital. In the aftermath of an attack, the most important thing is getting your business back quickly so that you can continue to serve your customers, and that begins with recovering your AD. However, it’s very difficult to get your AD back from scratch.

If you are the victim of a ransomware attack and need to recover your AD service in its entirety, you may believe that having a good domain controller back-up will be sufficient in achieving this. But this is not the case. A good domain controller back-up does not equate to a seamless and fast AD service recovery. For this to happen,
organisations need to practise the recovery process in the same way they would a fire drill, following the detailed Microsoft AD Forest Recovery Guide. Equally, it’s worth looking for solutions to support this process – some can revert changes down to the attribute level or to protect objects when detected, automating the recovery process, and ensuring that it is executed both effectively and accurately in high-pressure situations.

**Time to get serious**

Given the threats, it is important that organisations act if they are to protect themselves properly. Targeting Active Directory and modifying it to suit the attacker is a common tactic taken by today’s cyber criminal – and it won’t be going away any time soon.

Companies have invested huge sums in applications that integrate with AD. Yet these same applications are simply unable to work directly in a cloud environment, owing to the differing architectures between the cloud and on-prem setups.

Smaller companies will be quicker to adapt and might even be able to change a holistically cloud-based model eventually. However, for those medium-size and larger organisations that have been using potentially thousands of different applications for many years, the transition process will be much longer.

But this does not mean that AD security should be avoided altogether. Indeed, those that are serious about the security and integrity of their AD need to be looking for additional ways to gain visibility into every AD change and have the ability to revert or recover when necessary.

Putting yourself in this position sooner, rather than later, is key.

**About the author**

Guido Grillenmeier is Chief Technologist with Semperis. Based in Germany, Grillenmeier has been a Microsoft MVP for Directory Services for 12 years. He spent 20-plus years at HP/HPE as chief engineer. A frequent presenter at technology conferences and contributor to technical journals, Grillenmeier is the co-author of Microsoft Windows Security Fundamentals. He’s helped various customers secure their Active Directory environments, and supported their transition to Windows 10/m365 and Azure cloud services.

**References**


---

**Layering identity and access management to disrupt attacks**

**Duncan Godfrey, Auth0**

The protection of digital identities forms an integral and growing part of customer relationships. Integral, because the threat of data breaches is ever present and constantly evolving; companies operate in the uneasy position of being one hack away from a potential loss of customer trust. Growing, because customers understand more than ever the worth of their digital identities.

To ensure that the login experience delivers on all levels, companies are having to manage an ever-more complex authentication environment, one in which federated identities streamline logins, and rogue access attempts are identified and thwarted. Protecting customer identity and access management (CIAM) services against online threats is therefore critical for robust cyber security, and to protect corporate reputation.

The Open Web Application Security Project (OWASP) lists broken access control as the most critical security risk to web applications (with the closely related identification and authentication failures in seventh place). In advising companies to use its list to help ensure that web applications minimise risks, OWASP calls out the often-incorrect implementation of authentication and session management that allows attackers to compromise credentials, such as passwords and session tokens.

**Threats against identity**

A range of attack types target CIAM services, among them fraudulent registrations and large-scale credential stuffing.

Attackers may create large numbers of fake accounts to take advantage of signup bonuses, to spread misinformation or to cause damage. This is known as a fraudulent registration attack and unfortunately, it is prevalent. In fact, analysis shows that around 15% of new account registration
attempts can be attributed to bots.² Too many fake accounts may prevent legitimate users from registering. Some will be a cynical attempt to gain access to benefits that come from registration, such as incentives or access to limited-edition goods.

The good news is that thoughtfully implementing your identity security can guard against fraudulent registrations to protect legitimate customers, bolster corporate reputation for reliability and avoid the potential costs of addressing security issues exposed by puppet accounts.

“To get past MFA, attackers have to dramatically increase the time and effort they put into compromising an account, which makes it infeasible to do at scale”

Credential-stuffing attacks are the biggest threat facing identity systems. They are made possible through the combination of password reuse and large-scale data breaches. They gift cyber criminals access to accounts when they try stolen login details against a range of sites. So widespread is the problem that analysis shows that in the first 90 days of 2021, breached passwords were detected at an average of more than 26,600 per day. Credential stuffing accounted for 16.5% of login traffic in the same timeframe, while daily peaks reached higher than 40%.

To protect access to their services and mitigate the risk of such attacks, organisations can – and should – look towards robust CIAM. As customers continue to become more cyber security-savvy, they are likely to ask questions about how companies will protect them and their data. This forms part of a relationship of trust, which is essential as companies seek long-lasting relationships with customers.

The login is where customers experience identity security – it needs to live up to their expectations of protection and privacy, but also convenience. In other words, customers will return when their online experience is simple, safe and enjoyable.

A layered approach

Robust and resilient CIAM is critical in the fight against attacks on digital identities. Traditionally, multiple security products or solutions would have operated together at different layers or locations, such as endpoint, network and cloud, to provide a robust defence. Now, a layered approach to CIAM involves defensive measures before and throughout the authentication workflow.

The challenge is to develop and implement security measures that strike the right balance between increasing friction for attackers, without disrupting the experience for the genuine user. Whether CIAM solutions are in-house, or provided by an identity-as-a-service platform, a number of measures should be applied.

Weak and common passwords are the foundation of many hacks. To foil brute force attacks that rely on such ill-advised security behaviour, enforce password length and complexity, and advise on frequency of password change. It is also advisable to prevent users from repeating their passwords and to compare potential passwords against a dictionary to prevent common choices being registered.

Despite these measures, users will inevitably reuse their passwords across multiple sites and accounts. In fact, one report revealed that 73% of online accounts use duplicated passwords and that more than half (54%) of consumers use five or fewer passwords across their entire online life.³

A breach in one service provides the fuel that attackers need for credential-stuffing attacks, potentially threatening many services. By comparing user passwords against lists of breached credentials, application providers can issue warnings when users are at risk and encourage them to reset their passwords.

Basic steps

There are some basic steps that organisations can take to reduce their vulnerability to login attacks.

Generic failure messages only: Failed login attempts from cyber attacks can still yield useful results for perpetrators. They can gain valuable intelligence from the process, especially if the application returns more than a single generic error message. Detailed failure messages can assist would-be hackers by providing information about users registered in the system. To prevent this from happening, keep attackers in the dark by returning only generic messages.

Limit failed login attempts: Of course, credential stuffing and other large-scale forms of attack are likely to trigger many failures for each successful login. Mitigating attacks starts with limiting failed login attempts and monitoring for spikes in failed logins or a huge number of sign-ups. Evidence of such behaviour should be used to detect attacks and trigger countermeasures, such as challenging attempted logins with a Captcha, to shut them down and prevent future attempts.

Reduced-friction multi-factor authentication (MFA): MFA is essential for securing digital identities as a basic username and password combination does not provide sufficient protection. Passwords are in perpetual risk of being breached and, when they provide the only key to the door,
there is no additional defence once they’ve been compromised. To get past MFA, attackers have to dramatically increase the time and effort they put into compromising an account, which makes it unfeasible to do at scale. So compelling is the case for MFA, that Microsoft suggests accounts are more than 99.9% less likely to be compromised when it is used.4

Limiting friction
Having said that, requiring additional factors for authentication stands to introduce friction to the login process, something businesses and users want to avoid. Friction delays access, causes frustration, hampers productivity and, in the worst of cases, can lead to session abandonment. MFA must strengthen security without inconveniencing users. Application providers can limit friction through step-up authentication, adaptive MFA and WebAuthn-enabled biometric methods.

• **Step-up authentication** allows some resources to be accessed with one set of credentials but requires more credentials (as provided by MFA) to access sensitive resources. This adapts authentication to the importance of the resource and the risk level, should it be exposed.

• **Adaptive MFA** only engages MFA when an interaction is deemed risky, as determined by behavioural data. This can include the user attempting a login from a new device, suddenly logging in from a location too far from their previous attempt to be plausible, or attempted access from a suspicious IP address.

• **WebAuthn-enabled device biometrics** (facial recognition, fingerprint identification etc) provide the best combination of high security and low friction. This method of strong authentication holds tremendous appeal for users and application providers and, as such, uptake is likely to grow substantially over time.

Secure sessions
A server-side, secure session manager that generates a new session ID after login provides another defensive measure. It’s important not to put session IDs in URLs, and to securely store them and invalidate them after a user has ended their session.

Application and service providers should also encrypt password databases. Encryption ensures that, should a database become compromised, it is of no use to the hackers. This helps protect against stolen password use and also makes an organisation a less appealing target in the first place.

Also, the dangers of default settings in a security context are well publicised. Many users leave default admin credentials unchanged, making them an obvious target and leaving systems vulnerable to attack. It’s a simple open back door for cyber criminals that mustn’t slip through the net.

Zero trust means securing identity
Secure access used to be about securing perimeters. Now, increasingly, the emphasis is on zero trust, which comes down to securing identity. That has important consequences for CIAM with the exponential rise of credential-stuffing attacks, fraudulent registrations and the widespread use of breached credentials being key areas of concern for security professionals managing digital identities.

A comprehensive CIAM strategy helps mitigate risk, reduce costs through process automation, improve business agility, enhance the customer experience, and potentially generate new revenue streams.

To ensure these benefits, heightened authentication security measures must be implemented with not only privacy and protection in mind, but also the user experience. When enhanced security increases friction, customers and users may have an unacceptable online experience. Instead, robust authentication needs to come from understanding identity-related threats, by taking a layered approach to CIAM and through safeguarding the user experience.

About the author
Duncan Godfrey is VP of security engineering at Auth0, responsible for making sure that Auth0 is secure by design. He has over 15 years of experience in information security and has worked for BT, the UK Government and Amazon. Originally from the UK, Godfrey lives in Austin, Texas.

References
I have gained inspiration from the Human Factors in Diving community (https://bit.ly/3j1S5xy) to start an ‘A-Z of cyber security’, which we’ll build over the course of the coming months.

**A: Awareness.** Whenever security professionals talk about cyber security, they bemoan the lack of awareness. This assumes that the only way to get people to behave more securely is to narrow the knowledge gap – in other words, ‘if only they knew, they would act securely’.

Many awareness-raising efforts do their best to deliver the information that people need, but they fail to appreciate two important truths. First, information is not the same as knowledge. Most trainers impart information, but attendees seldom get an opportunity to apply their new insights. For information to become knowledge, they need the ability to convert the information and apply it, knowing when to do this.

Second, situational awareness is key. (See: MR Endsley, ‘Designing for situation awareness in complex systems’ in Proceedings of the Second International Workshop on symbiosis of humans, artifacts and environment.) This starts with sensory awareness – what we see and hear. The mind then attempts to make sense of what a person sees and hears based on previous experiences. Note the word ‘experiences’ – not merely information that people have been exposed to, but experiences in applying the information. The final step builds on this sense-making to predict the future – to anticipate what might happen next based on the actions people decide to take.

This means that merely imparting information to employees and checking an awareness-raising box is suboptimal. Awareness is necessary but not sufficient.

Awareness efforts must give people the opportunity to apply the information and to develop new skills. This will close the knowledge gap and also improve situational awareness, which will have been honed during experience-building training exercises.

**B: Briefing.** Many trainers consider that they have briefed employees during awareness-raising endeavours. The kind of briefing that few engage in is related to giving people sufficient information to enhance their just-in-time situational awareness. For example, one employee might spot a phishing message and report it to the security officer. The officer might send an email to warn all staff about the phishing message. Employees are likely to see the warning only after they have opened the rogue email. There is a need for another channel to ensure that people are warned before they process the phishing email.

Of course, the security officer might be able to remove the email from all the employees’ inboxes, but that also misses a valuable opportunity to create a learning experience. It might well be preferable to forewarn and forearm employees, using a different channel. This allows them to view the phishing email knowing exactly what it is. This builds those experiences that they can rely on to enhance their day-to-day situational awareness.

**C: Communication.** Briefing is related to effective communication. The sender of any cyber security-related message has to be aware of: (1) the recipient’s likely response to the message based on the language and terminology it uses; and (2) the emotions it is likely to elicit.

In terms of the first, keep it simple and don’t use acronyms. Make it actionable – tell them what to do with the information you’re communicating. For example: ‘if you see this email, delete it but don’t report it – we already know about it’. In terms of the latter, ensure that negative emotions such as fear or shame are not triggered. This is not conducive to durable experiences they can rely on.

---

**EVENTS**

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.

5–8 December 2021 **Security Weekly Unlocked**
Florida, US
https://events.securityweekly.com/unlocked2021

9–10 December 2021 **ICCS**
Cardiff, UK

10–13 January 2022 **FloCon**
Virtual event
https://bit.ly/2f0WyUm

2–4 February 2022 **IT-Defense 2022**
Berlin, Germany
https://bit.ly/3mh1Ahj

7–10 February 2022 **RSA Conference**
San Francisco, CA, US
www.rsaconference.com

8–9 February 2022 **HackCon**
Oslo, Norway
www.hackcon.org/english/

1 March 2022 **NullCon**
Goa, India
www.hackcon.org/english/
https://nullcon.net/goa-2022

2–3 March 2022 **Cloud & Cyber Security Expo**
London, UK
www.cloudsecurityexpo.com

5 March 2022 **B-Sides Tampa**
Tampa, FL, US