FEBRUARY 2024

FOURTH QUARTER

Adversarial Threat Report: Countering the Surveillance-for-Hire Industry & Influence Operations

Ben Nimmo, Global Threat Intelligence Lead
Margarita Franklin, Director, Public Affairs, Security
Dr. Lindsay Hundley, Influence Operations Policy Lead
David Agranovich, Policy Director, Threat Disruption
Margie Milam, IP Enforcement & DNS Policy Lead
Mike Dvilyanski, Head of Threat Investigations

Meta

TABLE OF CONTENTS

Purpose of this report 3

Executive summary 4

Countering the surveillance-for-hire industry

Trends & notable tactics, techniques and procedures (TTPs) 7

Post-disruption analysis 10

Hardening product security against spyware threats 12

Call to action: recommendations for countering spyware 14

Countering coordinated inauthentic behavior

China-based network 18

Myanmar-based network 19

Ukraine-based network 20

Analysis: **Lessons learned from Russian influence operations related to war in Ukraine**

Overt influence operations (state-controlled media) 21

Covert influence operations 23

Update on our work against domain name abuse 28

Appendix: Threat indicators 29

PURPOSE OF THIS REPORT

Our public threat reporting began over six years ago when we first shared our findings about coordinated inauthentic behavior (CIB) by a Russian covert influence operation. Since then, we have expanded our ability to respond to a wider range of adversarial behaviors – from cyber espionage to spyware – as global threats have continued to evolve. To provide a more comprehensive view into the risks we tackle, we've also expanded our regular threat reports to include other threats and our detailed insights — all in one place, as part of our quarterly reporting. In addition, we're also publishing threat indicators to contribute to the security community's efforts to detect and counter malicious activity elsewhere on the internet (See Appendix).

We expect the make-up of these reports to continue to evolve in response to the changes we see in the threat environment and as we expand to cover new areas of our Trust & Safety work. This report is not meant to reflect the entirety of our security enforcements, but to share notable trends and investigations to help inform our community's understanding of the evolving threats we see. We welcome ideas from our peers to help make these reports more informative.

For a quantitative view into our enforcement of our Community Standards, including content-based actions we've taken at scale and our broader integrity work, please visit Meta's Transparency Center here: https://transparency.fb.com/data/.

EXECUTIVE SUMMARY

In this Adversarial Threat Report, we're sharing updates on our work against the surveillance-for-hire industry, our Q4 takedowns of new CIB networks in China, Myanmar and Ukraine, and an annual update on the adversarial trends we've identified in the two years since Russia began its full-scale war against Ukraine.

Countering spyware: In our third annual report, we share notable trends and tactics across our investigations into the surveillance-for-hire industry targeting people around the world. It includes findings related to eight firms from Italy, Spain and the United Arab Emirates: Cy4Gate; RCS Labs; IPS Intelligence; Variston IT; TrueL IT; Protect Electronic Systems; Negg Group; and Mollitiam Industries.

They targeted iOS, Android, and Windows devices. Their various malware included capabilities to collect and access device information, location, photos and media, contacts, calendar, email, SMS, social media and messaging apps and enable microphone, camera and screenshot functionality. Their scraping, social engineering and phishing activity targeted Facebook, Instagram, X (formerly Twitter), YouTube, Skype, GitHub, Reddit, Google, LinkedIn, Quora, Tumblr, VK, Flickr, TikTok, SnapChat, Gettr, Viber, Twitch and Telegram.

Our post-disruption analysis suggests that a comprehensive threat disruption approach across our industry and society can impose substantial friction on spyware groups and force them to expend more resources to hide and spread their activities across many services online to stay afloat. We've observed that our continuous, repeated disruptions have led to a decrease in activity by these groups on our apps. We share our call to action with specific policy recommendations, in addition to updates on our defense-in-depth approach to hardening our product security against spyware.

Lessons learned from Russian influence operations related to the war in Ukraine: February marks two years since Russia's full-scale invasion of Ukraine and ten years since Russia's annexation of Crimea. Our teams remain on high alert to monitor for adversarial changes related to this war as Ukraine or Ukraine-related issues remain the largest focus of Russian-origin influence operations.

Our enforcements against Russian state controlled media led to posting volumes declining by 55% and engagement levels by 94% compared to pre-war levels globally and across languages, according to the latest research by Graphika. For covert influence operations, since 2022, we've seen fewer attempts to build complex deceptive personas in favor of thinly-disguised, short-lived fake accounts in an effort to spam the internet, hoping something will "stick." Despite these noisy attempts, we've seen a consistent decline in the followings of Russian-origin CIB campaigns.

While we expect spammy attempts at throwing large volumes of accounts across many internet services and websites to continue in 2024, our threat research shows that, historically, the main way that CIB networks get through to authentic communities is when they manage to co-opt real people - politicians, journalists or influencers - and tap into their audiences. Reputable opinion-makers represent an attractive target and should exercise caution before amplifying information from unverified sources, particularly ahead of major elections.

Finally, unlike recent China-origin influence operations which engaged on both sides of partisan divide in the US, the Russian-origin campaigns of late tended to stick to a particular side on any given issue in the West. More often than not, it's been the side that is less supportive of Ukraine.

These and other insights from our threat research help inform our efforts to protect public debate around the world, including ahead of elections since many deceptive tactics appear in Ukraine first.

New CIB networks disrupted in Q4'2023

- **China:** We removed a network in China that targeted US audiences. It posed as members of US military families and anti-war activists across multiple internet services including Medium, YouTube and petitions platform rootsaction[.]org. We found this activity as part of our internal investigation and removed it before this network was able to build an audience.
- **Myanmar:** We removed a network in Myanmar that targeted domestic audiences and posed as members of ethnic minorities across different platforms, including Telegram, X (formerly Twitter), YouTube and the network's own websites. We found this activity as part of our internal investigation, and linked it to individuals associated with the Myanmar military.
- **Ukraine:** We removed a network in Ukraine that targeted audiences in Ukraine and Kazakhstan. We found the full scope of this activity after reviewing information shared with us by our peers at Google.

Countering domain name abuse globally: Many threat actors continue to utilize domain name infrastructure in their malicious operations across the internet – from cyber espionage to covert influence campaigns and spyware firms. We recently resolved a legal case against Freenom, a country code domain registry provider, whose domain names accounted for over half of all phishing attacks involving country code top-level domains (ccTLDs). This settlement resulted in Freenom announcing that it will exit the domain name business, including its operation of the country-code registries. While Freenom winds down its domain name business, it has agreed to treat Meta as a trusted notifier and it will also implement a block list to address future phishing, DNS abuse, and cybersquatting.

01

Countering the surveillance-for-hire industry

This is our third annual report on the surveillance-for-hire industry and its indiscriminate targeting of people around the world.[1] We continue to investigate, study and take action against malicious violating activity linked to spyware vendors globally, including taking down their accounts on our apps and blocking malicious links from being shared on our platform.

It's important to note that we often have only a limited view into their activity as these firms target people across many internet surfaces at once, using each service differently to enable various stages of the surveillance attack chain.

As a reminder, we typically observe three phases of targeting activity by commercial spyware players that make up their "surveillance attack chain": **Reconnaissance, Engagement, and Exploitation**. Each phase informs the next and often they repeat in cycles. While some of these entities specialize in one particular stage of surveillance, others support the entire attack chain from start to finish. Although public debate often focuses on the exploitation phase, it's critical to disrupt the entire lifecycle of the attack because the earlier stages enable the later ones. If we can collectively tackle this threat earlier in the surveillance chain, it would help stop the harm before it gets to its final, most serious stage of compromising people's devices and accounts. More on how these phases work and what they entail is in our first report on this threat.

[1] See Meta's 2021 report and 2022 report on countering the surveillance-for-hire industry.

Trends & Notable Tactics, Techniques and Procedures (TTPs)

In this report, we're sharing our findings about six separate networks of accounts linked to eight firms from Italy, Spain and the United Arab Emirates. Here are some notable insights:

COMPLEX CORPORATE RELATIONS

We continue to see a complicated web of corporate structures among the entities engaged in these malicious activities, making attribution of abusive activities more challenging. Some of them acquire smaller vendors for a particular set of capabilities, likely to broaden the range of services they can offer and cover the entire surveillance attack chain. Some appear to engage additional firms with relevant expertise – like pen-testing or forensics – to facilitate surveillance operations, including to help validate whether their exploits are working as intended.

For example, we removed a network of accounts on Facebook and Instagram linked to **Italian surveillance-for-hire company Cy4Gate, which is associated with a larger defense contractor called ELT Group.** Our latest research found Cy4Gate using fake accounts with GAN profile photos, likely to scrape public information about its targets.[2] Further, Cy4Gate's surveillance-for-hire activity was previously reported by researchers at CitizenLab and Motherboard, including a now-inactive website that spoofed WhatsApp. There, Cy4Gate offered a trojanized version of WhatsApp for iOS. Our malware analysis of this exploit showed that its capabilities included collecting and accessing device information, location, photos and media, contacts, calendar, email, SMS, Telegram, Skype, Viber, Facebook, Instagram, LinkedIn, Signal, WhatsApp, and enabling microphone, camera and screenshot functionality.

We also took down a network of accounts on Facebook and Instagram linked to **an Italian spyware firm called RCS Labs owned by Cy4Gate.** This network consisted of multiple clusters of activity, including those operated from Italy, Kazakhstan and Mongolia. In addition to using fake accounts to test their malicious capabilities targeting Android and iOS devices, RCS Labs and its customers used tactics that included social engineering and phishing across the internet, including Facebook, Instagram, and LinkedIn. Their fictitious personas posed as protestors, journalists and young women to trick people into sharing their emails and phone numbers, as well as clicking on malicious links. These would typically be one-time-use, trackable urls – generated through freely available IP

[2] We've reported on the use of photos generated using artificial intelligence techniques like generative adversarial networks (GAN) by various threat actors since 2019, more details here.

tracking services – meant to identify and trace their targets' IP addresses. We also found RCS Labs and its customers embed canary tokens in Word documents that were disguised as news articles or anti-government petitions. This is also likely designed to track the targets' IPs and profile their devices, enabling the early stage of the surveillance attack chain – reconnaissance. Targeting by RCS Labs and its customers included journalists, activists and dissidents in Azerbaijan, Kazakhstan and Mongolia.

CUSTOMERS LEVERAGE MULTIPLE SUPPLIERS

By continuing to democratize access to spyware, this industry enables customers to use multiple malicious tools at once to complete the surveillance attack chain without relying on one vendor as a single point of failure. This makes it harder for any one threat research team to fully understand the activity we each see and identify who its ultimate beneficiary might be. It also makes it much more challenging for the targets to understand the fullest extent of surveillance aimed at them across the internet and hold those responsible to account, including in court. *See our call to action and recommendations for tackling this threat here.*

As an example, we removed a network of accounts on Facebook and Instagram linked to **an Italian surveillance-for-hire firm, IPS Intelligence**, that advertises its data collection and surveillance technologies**.** On our platform, this group's activity manifested primarily as scraping public information, using fake accounts with GAN-generated profile photos and targeting people in Italy, Tunisia, the US, Malta, Oman, Turkey, France, Zambia, Germany, and Mexico. These accounts were operated primarily from Italy and Tunisia. This activity appeared to target many services across the internet, including Facebook, Instagram, X (formerly Twitter), YouTube, Skype, GitHub, Reddit, Google, LinkedIn, Quora, Tumblr, VK, Flickr, TikTok, SnapChat, Gettr, Viber, Twitch and Telegram.

Notably, a customer of IPS Intelligence in Tunisia appeared to engage in separate phishing and social engineering activity, including the use of IP logging links to trace its targets' IP addresses, demonstrating how the surveillance-for-hire industry as a whole continues to supply different capabilities to the same clients to enable the full surveillance attack chain. Finally, our analysis showed this group using instrumentation and debugging tools for Android. While it's unclear what the ultimate goal might be, the presence of this tooling suggests an intentional tampering with the intended functionality of the services IPS Intelligence and its customers target.

TESTING EXPLOIT CAPABILITIES

Our investigations continued to uncover early activity by spyware firms where they try to test their capabilities to compromise their own accounts and devices, and then exfiltrate data from them.

Importantly, these findings allow us to identify rare early signals and understand where spyware firms might be investing in developing their capabilities so we can continue to harden our products' security and share our findings with industry peers to inform our collective defenses. This is particularly important for platforms like ours because when people's devices or browsers are compromised with malware and taken over without their knowledge, we may not always have enough signal at the app level to determine that it is no longer controlled by its rightful owner alone. *See our approach to reducing and hardening the attack surface here.*

For example, we took down a network of accounts on Facebook and Instagram linked to a **Spanish spyware firm called Variston IT, its Italian subsidiary and exploit developer TrueL IT, and a UAE-based firm called Protect Electronic Systems.** The exploit tooling by Variston was reported on by Google, as well as Variston's past targeting activity in the United Arab Emirates. On our platforms, they used fake accounts for exploit development and testing, including sharing of malicious links and placing calls between their own accounts in an apparent attempt to validate iOS and Android-targeting capabilities.

We also took down a network of accounts on Facebook and Instagram linked to **Italian spyware firm Negg Group.** It was previously reported by security researchers to have developed zero-day exploits targeting iOS and Android and deployed to target people in Italy and Malaysia across the internet. While we haven't seen targeting of authentic users on our platform, we observed Negg Group testing the delivery of its spyware – targeting iOS, Android and Windows – and its exfiltration capability between their own accounts.

In another example, we removed a network of accounts on Facebook and Instagram linked to **a Spanish firm, Mollitiam Industries, that advertises a data collection service and spyware targeting Windows, MacOS and Android.** Mollitiam Industries and its customers ran fake accounts which they used for testing malicious capabilities among their own accounts and scraping public information. Similar to other surveillance-for-hire firms, they used IP-logging links aimed at tracing their targets' IP addresses. They also engaged in phishing and social engineering targeted primarily at people in Spain, Colombia and Peru, including the political opposition, journalists, anti-corruption activists and activists against police abuse.

Post-Disruption Analysis

Our approach to threat disruption: We know that spyware firms are extremely motivated and unlikely to stop their cross-internet activity as a result of a single takedown by any one platform. That's why our work doesn't stop with the initial disruption. We take a defense-in-depth approach where we attack the problem from many angles at once – continuously, over time – so we keep degrading malicious capabilities while constantly improving our defenses against them. This includes:

- Recidivism tracking and enforcement using automated detection and expert investigations,
- Sharing insights and threat indicators with our industry peers and researchers,
- Public attribution,
- Public reporting to ensure that our threat research is accessible to platforms we may not have direct contact with,
- Legal actions,
- Sharing with governments and regulators to help inform stronger defenses,
- Alerting the people who we believe were targeted by spyware so they can take appropriate actions to protect themselves.

Threat reporting tensions: As we do this work, we have to balance a number of competing challenges. For example, there is an ongoing tension between the importance of exposing malicious activities and how our threat reporting may tip off the bad actors behind them. We continuously evaluate these trade-offs and weigh many factors, including whether exposure is likely to:

- Degrade technical malicious capabilities;
- Enable people who we believe were targeted by malicious groups to better protect themselves across the internet;
- Raise the cost of malicious targeting, including the time spent changing TTPs following our detection and exposure;
- Reduce our visibility into malicious TTPs due to adversarial adaptation and improved operational security by bad actors.

Expected adversarial adaptation: In response to disruptions, we expect the surveillance-for-hire industry to keep trying to adapt to stay afloat. And, as is typical for security threats, we monitor for

these changes and feed them back into our defenses to keep ahead. We do this through what we call post-disruption analysis, which helps us study and understand the impact of various security efforts on the overall behavior and TTPs of these malicious groups.

Conducting post-disruption analysis: While we have to be careful in sharing our insights to avoid tipping off spyware vendors, here are some examples of the spectrum of changes and attempted operational security (OpSec) improvements we've seen from the spyware firms we've been tracking and countering over the years:

- **Incremental changes in tactics:** As part of their adversarial adaptation, many malicious groups abandoned the reported tactics they had originally used, including GAN profile photos likely used to appear more authentic. And some switched to new tools to enable various operational activities like tracking phishing campaigns with marketing services.
- **Changes in malicious tooling:** Some of these groups claimed to have removed malicious capabilities like scraping of our apps from their product. Others patched apparently mis-configured product features and code that could give away their malicious activity, and some still struggled to get customers to update to the latest, patched version of their products.
- **Changes in infrastructure:** Post-disruption recidivism often involved attempts to move to new internet infrastructure, including shifting it to different countries. This may be particularly accessible for spyware firms that are part of larger multinational holdings.
- **Broader changes in operations:** Over time, likely in response to detection and regulatory pressures in some regions, some of these vendors ceased to operate under the same names and structure, and their employees shifted from one spyware firm to another. We've also seen some pause their attempts at malicious activity on our apps for periods of time.

Overall, our post-disruption analysis has shown that a comprehensive threat disruption approach can impose substantial friction on spyware groups and force them to expend more resources to hide and spread their activities across many services online to stay afloat. Our disruptions – when applied continuously and repeatedly over time – have led to a decrease in activity by these groups on our apps. More generally, broader actions against spyware providers by our industry and researchers have proven impactful in shining a light on this indiscriminate abuse and enabling transparency and accountability on behalf of the people targeted.

Hardening Product Security Against Spyware Threats

As part of our threat disruption efforts, our security engineering teams work to strengthen the security of our products, including based on the insights we gain through threat research into the surveillance-for-hire industry. In addition to proactively discovering and fixing these vulnerabilities in our apps, we also invest in reducing attack surface and exploit mitigation technologies. These technologies make exploitation of zero-day vulnerabilities difficult by eliminating or reducing the reliability of common techniques used by exploit developers. We prioritized these changes based on our understanding of exploitation techniques used to target zero-day vulnerabilities in similar applications, the results of our internal exploitation exercises, and threat research related to targeting of our apps.

Here are some updates on these efforts aimed to protect people using our apps.

MESSENGER: CONTROL FLOW INTEGRITY

We recently enabled Control Flow Integrity (CFI) on Messenger for Android. CFI is a compiler-provided exploit mitigation that applies to all native (C and C++) code in the application. It works by introducing constraints around forward edge control flow transfers, when a function pointer is used, by ensuring the destination address is valid. When an invalid destination address is discovered, the app will terminate. This makes common exploitation techniques like 'Return Oriented Programming' (ROP) harder to achieve.

Messenger users don't need to take any action to benefit from CFI other than updating their app. It is enabled for all Messenger Android builds starting with version v434.

WHATSAPP: VOIP MEMORY ISOLATION

We recently enabled a new memory allocator for the WhatsApp VoIP calling library. The allocator design includes numerous security properties that make exploitation of common memory safety vulnerabilities more difficult and less reliable. The allocator provides strong spatial isolation of memory allocations at random locations, strict checking of allocator data structures, sanitization of memory, and a more aggressive approach to page unmapping. This spatial separation of memory allocations breaks some common exploitation techniques for certain vulnerability classes.

WhatsApp users don't need to take any action in order to benefit from this change other than updating their app. It is enabled for all WhatsApp builds for Android and iOS starting with version 2.23.21 and Windows 2.2345.6.

WHATSAPP: SILENCE UNKNOWN CALLERS

In 2023, WhatsApp enabled a feature known as 'Silence Unknown Callers'. It not only stops annoying spam calls but also reduces zero-click attack surface that threat actors often seek to exploit in various applications. The feature works by introducing a cryptographically verified and privacy-preserving mechanism that verifies for the server whether to allow or reject the call. This protects the client from processing attacker-controlled data via WhatsApp calling from callers it does not trust.

In order to benefit from this change WhatsApp users need to update the app and enable the feature in their settings. It is available for all WhatsApp builds for Android starting with version 2.23.10.18, and iOS after version 2.23.10.70.

***

As this report and security research by our peers have shown, the surveillance-for-hire industry often tries to leverage zero-day security vulnerabilities as a mechanism for installing spyware on the devices of people they target. We will continue hardening our products and services against this malicious activity. We also welcome feedback and technical research from the security community. In 2023, we increased the maximum payout for remote code execution vulnerabilities in our apps to $300,000 USD.

Call to Action: Recommendations for countering the surveillance-for-hire industry

In 2022, we released detailed recommendations for governments and industry for countering abuse and indiscriminate targeting of people around the world by the surveillance-for-hire industry. Since then, we've seen a number of significant developments from across society to combat spyware, including the US government's landmark Executive Order restricting the procurement, testing, and deployment of commercial spyware, the tech industry's joint recommendations for regulating the spyware industry, and the multistakeholder Paris Call for action to protect people from abusive spyware.

Despite increased scrutiny, the commercial spyware industry continues to grow and, as our investigations show, to indiscriminately target people around the world, including journalists, dissidents, human rights defenders and activists.

With such wide targeting across the internet, there are many stakeholders who have the responsibility and opportunity to help tackle potential human rights violations and abuse stemming from this industry. It includes tech platforms and services like hosting providers and domain registrars; investors; local, state and national regulators; and civil society including security researchers and investigative journalists. Building on our earlier recommendations, a large body of investigative work and our post-disruption analysis since 2021, we're sharing our latest recommendations on what an effective whole-of-society strategy could look like across public and private sectors and civil society.

TECHNOLOGY SECTOR

Threat Disruption model: Meta developed the Threat Disruption Model to address a range of adversarial online threats, combining a comprehensive set of levers aimed to degrade the ability of malicious groups to target and abuse people using our services. It includes: expert investigations, takedowns of accounts and blocking of infrastructure used by spyware firms, evidence-based public reporting and attribution of these cross-internet abusive operations, information sharing with researchers and other online services targeted by these actors, continuous improvements of the security of our products and legal action against violating activities by spyware vendors.

Blocking abusive activity: Industry efforts to counter commercial spyware begin with ensuring that each of our security teams are hunting for surveillanceware targeting our respective services in all

phases of the surveillance attack chain: reconnaissance, engagement, and exploitation. These investigations and the takedowns that follow should be designed to maximally degrade malicious capabilities by blocking as much of the enabling infrastructure as possible at once and by sharing critical threat research, including malware signatures, technical analysis of malware capabilities and attribution, with others in the defender community and publicly, where appropriate.

Alerting people targeted: Various industry security teams may also have a unique visibility into who might be the people targeted by spyware across the internet. They should endeavor to provide appropriate alerts to these individuals to both enable them to secure their online presence and to protect themselves offline from potential risks.

Post-disruption analysis and iteration: While we know that commercial spyware providers are inherently adversarial and they actively adapt to detection and enforcement, interventions by different industry security teams significantly vary. Collectively, we have limited insight into the impact of each intervention in deterring abusive spyware activity that tend to target many online surfaces at once to enable surveillance operations. To standardize our approaches and efficiently innovate, industry teams should conduct post-disruption analysis designed to evaluate each intervention and its effect on malicious behavior and share their findings with industry partners and the public, where appropriate.

Hardening product security: Finally, industry security engineering teams should partner with threat intelligence teams to identify exploit development of vulnerabilities in their products – including open source libraries – that may enable spyware to target people. In doing so, they should invest in building tools and capabilities designed to narrow and harden the attack surface and provide potential targets with security tools to better protect themselves online.

FINANCIAL SECTOR

Private surveillance companies, like any business, often rely on outside support, including investment, to develop their capabilities. Investors therefore have the responsibility - under the UN Guiding Principles on Business and Human Rights - to ensure that their funding does not enable abusive and indiscriminate targeting of people, including dissidents, journalists, and democracy activists, or enable human rights abuses.

Investors, private equity firms, and venture capitalists should consider developing their own responsible investment guidelines, including human rights due diligence requirements for their portfolio companies. Such guidelines should consider human rights due diligence prior to investment regarding the companies' human rights policies and procedures, including how the

surveillance-for-hire firms intend to protect the privacy of people's data, the types of customers they intend to sell their services to, and how they plan to implement audit or risk assessment systems and accountability measures for their customers' compliance with human rights standards. Investors, private equity firms, and venture capitalists should consider publishing insights emanating from such due diligence in order to improve transparency across the commercial spyware industry and protect investors from the risks associated with abusive spyware companies.

REGULATORY LEVERS

Our investigations shared in this threat report identified several surveillance-for-hire firms operating in the European Union that continue to sell commercial spyware services to customers targeting people around the world, including in the EU itself. While the EU has one of the world's most comprehensive data privacy regimes, mercenary spyware companies, which access, store, and transmit EU user data, largely disregard EU privacy standards.

Employing data protection standards: EU Data Protection Authorities have a unique opportunity to protect the privacy of EU citizens by requiring commercial spyware companies based in or operating in Europe to comply with existing data protection regulations including the requirement to use an appropriate lawful basis before collecting, storing, or transmitting their data; storing EU user data in legally compliant ways; and providing EU citizens with the ability to request the deletion of their data. The EU should also explore mechanisms for limiting the amount and type of data accessed and how long it is retained by spyware companies and their clients; the UN Special Rapporteur for Human Rights and Counter-Terrorism has offered useful bounds for data access in the deployment of surveillance technologies.

Requiring human rights due diligence: As a leader in the global human rights space, the EU has an opportunity to set a standard and help curb malicious use of surveillance technology. In line with European Parliament's PEGA Committee recommendations, the EU should require commercial spyware companies based or operating in the EU to conduct human rights due diligence, in accordance with the standards of the UN Guiding Principles on Business and Human Rights, and comply with existing EU risk assessment mechanisms.

Requiring transparency into spyware customers: Commercial spyware companies sell sophisticated cyber capabilities, but one of their most valuable products is the layer of secrecy and deniability they offer to their customers. This deniability enables egregious human rights abuses. It also complicates government and civil society efforts to identify and hold the people behind targeting accountable, making it particularly difficult for targets to seek legal remedy for the abuse

of their privacy and security. Globally, governments should extend our 2022 recommendations and require commercial spyware companies to retain information about their customers and audit and log the targeting that customers conduct using their products. EU privacy regulations again provide a useful example, and the EU should continue to lead by creating a legal mechanism for targeted individuals to obtain information about the end customers targeting them and the companies providing the technology to do so.

Enabling legal remediation for spyware targets: Targets of abusive spyware often face significant difficulty in seeking remediation through the courts. As noted by the PEGA Committee recommendations, the EU should lead by developing clear standards for pursuing legal cases against spyware companies and their clients. They should also provide support for people who were targeted to pursue potentially costly and challenging legal action against spyware companies and their clients, including by providing them with technical expertise like forensic analysis to strengthen their cases.

Procurement standards for government entities: The US government laid out clear guidelines for US departments and agencies seeking to procure, test, and deploy commercial spyware capabilities in the landmark Executive Order in March 2023. EU member states should consider publishing similar guidelines and auditing standards for procurement, testing, and deployment of spyware capabilities in their own countries.

***

We believe that our latest recommendations for countering the surveillance-for-hire industry are applicable to a wide range of defender teams – from tech companies and the financial sector to governments and civil society. Our hope is to see them serve as a force multiplier in raising our collective defenses against the abuse by spyware.

02

Countering coordinated inauthentic behavior

We view CIB as coordinated efforts to manipulate public debate for a strategic goal, in which fake accounts are central to the operation. In each case, people coordinate with one another and use fake accounts to mislead others about who they are and what they are doing. When we investigate and remove these operations, we focus on behavior rather than content — no matter who's behind them, what they post or whether they're foreign or domestic.

Continuous CIB enforcement: We monitor for efforts to come back by networks we previously took down. Using both automated and manual detection, we continuously remove accounts and Pages connected to networks we took down in the past.

We included threat indicators for several CIB networks in the Appendix to help inform further cross-internet threat research by our industry peers and security researchers.

China

We removed 33 Facebook accounts, six Pages, six Groups, and four Instagram accounts for violating our policy against coordinated inauthentic behavior. This network originated in China and targeted US audiences.

The individuals behind this activity used fake accounts to post content, manage Pages and Groups, and pose as members of US military families and anti-war activists. Some of these Pages and Groups were focused on military themes, particularly US aircraft carriers. We removed this network before it was able to build an audience among authentic communities.

This operation posted links to news articles, commentary and memes in English, including apparent copy-pasted content from elsewhere on the internet, about the United States military, and criticism of US foreign policy towards Taiwan and Israel and its support of Ukraine. It was also active on YouTube, Medium, and petitions platform rootsaction[.]org, where it ran a petition claiming to have been written by Americans to criticize US support for Taiwan. As of February 5, the petition had attracted just over 300 signatures.

We found this network as a result of our internal investigation into suspected coordinated inauthentic behavior in the region. Our investigation found some links between this activity and the unattributed network from China we disrupted in September 2022.

- *Presence on Facebook and Instagram:* 33 Facebook accounts, six Pages, six Groups, and four Instagram accounts.
- *Followers:* About 700 accounts followed one or more of these Pages, about 2,300 accounts followed one or more of these Groups, and no accounts followed these Instagram accounts.

Myanmar

We removed 381 Facebook accounts, 88 Pages and 19 Groups for violating our policy against coordinated inauthentic behavior. This network originated in Myanmar and targeted domestic audiences in that country.

The individuals behind this activity used fake accounts to pose as members of ethnic minorities, manage Pages and Groups, comment on other people's content, and post long-form articles. Some of the network's Pages posed as fictitious news brands. These were active on multiple internet services, including Telegram, YouTube, X (formerly Twitter) and Viber.

The people behind this activity posted primarily in Burmese, and to a small extent in Kachin, Rohingya, Rakhine and Chin, about news, conflict, and ethnic armed groups in Myanmar. They shared original articles that praised the Burmese army and criticized the ethnic armed organizations and minority groups.

We found this activity as part of our internal investigation into suspected coordinated inauthentic behavior in the region. Although the people behind it attempted to conceal their identities and coordination, our investigation found links to individuals associated with the Myanmar military.

- *Presence on Facebook and Instagram: 381 Facebook accounts, 88 Pages and 19 Groups.*
- *Followers: About 149,000 accounts followed one or more of these Pages, and about 36,000 accounts joined one or more of these Groups.*
- *Advertising: About $60 in spending for ads on Facebook, paid for mostly in US dollars and Thai baht.*

Ukraine

We removed 1,020 Facebook accounts, five Pages, two Groups and 711 Instagram accounts for violating our policy against coordinated inauthentic behavior. This network originated in Ukraine and targeted audiences in Ukraine and Kazakhstan.

The people behind this activity used fake accounts to manage Pages and Groups, post long-form content and comment on other people's posts to make them appear more popular than they were. Some of these accounts used profile photos of real people copied from elsewhere on the internet, and regularly updated them with new photos to appear authentic.

The people behind this activity posted primarily in Russian about political events in Ukraine and Kazakhstan. In Ukraine, they posted supportive content about Viktor Razvadovskyi, a politician in Ukraine. In Kazakhstan, they commented on posts by news media like Radio Azattyq (the Kazakh service of Radio Liberty), and posted supportive commentary about the current government and critical commentary about the opposition.

We found the full scope of this activity after reviewing information shared with us by our peers at Google.

- *Presence on Facebook and Instagram:* 1,020 Facebook accounts, five Pages, two Groups and 711 Instagram accounts
- *Followers:* About 51,000 accounts followed one or more of these Pages, about 1,300 accounts joined one or more of these Groups and about 5,300 accounts followed one or more of these Instagram accounts.
- *Advertising:* About $30,000 in spending for ads on Facebook, paid for mostly in US dollars.

03

Analysis: Lessons learned from Russian influence operations related to war in Ukraine

This February marks two years since Russia's full-scale invasion of Ukraine and ten years since Russia's annexation of Crimea. Our teams remain on high alert to monitor for adversarial changes related to this war given that Ukraine remains the largest target of Russian-origin influence operations, with campaigns dating back to before 2014 and many deceptive tactics used elsewhere appearing in Ukraine first.[3]

Notably, even in cases where Ukraine was not a direct geographic target of the Russia-based influence operations, many of them still pushed narratives criticizing Ukraine onto audiences in other regions, including Europe, the Middle East and the US. This trend accelerated in the last two years since Russia began its full-scale invasion.

For the past two years, we've used different enforcements against influence operations – both covert and overt - and we continue to monitor for any changes that might help inform our efforts to protect public debate around the world, including ahead of major elections this year. Here is an update to our analysis from February of last year.

Overt influence operations

Two years ago, we took unprecedented steps to limit the spread of Russian state controlled media. The most recent research shows that these measures continue to result in sustained lower levels of activity and engagement globally. There are a number of takeaways worth highlighting to help the defender community assess and improve the effectiveness of different enforcement regimes.

Over the past two years, our enforcement measures included blocking Russian state controlled media from running ads, placing their content lower in people's feeds, and adding in-product nudges that ask people to confirm they want to share or navigate to content from these outlets. As

[3] Since 2017, we've disrupted twice as many CIB networks originating from Russia that targeted audiences in Ukraine or targeted international audiences with narratives about Ukraine, as we have domestic-focused, Ukraine-based CIB campaigns.

a reminder, we took these measures globally – across all languages – in addition to the usual transparency labels we apply to state media. Even outside of crises, we believe that people should know if the news they read is coming from a state-controlled publication. More on our approach to state-controlled media can be found here.

As part of our response to the war, we – alongside our industry peers – also complied with government requests from the European Union, United Kingdom, and Ukraine to geo-block some Russian state media outlets to restrict the ability to view these outlets' posts in these countries.

Additionally, in March 2022, the Russian government attempted to block or restrict access to Facebook and Instagram, as part of a wider attempt to cut Russian citizens off from the open internet, silence people and independent media, and manipulate public opinion.

Decrease in engagement and posting globally

In the first year, research showed that engagement with Russian state media dropped sharply: by August 24, 2022, posting volumes by these entities were down over 40% and engagement had fallen 80% compared to their levels the year prior. This trend persisted throughout the first year of the war.

Two years in, levels of activity and engagement have continued to decline globally, according to Graphika's latest research. It shows posting volumes went down 55% and engagement levels were down 94% compared to pre-war levels, while "more than half of all Russian state media assets had stopped posting altogether."

The research also shows that engagement levels fell significantly across all top target languages of the Russian state-controlled media outlets, including Russian, Spanish, English and Arabic, regardless of whether additional government restrictions were in place in particular regions.

This suggests that our global actions are effective in containing the spread of and engagement with content from Russian state media worldwide during the ongoing invasion, and in restricting the ability of posts from Russian state media to go viral, no matter the audience they aimed at. In fact, it appears that the softer social media enforcements yield sustained results in containing the reach of this content, while also providing transparency and context to people looking for these posts.

Because the potential for harm from misleading propaganda looms especially large in war times, our goal has been to limit the reach of Russian state media, particularly for people who might unintentionally encounter it in their feeds on our apps. However, we also believe that it's important

to preserve the ability for people who intend to find and view content from Russian state media to do so on our apps where it can be fact-checked and viewed alongside other counter-speech.

In that context, it's noteworthy that the decline in engagement was observed globally, and covered both regions and languages where governments had instituted geoblocks or bans, and regions where no such bans were mandated. This raises some important questions for further research about what the most effective approach may be to restricting state-controlled media online in times of war.

Addressing state media always involves striking an appropriate balance between ensuring people's right to access information and protecting against the Trust & Safety risks these outlets can pose. We will continue monitoring the situation, engage with experts, refine our enforcement measures and share our learnings with the broader defender community.

Covert influence operations

Everyone, everywhere, all at once

Ukraine has historically been targeted by a large number of different types of threat actors in Russia – from state entities to disinfo-for-hire groups.

Early CIB networks targeting Ukraine were often linked to various intelligence services. This includes operation "Secondary Infektion", which for the first time, we are now able to link to individuals associated with the Russian state.[4]

Since 2022, most CIB operations that we've detected were run by deniable entities such as marketing firms and troll farms, rather than state entities – notably the self-proclaimed "Cyber Front Z" (which was linked to individuals associated with past activity by the Internet Research Agency), and a pair of commercial companies.

These varied threat actors continued to add more and more online services to their campaigns – from Facebook, Twitter (now X), YouTube, LiveJournal, Blogspot, VK, OK, local blogs and forums, to websites and the comments sections of newspapers and TV stations, and later TikTok and Telegram.

[4] We first exposed Secondary Infektion in our threat reporting in May 2019. Following our sharing, open-source researchers, industry peers and various governments identified elements of this operation across more than 300 different social media platforms.

Throwing the kitchen sink

Russia-origin CIB networks and their varied operators have aimed an exceptionally wide range of techniques at Ukraine. Since 2014, these campaigns have targeted Ukraine with everything from fake "hacktivist" personas and impersonations of deceased journalists to forged leaks, coordinated mass commenting, and even a cardboard cutout of Russian soldiers for people to take selfies with. Given this much variety, it's not credible to define these TTPs as a single "Russian playbook".

As a notable change in tactics, since the 2022 invasion, we've found fewer attempts to build complex deceptive personas. Instead, we've seen more use of thinly-disguised, often short-lived fake accounts in an effort to spam the internet with inauthentic activity, apparently in the hope that something will "stick." We've taken down the two largest CIB networks, in terms of the number of accounts used, that we've ever seen from Russia. Both targeted Ukraine, and both prioritized posting volume over persona-building.

More noise, less impact

Despite this attempt at high-volume operations, we've seen a consistent decline in the overall followings of the Russian-origin CIB campaigns, even in cases with higher numbers of accounts on our apps. This stands in contrast with early CIB networks targeting Ukraine which had hundreds of thousands of followers.

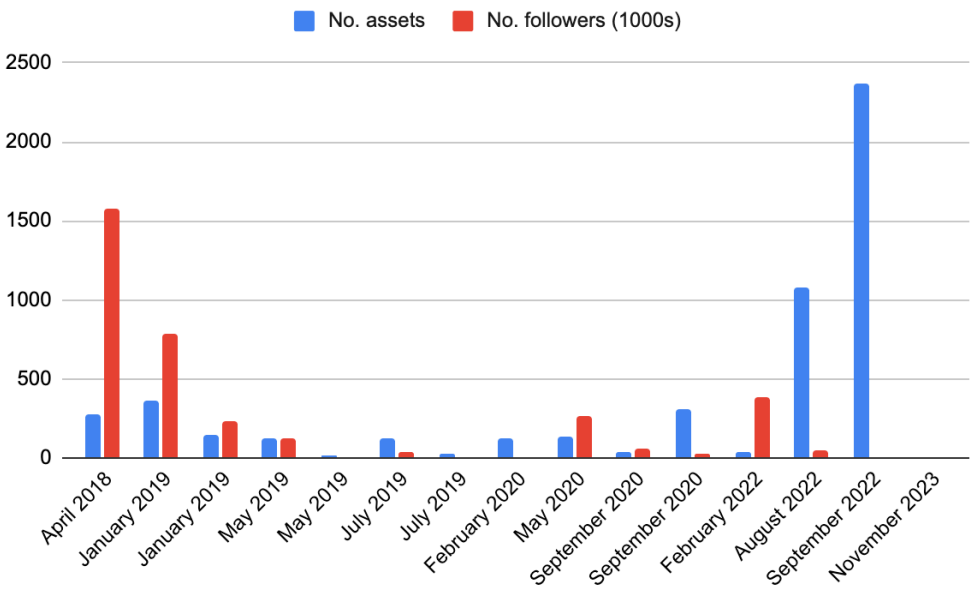

Image: Graph showing the number of accounts, Pages, Groups (blue) and thousands of followers (red) in Russia-origin CIB takedowns, 2017-23.

Our first takedown of a Russian-language network linked to the IRA had 1.5 million followers; the most recent had under 50,000. The use of spammy, fungible accounts may have been intended to compensate for this deficit by trying to achieve short-term exposure rather than building any trust or credibility. It can also be a reflection of the need to demonstrate volume as a metric of efficacy to the ultimate customers of these campaigns in a time of war.

The Russian operations targeting Ukraine are particularly aggressive and persistent, and we expect them to keep trying to find audiences across the internet.

Same hit, different day

In contrast to the wide range of tactics used by many different Russia-based campaigns over the past decade, some of the claims they made and narratives they promoted are remarkably consistent over time. This suggests that they may have been responding to a degree of centralized narrative guidance, particularly in moments of crisis.

For example, the day after an anti-aircraft missile supplied by Russia shot down Malaysian Airlines flight MH17 on July 17, 2014, Twitter accounts linked to the Internet Research Agency, English- and German-language social media accounts linked to Secondary Infektion, and a Russian-language "news" website linked to Russian military intelligence all called the shooting down a "provocation."

Other narratives surfaced repeatedly in different covert operations, in addition to state-controlled media, over the the last decade, including:

- **Bioweapons labs:** multiple Russia-origin networks accused the US of creating "bioweapons" in Ukraine, at least as early as 2014 and as recently as in 2023;
- **Ukraine arming terrorists:** several influence operations accused Ukraine of re-exporting Western-supplied arms to whichever terrorist group was making international headlines at the time (such as Islamic State or Hamas), at least as early as 2015 and as late as 2023;
- **Violence between Ukraine and neighboring nations:** several campaigns invoked historical violence between Poles and Ukrainians in an apparent attempt to stir up tensions between countries, as early as 2014 and as recently as 2023;
- **Western interference in Ukraine:** CIB campaigns accused Western countries of manipulating Ukraine as early as 2014 and as recently as 2022, with some even claiming that different countries (e.g., Poland and Germany) were planning to divide Ukraine's territory between themselves.

Such narratives have also been amplified by authentic voices over time, making narrative analysis an unreliable basis for identifying or attributing covert influence operations.

Narrative laundering

Throughout the last decade, we've observed attempts at "narrative laundering" where influence operations – both overt and covert – try to improve the credibility of their narratives by hiding their sources. Early CIB activity copy-pasted Russian state-controlled media articles without attribution, while state-controlled media and Russian embassies at times cited fake personas and fictitious new outlets as if they were authoritative, independent sources.

Crucially, these attempts at narrative laundering repeatedly targeted real influencers, politicians and media outlets in the West. Several early personas run by Russian military intelligence attempted to plant articles or leaks with Western news outlets. Operation "Secondary Infektion" reportedly emailed a set of hacked documents to UK-based politicians and activists ahead of the UK general election in 2019. A Russia-origin CIB network that focused on COVID-19 offered to pay social media influencers to push its content to their audiences.

As platforms make it harder for both overt and covert influence operations to gain new audiences, laundering narratives through independent voices is likely to become an even more appealing tactic.

Moving geographic targets

While so many Russia-origin influence operations targeted Ukraine, very few had Ukraine as their sole focus. Instead, they remained flexible and moved from one target country to the next, often in response to world events.

For example, some of the earliest CIB networks from Russia started by targeting opposition politicians inside Russia as early as in 2013. They then pivoted to targeting Ukraine in late 2013-early 2014. While many of these campaigns and their operators remained focused on Ukraine for the following decade, they also branched out to target other countries at different times, such as during Russia's military intervention in Syria, the doping-related Olympic ban on Russian athletes in 2018, US elections in 2016, 2018 and 2020, and the UK general election in 2019.

This agility is important to bear in mind as we look ahead to the many elections in 2024. For example, as we reported in November, "Doppelganger" – primarily focused on Ukraine and weakening its support by the West – has set up a series of websites themed around hot-button issues and elections in the US, Germany and France. Exposing such changes early is essential in

limiting their ability to build audiences across the internet and we'll continue to share threat indicators to help inform broader threat research and detection (see Appendix).

Lessons for 2024

As the large body of threat research shows, there are some key learnings that should inform our 2024 expectations related to the threat of influence operations that originate from Russia, including their targeting of Ukraine as we approach the two-year mark since the full-scale war began.

While Russia-origin influence operations remain active across the internet, and some are already showing a pivot towards election-related issues in the West, over the past decade they've increasingly struggled to build audiences and break through in the public discourse. While we expect spammy attempts at throwing large volumes of accounts and websites to continue, our threat research clearly shows that, historically, the main way that CIB networks get through to authentic communities is when they manage to co-opt real people - politicians, journalists or influencers - and tap into their audiences. As such, reputable opinion-makers represent an attractive target and should exercise caution before amplifying information from unverified sources.

Additionally, there's a difference between attempting an influence operation, and succeeding at it. Over-amplification of these unsuccessful attempts at influence can provide a boost for the people behind them to claim impact inside Russia so they can justify funding of these operations, regardless of their efficacy. Alternatively, when anyone who people disagree with is suspected of being a "Russian bot" – it erodes trust in the information environment and, again, plays into the hands of deceptive campaigns who are trying to create perception that they are everywhere.

Finally, unlike recent China-origin influence operations which engaged on both sides of partisan divide in the US, the Russian-origin campaigns of late tended to stick to a particular side on any given issue in the West. More often than not, it's been the side that is less supportive of Ukraine.

04

Update on our work against domain name abuse

As this report shows, many threat actors continue to utilize global domain name infrastructure in their malicious operations across the internet – from cyber espionage to covert influence campaigns and spyware firms. While we regularly block and publicize these malicious campaigns, they often continue to persist across the broader internet, including their websites and domains. This is because the mechanisms for redressing abusive domain names are not sufficient for the scale of the abuse that our industry and researchers see online today.

To help inform our industry's broader response to this abuse, we are sharing an update on our efforts to use domain name litigation against domain name registrars, registries, proxy providers, and others when they fail to cooperate or when they register imposter domain names, as prohibited by their agreements with ICANN, the organization charged with oversight of the domain name system.

In our August Q2'2023 Threat Report, we shared that we filed litigation against Freenom, a country code domain registry provider, whose domain names accounted for over half of all phishing attacks involving country code top-level domains (ccTLDs). Recently, we resolved this case through settlement and Freenom independently decided to exit the domain name business, including its operation of the country-code registries. While Freenom winds down its domain name business, it has agreed to treat Meta as a trusted notifier and it will also implement a block list to address future phishing, DNS abuse, and cybersquatting.

While this is a good example of this approach working, it is not easily scalable and is resource-intensive. This means that many targeted entities, including news outlets and civil society organizations, may not be able to pursue it. See our recommendations on how to improve the industry's efforts to tackle domain name abuse at scale.

As part of our continuous effort to help address the abuse of domain name systems by threat actors, next month, we'll be joining the US Chamber of Commerce at its Phishing Symposium to share our insights, learn from other industry peers and help inform our collective defenses against this threat.

Appendix: Threat indicators

The following section details unique threat indicators that we assess to be associated with the malicious networks we disrupted and described in this report. To help the broader research community to study and protect people across different internet services, we've collated and organized these indicators according to the Online Operations Kill Chain framework, which we use at Meta to analyze many sorts of malicious online operations, identify the earliest opportunities to disrupt them, and share information across investigative teams. The kill chain describes the sequence of steps that threat actors go through to establish a presence across the internet, disguise their operations, engage with potential audiences, and respond to takedowns.

This section includes the latest threat indicators and is not meant to provide a full cross-internet, historic view into these operations. It's important to note that, in our assessment, the mere sharing of these operations' links or engaging with them by online users would be insufficient to attribute accounts to a given campaign without corroborating evidence.

CHINA-BASED CIB NETWORK

Tactic	Threat indicator
Acquiring assets	
Acquiring Facebook accounts	33 accounts
Acquiring Facebook Groups	6 Groups
Acquiring Facebook Pages	6 Pages
Acquiring Instagram accounts	4 accounts
Acquiring YouTube channels	https://www.youtube[.]com/@KristenPreece/
Acquiring accounts on online forums	https://medium[.]com/@katelynwilsonkwk

	https://medium[.]com/@kristenprce
	https://diy.rootsaction[.]org/petitions/tell-kevin-mccarthy-that-selling-weapons-to-taiwan-is-selling-war-to-us?just_launched=true
Disguising assets	
Posing as non-existent person	Posing as anti-war activists
	Posing as military family members
Backstopping	The network maintained some of its personas across multiple social media sites
Evading detection	
Copying authentic content	The network copied some of its content from authentic news outlets
Targeted engagement	
Acquiring followers for Facebook Pages	About 700 accounts followed one or more of these Pages
Acquiring followers for Facebook Groups	About 2,300 accounts followed one or more of these Groups
Acquiring Instagram followers	No accounts followed these Instagram accounts
Posting into specifically themed Groups	Some of these Pages and Groups were focused on military themes, particularly US aircraft carriers

MYANMAR-BASED CIB NETWORK

Tactic	Threat indicator
Acquiring assets	
Acquiring Facebook accounts	381 accounts
Acquiring Facebook Pages	88 Pages
Acquiring Facebook Groups	19 Groups
Acquiring domains to support influence operations	hminewai[.]com
	kothet[.]com
	banyunt[.]info
Acquiring YouTube channel	https://www.youtube[.]com/@KoThet-pv9fm
Acquiring X account	https://twitter[.]com/KoThet969
Acquiring Telegram channels	http://t[.]me/factcheckmm
	http://t[.]me/daily21news
	http://t[.]me/manawpyomay
	http://t[.]me/tgitharmm
	http://t[.]me/evidenceoftruth

	http://t[.]me/myitsone1990
	http://t[.]me/inntharlay
	http://t[.]me/dgf21news
Acquiring Viber channels	Some of the operation's brands were also active on Viber.
Disguising assets	
Posing as non-existent person	Posing as members of ethnic minorities (Kachin, Rohingya, Rakhine and Chin)
Backstopping	The network maintained some of its personas across multiple social media sites
Targeted engagement	
Advertising to promote content	About $60 in spending for ads on Facebook, paid for mostly in US dollars and Thai baht
Acquiring followers for Facebook Pages	About 149,000 accounts followed one or more of these Pages
Acquiring followers for Facebook Groups	About 36,000 accounts joined one or more of these Groups
Posting about individuals or institutions	They shared original articles that criticized the ethnic armed organizations and minority groups.
	They shared original articles that praised the Burmese army

UKRAINE-BASED CIB NETWORK

Tactic	Threat indicator
Acquiring assets	
Acquiring Facebook accounts	1,020 accounts
Acquiring Facebook Pages	5 Pages
Acquiring Facebook Groups	2 Groups
Acquiring Instagram accounts	711 accounts
Disguising assets	
Adopting visual disguise	Some of these accounts used profile photos of real people copied from elsewhere on the internet, and regularly updated them with new photos to appear authentic
Targeted engagement	
Advertising to promote content	About $30,000 in spending for ads on Facebook, paid for mostly in US dollars
Acquiring followers for Facebook Pages	About 51,000 accounts followed one or more of these Pages
Acquiring followers for Facebook Groups	About 1,300 accounts joined one or more of these Groups
Acquiring followers for Instagram accounts	About 5,300 accounts followed one or more of these Instagram accounts

Engaging with specific audience	In Kazakhstan, they commented on posts by news media like Radio Azattyq (the Kazakh service of Radio Liberty)
Posting about individuals or institutions	In Kazakhstan, they posted critical commentary about opposition voices
	In Kazakhstan, they posted supportive commentary about the current government
	In Ukraine, they posted supportive content about Viktor Razvadovskyi, a politician in Ukraine

Continuous enforcement: novel indicators from recidivist attempts

We monitor for, and enforce against, efforts to come back by networks we previously removed for CIB, cyber espionage and other policy violations. Some of these networks may attempt to create new off-platform entities, such as websites or social media accounts, as part of their recidivist activity.

We're sharing some of these novel threat indicators related to recidivism attempts to enable further research by the open-source community into any related activity across platforms. It's important to note that, in our assessment, the mere sharing of these operations' links or engaging with them by online users would be insufficient to attribute accounts to a given campaign without corroborating evidence.

SIDECOPY, PAKISTAN-BASED CYBER ESPIONAGE GROUP

As a reminder, cyber espionage actors typically target people across the internet to collect intelligence, manipulate them into revealing information, and compromise their devices and accounts. When we disrupt these operations, we take down their accounts, block their domains from being shared on our platform, and notify people who we believe were targeted by these malicious groups. We also share information with security researchers, governments, and our industry peers where appropriate so they too can take action to stop this activity.

In 2021, we shared our research into a threat actor in Pakistan known in the security community as SideCopy. We continued to monitor and take action against their recidivist attempts. As a result, we found and analyzed new, previously unreported malware families deployed by SideCopy across the internet. These two new families – Panther and Cyrus – target Android devices. We believe that both of these new strains are custom-developed by SideCopy or someone it's working closely with.

Panther Malware for Android: This malware family masquerades as chat apps for Android devices and includes capabilities to capture call logs, contacts, files, device information, SMS, and enable the microphone to record audio. It's also capable of running shell commands, while also checking for and retrieving files with a specific path and name if they exist on the device.

Cyrus malware for Android: This malware family also masquerades as chat apps for Android devices, including one called Conversations. It contains the full functionality of the chat application

in addition to malicious code. Cyrus is capable of collecting: call logs, contacts, device information, network information, operating system information, and Google account information.

Tactic	Threat indicator
Acquiring assets	
Acquiring domains hosting PJobRAT malware	toolkitapi[.]xyz
	itechcube[.]xyz
Acquiring domain hosting PJobRAT and Panther malware	connectsol[.]xyz
Acquiring domains hosting Cyrus malware	kuicksy[.]in
	fizzled[.]in
Acquiring domains to support malware campaigns	liberaltraveller[.]blogspot[.]com
	luxuriantexplorer[.]blogspot[.]com
	1cchat[.]blogspot[.]com
	sanga1l[.]blogspot[.]com
	lifestylepractices[.]blogspot[.]com
	seaviewsunsetbeauty[.]blogspot[.]com
	popularhistorywars[.]blogspot[.]com

	livingstyleimplementation[.]blogspot[.]com
	bharatnamaste[.]joinmyapp[.]xyz
	twinkle[.]joinmyapp[.]xyz
	adultchatroom[.]in
	chit-chat[.]joinmyapp[.]xyz
Developing custom malware for Android	**CryptChat (Panther) hash:** e63f7718c84c3b325c7da61a58dea7f8cb5b4eaa27b78eb7d78b4771400d3bc5
	Gorilla Jung (Panther) hash: 7d2b44266643ac92bc967dfdc8212b3709f66927c01f3d5c2dacff4e69fca42e
	Gorilla War (Panther) hash: c4b2bb4c34bef590d3af566c04af2f96ea4384c318b33e1f4f0e38c183444eda
	IndiModels (Panther) hash: 6c9fdb8dd485175d0d667c739fab4caf579a6bdc1b6d59dd899a9e64e5a437f9
	Jihad Pakistan (Panther) hash: fe40ff990b094b3be329dda8d90372cbb324c03424dfa8aedce715a66dd11960
	TChat (Panther) hash: 967b3818434585ed71adc693aa955b64c8855b3bf2c80bfae1a4db86bdc77978
	Fizzled (Cyrus) hash: 1eabae70651e4add259b4176a2a50de2c8ac9a526aaea2e0b2657082593ceff6

	Kuicksy (Cyrus) hash: d659be4ae2e65369ac6d5fc7e47d257f57f3057b6e335955593491aa1dcd6712
Disguising assets	
Disguising lures for Panther malware as blogs	historybooksandinfo[.]blogspot[.]com
	uniqueaccessoriesblog[.]blogspot[.]com
	dependablework[.]wordpress[.]com
	chitchatone[.]blogspot[.]com
	bookreadinginfo[.]blogspot[.]com
Disguising lures for PJobRAT malware as blogs	chitchate[.]wordpress[.]com
	talkto1[.]wordpress[.]com
	bluechatt[.]wordpress[.]com
Disguising lures for Mayhem malware as blogs	conversemore[.]wordpress[.]com
	ourhome234[.]blogspot[.]com
	f2reebooks[.]blogspot[.]com
Coordinating and planning	
Using domains and subdomains for command and control (C2) of	mrsuriyais[.]onthewifi[.]com

Panther malware family	
	whoseethe[.]ddnsking[.]com
Using domains and subdomains for command and control (C2) of Cyrus malware family	androidfirebase[.]in
	officetemplatecv[.]in
Using domains and subdomains for command and control (C2) of Mayhem malware family	hopat[.]webhop[.]me

YARA Rules

PANTHER

```
rule panther {

  meta:

    source = "Meta"

    description = "Rule for Android malware called panther"


  strings:

    /* Hard coded strings */

    $str_hc_1 = "get(MApp.CTX)"

    $str_hc_2 = "acArr.toString()"


    /* Heuristics around SMS retrieval */

    $str_sms_1 = "$isFromCmd"

    $str_sms_2 = "smm-"

    $str_sms_3 = "CTX.contentResolver"


    /* Heuristics around external storage file retrieval */

    $str_external_storage_1 = "fls-"

    $str_external_storage_2 = "bbthree"

    $str_external_storage_3 = "classStorageVolume.getMethod"


    /* Heurstics on contact retrieval */

    $str_contacts_1 = "cnt-"

    $str_contacts_2 = "numberCursor.getString(indexNumber)"

    $str_contacts_3 = "nameCursor.getString(idexDisplayName)"
```

```
    /* Heuristics on call log retrieval */
    $str_cl_1 = "clg-"
    $str_cl_2 = "cursor.getString(indexType)"
    $str_cl_3 = "cursor.getString(indexDuration)"

    /* Heuristics from helper functions */
    $str_hlper_1 = "DcFile(name="
    $str_hlper_2 = "cursor.getString(indexColName)"
    $str_hlper_3 = "decode(KratosConstants"
    $str_hlper_4 = "panther/buddy.php"
    // base64 of panther/buddy.php with different padding in front
    $str_hlper_4_b64_0 = "cGFudGhlci9idWRkeS5waHA="
    $str_hlper_4_b64_1 = "L3BhbnRoZXIvYnVkZHkucGhw"
    $str_hlper_4_b64_2 = "YW50aGVyL2J1ZGR5LnBocA=="
    $str_hlper_5 = "aHR0cDovL21yc3VyaXlhaXMub250aGV3aWZpLmNvbT" // subset of b64 encoded
C2

    /* Heuristics on audio recording */
    $str_hotmic_1 = "shll-"
    $str_hotmic_2 = "commandResult.getStderr()"
    $str_hotmic_3 = "rec aud perm denied"

    /* Alternate comms channel over socket */
    $str_socket_1 = "nn-k-f-s"
    $str_socket_2 = "ioSocket.id()"
```

```
    /* Location tracking */

    $str_loc_1 = "CREATE TABLE IF NOT EXISTS loco(id INTEGER PRIMARY KEY, uniquevalue
TEXT UNIQUE, lat TEXT, lng TEXT, alt TEXT"


  condition:

    uint32be(0) == 0x6465780a and (

        all of ($str_hc_*) or

        all of ($str_sms_*) or

        all of ($str_external_storage_*) or

        all of ($str_contacts_*) or

        all of ($str_cl_*) or

        2 of ($str_hlper_*) or

        all of ($str_hotmic_*) or

        all of ($str_socket_*) or

        $str_loc_1

    )

}
```

CYRUS

```
rule cyrus {

  meta:

    source = "Meta"

    description = "Rule for Android malware called cyrus"


  strings:

    $upload_url0 = "cyrus"

    $upload_url1 = "initc.php"

    $upload_url2 = "usload.php"


    $storage_file0 = "os.txt" fullword

    $storage_file1 = "rsm.txt" fullword

    $storage_file2 = "scl.txt" fullword

    $storage_file3 = "rcl.txt" fullword

    $storage_file4 = "rcn.txt" fullword


    $log_msg0 = "restrat"

    $log_msg1 = "hi back restarting!! :D"

    $log_msg2 = "success, path: "


    $class0 = "eu/siacs/conversations/services/GoogleService"

    $class1 = "eu/siacs/conversations/services/MiscService"

    $class2 = "eu/siacs/conversations/services/Service"

    $class3 = "eu/siacs/conversations/services/UpService"

  condition:
```

```
    uint32be(0) == 0x6465780a and (
      2 of ($upload_url*)
      and 3 of ($storage_file*)
      and 1 of ($log_msg*)
      and 1 of ($class*)
    )
}
```

DOPPELGANGER: BRANDS & SPOOFED DOMAINS

We first exposed an influence operation known as “Doppelganger” in September 2022. We continued to share new findings and threat indicators related to this ongoing, cross-internet campaign on GitHub, in a machine-readable format. This includes domains that spoof news and government websites, Doppelganger’s own brands, and hundreds more that Doppelganger uses to redirect people to its spoofed and branded websites..