# [Time - 0.54 mins] Introduction (Sruthi):

## Slide 2:

Good morning everyone. I'm Sruthi. Today, we are here to present about the threats social media end users today currently face.

#### Slide 3:

Here is the flow of the contents for this presentation.

## Slide 4:

(skip)

#### Slide 5:

Social media has become integral to our daily lives, and its ability to connect billions across the globe has made it a powerful tool for communication. In total, 64% of the world's population are now social media users.

#### Slide 6:

However, its popularity has also led to misuse, and actions such as phishing, identity theft, cyberstalking and catfishing have gained popularity over the recent years.

#### Slide 7:

We identified 3 main elements of social media to be :

- 1) The users and content creators
- 2) The platform features
- 3) the content itself

And susceptibility of social media platforms due to the three factors:

- 1) Ease of manipulation of consumers,
- 2) ease with which attackers can exploit platform features and
- 3) familiar nature of popular content sharing options.

We illustrate how each factor affects the strength of social media against attacks with an analysis of a case study each.

# [Time - 3.46 mins] Case Study 1 (Sruthi):

### Slide 8:

The first case study highlights the ease of manipulation and projection of content in social media to the creator's intent.

#### Slide 9:

Social media post popularity is often driven by the number of interactions like likes, comments, and shares, and thus demand attention-grabbing headlines to reinforce users' existing biases. This often gives rise to disinformation, which is the deliberate spread of false information to

deceive. In the following case studies, we'll use misinformation and disinformation interchangeably for simplicity.

# Slide 10:

The rapid spread of media across social media platforms can be attributed to the Guru - Follower Networks. The "guru" accounts are typically seen posting messages actively, which is forwarded without verification by large numbers of devoted users known as "followers".

#### Slide 11:

The consistency with which the information floods the public's feeds with minimal changes only to the headlines further reshapes the public's beliefs and constructs.

#### Slide 12:

Threat actors, well aware of how easily information spreads, exploit this by feeding manipulated images that reinforce their existing beliefs, thereby perpetuating the cycle.

# Slide 13:

The 'Ghost of Kyiv' legend illustrates how easily content can be manipulated to shape narratives. This rumored Ukrainian pilot was celebrated for allegedly downing 40 Russian planes, a story amplified by social media. However, experts doubt the claim, noting that such achievements are unrealistic.

# Slide 14:

When we run a simple reverse image on this viral picture of the supposed "Ghost of Kyiv", we discover that this image was actually posted by the Defence of Ukraine in 2019 and given a completely different narrative in 2022.

#### Slide 15:

Here, we perform noise analysis of this image. When images are tampered with, they often leave visible traces in the noise. In this case, this can be seen at the pilot's head, the insignia on his arm and the Ukrainian flag in the background as shown on the right.

#### Slide 16:

We can now put these two concepts together. Let's observe how easy manipulation of content on social media allows threat actors to exploit gaps present in the rapid nature of content sharing that is propelled by the guru - follower networks in the context of COVID 19.

# Slide 17:

This case study explores the COVID-19 vaccine disinformation, where anti-vaxxers leveraged social media to spread false narratives and gather public support. Their tactics included social engineering, fake ads, misleading links, and even deepfakes, aiming to undermine trust in vaccines.

### Slide 18:

To understand the anti-vaxxers' TTPs, we map anti-vaccine misinformation campaigns onto the cyber kill chain framework and show a systematic approach to influencing public opinion. The campaign begins by gathering information about vaccine opinions, then spreads manipulated content through social media and impersonated accounts to amplify misinformation. As it gains traction, the campaign leverages shares, posts, and rallies to reinforce anti-vaccine beliefs and undermine public confidence in vaccines.

#### Slide 19:

Let's now look at some popular conspiracy theories on the internet, starting with the Magnet Vaccine Conspiracy Theory. The image on the right shows the clusters of communities of accounts discussing the topic that COVID-19 vaccines contain metallic particles that can record mental activity and transmit it to a computer. The text in each cluster shows the accounts leading the conversations, and the size of each cluster reflects the volume of accounts fully on board with this claim.

## Slide 20:

Another conspiracy theory states that the COVID-19 vaccine introduced microchips into the human body. What's surprising is that one in five Americans reportedly believed this.

#### Slide 21:

This slide shows real twitter posts of people justifying their claims with pseudoscience.

# Slide 22:

The effortless spread of such disinformation, often with an underlying political agenda, contributes to cognitive hacking in the public. Attackers "hack" the victim's cognition and change their perception and beliefs to their content to gain support without further analysis.

# [Time: 2.5 mins] Case Study 2 (Anika):

# Slide 23:

Let's dive into the 2nd case study, highlighting how social media can be used to steal data through malware.

#### Slide 24:

CoralRaider is a threat actor group focused on stealing data using social engineering and living-off-the-land techniques

#### Slide 25:

CoralRaider's impact has a global reach, with targets worldwide.

#### Slide 26:

Here's an overview of their attack chain and how their malicious payload exploits the victim's system to exfiltrate data.

## Slide 27:

The malware is delivered via Facebook malvertisement campaigns aimed at personal and business accounts.

The left image offers individuals explicit content while the one on the right offers businesses Microsoft templates.

Clicking the link triggers a drive-by-download of malicious shortcut files.

#### Slide 28:

We found that the files were hosted on Google Drive, flagged suspicious by VirusTotal, though other tools provided no insights.

#### Slide 29:

The shortcut file is masqueraded as a common file, like PDF, through file extension spoofing. Once clicked, it downloads and executes a HTML application. An obfuscated script triggers PowerShell, performing tasks like bypassing user access control and running RotBot.

#### Slide 30:

An analysis of the hash values of the IOCs shows they are linked to the shortcut files and are flagged as malicious by VirusTotal.

#### Slide 31:

RotBot, a RAT tool running through PowerShell, disguises itself as a legitimate Windows process to evade detection, perform reconnaissance and download the C2 server connection and an information stealer called XClient Stealer.

# Slide 32

HTTP requests are sent to the Telegram C2 server with confidential data. Below shows the code used for this.

# Slide 33

This is the link analysis of the attack process I just explained.

## Slide 34

CoralRaider attacks have been attributed with high confidence to being Vietnam-based. With the IP address found from the C2 server, a Vietnam address and phone number are found in a Whols query.

#### Slide 35:

Vietnamese is also commonly used in attacker files.

# Slide 36:

This case study shows that both personal and business accounts face risks, with individuals vulnerable to data theft and businesses vulnerable to reputational and financial losses.

# Case Study 3:

# Slide 37:

The 3rd case study focuses on malicious content mediums, specifically how a twitter user tried to incorporate memes to act as a C&C server using image steganography.

## Slide 38:

Social media platforms promote content creation through images and short text. However, threat actors can exploit this by hiding malicious payload in seemingly benign images, which is a misuse of image steganography.

## Slide 39:

Some reasons Threat actors employ this include:

- 1. Image formats being less suspicious than executables.
- 2. Image being a common digital media format, ideal for hiding C&C channels in legitimate activity.
- 3. And it being able to bypass security and exploit image parsing vulnerabilities.

### Slide 40:

the main objective was identified to be data theft.

The memes acted as command-and-control for malware already on victims' machines.

# Slide 41:

Here is a screenshot of the account, which has since been taken down by twitter

# Slide 42:

This malware is actually a trojan which targets windows machines, distributed by phishing attacks. Once installed it is able perform various commands as u can see on the top right image. It will open the page of the twitter account and look for images using the html <image> tag. After which, it will parse the commands hidden in the metadata of images and execute them on the infected machine. Therefore, posting 'memes' on the Twitter account actually serves as a mechanism for the attacker to activate the trojan.

## Slide 43:

From the source code of the malware, we can see that it uses pastebin as a c&c server to exfiltrate data, however the ip is a local address, which means it might have just been a proof of concept.

The method which the attackers used is very unique, because none of the tweets can actually cause an infection, its just a means to command machines that are already infected Attacks using steganography or popular social media platforms for C&C are not new, but Combining both makes it harder for the defender to notice as traffic to and from twitter would seem normal. Also for defenders, it's very impractical to scan the large amount of data produced on social media everyday for malicious activity

## Conclusion:

Slide 44:

In conclusion,

## Slide 45:

Our presentation focused on three key elements of social media and how they can be exploited by threat actors, with Each case study focusing on one of the key factors.

We also highlighted how Social media could assist threat actors in different phases of the cyber kill chain.

Slide 46:

References

Slide 47:

That's all for our presentation. Thank you for listening

# Random q and a steg info :

- In most cases these Images will not immediately execute malicious activity after downloading. Most of the time these innocuous images are used as tools to aid attacks.
- The focus of case study 3 is to show how attackers make use of social media accounts as a medium to execute part of an attack.
- Don't have much information on how the commands were embedded into the images using steganography. Nor do we know who. (little information)
- **The relationship with end users** while not directly affecting the end user .....engaging with malicious meme content giving it more traction could allow for its longer presence. Affect trust of users. Fear mongering.
- Steganography gives **living of the land** core in certain contexts. Provides means and prevents being detected.
- Why dont we see much of steg attacks?
  - Most social media sites compress and strip metadata .. sometimes even change pixel values. Hence trying to use images to execute malicious code or assist is very difficult. HIgh effort ( must understand how compression in sites work in detail etc) little outcome.
  - Low throughput (information bits/total bitssent)
  - Effectiveness also depends on the format of the image. Lossless formats like png are better..but larger file size tradeoff. lossy formats jpeg are smaller in size but dataloss.
- Maybe could combine Image steganography with social media sites that support polyglot files (files that support reading in two file formats) ....

( source: me = if telegram allows larger file sizes and if telegram supports polyglot files then this could be exploited..idk)

# Stegprocess:-

(very basic explanation. We dont have enough info to know if this is what the threat actor used )

- Change pixel LSB . not much change in physical appearance of image.
- Then decode it by reading all the Isbs .
- Currently steganography is much more advanced than this. You have cost functions
  that tell you which are appropriate zones of the image to hide information and matrix
  embedding techniques] that allows you to hide a lot of information modifying very few
  pixels.
- What steg process can do?
  - Buffer overflow

Polyglot files- files that support reading in two file formats. Telegram allows for large sized files and many file types. If telegram allows for polyglot files then image steg can be used to exploit this

https://crypto.stackexchange.com/questions/87101/would-this-image-steganography-technique-be-effective-at-avoiding-detection

## QnA (from ChatGPT):

1. How do you differentiate between organic social media interactions and potential malicious activities in real-time? Are there any automated tools or methods you recommend for this?

#### Answer:

Differentiating between organic and malicious activity in real-time can be challenging due to the sheer volume of interactions on social media. Automated tools like \*\*machine learning algorithms\*\* (e.g., anomaly detection and clustering algorithms) are commonly used to analyze patterns of engagement. For example, monitoring the frequency and nature of interactions like likes, comments, shares, and follower growth can help detect sudden spikes in activity that are typically associated with bots or coordinated campaigns. Tools such as \*\*Social Media Analytics (SMA) platforms\*\* (e.g., \*\*Brandwatch\*\*, \*\*Hootsuite Insights\*\*) and \*\*threat intelligence platforms (TIPs)\*\* like \*\*Anomali\*\* can help detect patterns of malicious behavior. Additionally, sentiment analysis can be used to detect abnormal levels of negative sentiment or agitation, which could indicate coordinated disinformation campaigns.

# \*\*Follow-up:\*\*

Disinformation campaigns often involve a strategic mix of organic and inorganic engagement. Machine learning models trained on historical data can often identify coordinated activities by

spotting anomalies in engagement patterns—such as sudden bursts of activity or accounts with fake follower counts. Automated systems like \*\*Botometer\*\* can also be used to detect bot-driven interactions.

2. You mentioned that the "Guru-Follower Networks" amplify misinformation. From a CTI perspective, what are the best ways to identify and track these networks without violating privacy concerns?

#### Answer:

From a CTI perspective, \*\*network analysis\*\* tools (like \*\*Gephi\*\* or \*\*Maltego\*\*) are used to map and analyze social networks to identify patterns of influence. By analyzing the interactions between a central "guru" and their "followers," we can identify manipulation tactics, like \*\*amplification\*\* or \*\*astroturfing\*\* (fake grassroots support). These tools allow you to trace connections, identify fake accounts, and detect coordinated behavior patterns. While privacy concerns are crucial, ethical research often focuses on publicly available data (open-source intelligence or \*\*OSINT\*\*) to map the relationships between accounts and detect suspicious networks without violating privacy.

# Follow-up:

Detection of these networks doesn't necessarily involve private user data but focuses on metadata and interactions (e.g., frequency of retweets, cross-posting between accounts). Combining these tools with natural language processing (NLP) to detect repeated narratives can be highly effective in identifying misinformation campaigns.

### \*\*On Case Study 1 - Disinformation and Cognitive Hacking:\*\*

3. \*\*In the "Ghost of Kyiv" example, how can threat actors manipulate metadata or image noise analysis to make disinformation even harder to detect?\*\*

# \*\*Answer:\*\*

Threat actors can manipulate \*\*image metadata\*\* (EXIF data) and \*\*image noise\*\* to make the origins of the image harder to trace. For example, by altering the \*\*image compression\*\*, adding noise, or even adding false metadata (e.g., fake creation dates or authorship), they can make it harder to spot image manipulation. More sophisticated techniques include \*\*image splicing\*\* or using \*\*deep learning models\*\* to create realistic fake images that appear authentic on the surface, making them difficult to detect using traditional noise analysis or metadata tools.

4. \*\*You discuss cognitive hacking and disinformation spread via social media. What role do machine learning and NLP play in identifying the origin and trajectory of a misinformation campaign?\*\*

<sup>\*\*</sup>Answer:\*\*

<sup>\*\*</sup>Machine learning (ML)\*\* and \*\*natural language processing (NLP)\*\* are crucial for analyzing and tracking the spread of disinformation. ML can be trained on large datasets of known

misinformation to detect emerging patterns in language, content, or behavior indicative of disinformation. For example, \*\*text classifiers\*\* can categorize content as likely being fake news or biased. \*\*NLP techniques\*\* like sentiment analysis and entity recognition help in tracking how narratives evolve over time, which can be useful for identifying coordinated campaigns. ML can also track the trajectory of false narratives by analyzing engagement across multiple platforms and mapping the spread of the content using \*\*temporal analysis\*\* and \*\*social network analysis\*\*.

# \*\*Follow-up:\*\*

Tools like \*\*Topic Modeling\*\* (using Latent Dirichlet Allocation, or LDA) and \*\*Word2Vec\*\* can be used to identify the semantic similarity of content across different posts, allowing for the detection of emerging disinformation topics.

5. \*\*In the case of the anti-vaxxer disinformation campaigns, how do you see AI and automated tools helping to prevent the spread of similar false narratives in the future?\*\*

#### \*\*Answer:\*\*

Al can be used to \*\*automatically flag and filter misinformation\*\* based on patterns observed in known fake news. Platforms can employ \*\*Al-driven moderation tools\*\* that use NLP to automatically detect disinformation and block or flag it before it spreads widely. Al can also detect \*\*deepfakes\*\* and \*\*manipulated media\*\* by analyzing inconsistencies in video/audio patterns. Additionally, \*\*bot detection algorithms\*\* can spot inauthentic accounts spreading false narratives. Collaborative systems, where platforms share threat intelligence about misinformation sources, can help detect and respond faster. Al can also be used to build \*\*counter-narratives\*\* by identifying the key influencers driving the disinformation and targeting them with fact-checking content.

## \*\*Follow-up:\*\*

Technical safeguards might include \*\*machine learning-based content flagging systems\*\*, where posts are flagged if they meet certain criteria (e.g., sensationalist headlines or text that mimics known disinformation patterns), with human oversight for review.

### \*\*On Case Study 2 - Malware Delivery via Social Media:\*\*

6. \*\*With the CoralRaider group using Facebook for malvertising, how do you assess the risk of these techniques evolving to bypass traditional security measures, such as sandboxing or URL filtering?\*\*

#### \*\*Answer:\*\*

\*\*Malvertising\*\* techniques, like those used by CoralRaider, will likely evolve to bypass traditional defenses by utilizing \*\*living-off-the-land\*\* tactics, which minimize the need to download traditional executables. Malicious payloads may leverage legitimate services like \*\*Google Drive\*\* or cloud storage to deliver malware, which can evade detection. \*\*Fileless malware\*\* techniques, which execute directly in memory (such as PowerShell scripts), can

bypass traditional signature-based detection systems. Evolving to use \*\*multi-stage infection\*\* processes (i.e., payloads delivered through a chain of decoy actions) may also make detection harder.

# \*\*Follow-up:\*\*

Detecting these techniques involves monitoring \*\*unusual behavior\*\* on systems (e.g., anomalous network traffic to C2 servers), \*\*analyzing IOCs\*\* like suspicious file hashes or command-and-control patterns, and \*\*behavioral analysis\*\* using tools like \*\*EDR (Endpoint Detection and Response)\*\* systems. Sandboxing may need to evolve to simulate more complex attack chains.

7. \*\*Could the presence of malware in social media advertisements be detected through the analysis of network traffic patterns (e.g., beaconing to C2 servers) or are there other more effective detection mechanisms?\*\*

## \*\*Answer:\*\*

Yes, network traffic analysis is a key method for detecting malware in social media ads.

\*\*Beaconing\*\* behavior, where malware communicates with command-and-control (C2) servers, is often a telltale sign of infection. By monitoring \*\*anomalous traffic\*\* to suspicious or known malicious IP addresses, defenders can detect these behaviors. Other methods include \*\*DNS traffic analysis\*\* to look for unusual domains or traffic patterns and \*\*traffic analysis\*\* of social media interactions to detect unusual link patterns or API calls that might indicate malvertising.

# \*\*Follow-up:\*\*

Using \*\*intrusion detection systems (IDS)\*\* or \*\*intrusion prevention systems (IPS)\*\* that analyze DNS queries, HTTP requests, and other network traffic can alert defenders to suspicious C2 communications. Additionally, \*\*sandboxing\*\* suspicious links from ads in a controlled environment to observe their behavior can also be effective.

### \*\*On Case Study 3 - Image Steganography as a C2 Mechanism:\*\*

8. \*\*You discuss using memes and images as a C2 channel. From a CTI perspective, how could defenders detect this type of steganography without direct access to the malware or infected machine?\*\*

# \*\*Answer:\*\*

Detecting \*\*image steganography\*\* without access to the malware itself involves monitoring \*\*network traffic\*\* for suspicious patterns. By analyzing the traffic between the infected device and the C2 server, defenders can identify unusual requests, such as large numbers of image downloads or uploads, that might indicate the use of images as a C2 channel. Additionally, \*\*image analysis tools\*\* can be used to detect hidden data in images by identifying anomalies in the image format or content. For example, \*\*statistical analysis\*\* of pixel patterns can sometimes reveal the presence of hidden information.

\*\*Follow-up:\*\*

Combining \*\*network traffic monitoring\*\* with \*\*file integrity monitoring\*\* on endpoints to detect unusual image file sizes or altered metadata can also help detect steganographic activity.

9. \*\*Given that attackers are leveraging social media for C2 through image metadata, how do you see this impacting the future of C2 server detection techniques, particularly in non-traditional communication channels?\*\*

\*\*Answer:\*\*

This evolution of C2 mechanisms highlights the need for more advanced \*\*traffic analysis\*\* and \*\*anomaly detection\*\* systems. In non-traditional channels, defenders will need to focus on detecting \*\*covert communication patterns\*\* rather than simply looking for traditional C2 signals. For example, monitoring \*\*web traffic for unusual requests\*\* or \*\*social media scraping tools\*\* could help identify suspicious activity. There may also be an increasing reliance on \*\*behavioral analytics\*\* to detect malicious activity even when it's using legitimate communication channels (

like image posts).

\*\*Follow-up:\*\*

\*\*Digital forensics tools\*\* that analyze \*\*image metadata\*\* and \*\*file system timestamps\*\* may need to evolve to detect C2 communication in metadata, in addition to traditional methods of looking at communication protocols.

### \*\*On Social Media Platforms and CTI:\*\*

10. \*\*How do you think platforms like Twitter or Facebook should adapt their security policies to detect and mitigate covert C2 operations, such as the one described in your third case study?\*\*

\*\*Answer:\*\*

Platforms should focus on detecting and blocking covert \*\*command-and-control (C2)\*\* operations by implementing \*\*advanced machine learning algorithms\*\* that analyze communication patterns, metadata, and unusual content sharing. This could include the use of \*\*image recognition tools\*\* that identify steganography or hidden messages within images. Additionally, platforms should implement stronger monitoring of \*\*suspicious behavior\*\* such as large-scale sharing of the same images, unusual spikes in post frequency, or coordination between users that suggests a C2 operation.

\*\*Follow-up:\*\*

Platforms can also collaborate with cybersecurity researchers to share data on potential C2 channels or malicious campaigns. Privacy concerns can be balanced by ensuring that only publicly available data (e.g., images, posts) is analyzed in the detection process.

\_\_\_

---

### \*\*General Cyber Threat Intelligence Questions\*\*

1. \*\*How can threat actors leverage public social media data for reconnaissance, and what methods can defenders use to detect and mitigate this activity?\*\*

\*\*Answer:\*\* Threat actors use open-source intelligence (\*\*OSINT\*\*) to gather information on individuals or organizations from public social media data, including posts, location tags, job roles, and connections. This reconnaissance can help them craft spear-phishing emails or identify vulnerabilities.

Defenders can detect reconnaissance by monitoring unusual profile visits or connections, employing tools like \*\*SOCMINT\*\* (Social Media Intelligence) or using deception tactics such as honeypot profiles to detect malicious inquiries. Mitigation includes limiting public visibility through privacy settings and employee awareness training.

2. \*\*What role does Telegram play in facilitating cybercriminal activities in Southeast Asia, and how can threat analysts track and infiltrate these groups ethically?\*\*

\*\*Answer:\*\* Telegram is widely used for its end-to-end encryption and anonymity features, making it a hub for dark web-like activities. Analysts can track such groups by monitoring public channels and using threat intelligence platforms that integrate with Telegram.

Ethical infiltration involves focusing only on publicly available information and obtaining explicit permissions when monitoring private groups. Analysts can identify patterns of malicious activity (e.g., shared malicious links or leaked data dumps) to gather actionable intelligence.

3. \*\*How can defenders identify signs of ransomware negotiation strategies being discussed in dark web forums or Telegram groups?\*\*

\*\*Answer:\*\* Key signs include discussions on data decryption, specific organization names, and shared proof of compromise (e.g., file samples or screenshots of encrypted systems). Threat analysts can deploy \*\*dark web crawlers\*\* or subscribe to paid CTI feeds that track ransomware activity.

Defenders can also monitor known IOCs, such as email addresses, wallets, or IP addresses shared in forums, to cross-reference them with ongoing incidents in their environment.

---

### \*\*On Case Study 1: Disinformation and Manipulation of Public Opinion\*\*

4. \*\*What CTI indicators suggest that a disinformation campaign is being coordinated on social media?\*\*

\*\*Answer:\*\* Indicators include high-volume posting with repeated narratives, rapid sharing of content from newly created accounts, and accounts that amplify specific hashtags. Network analysis often reveals clusters of interactions among fake accounts.

Analysts can use tools like \*\*Maltego\*\* or \*\*Social Mapper\*\* to map relationships between accounts and analyze posting behavior to identify orchestrated activity.

5. \*\*How can analysts track the propagation of disinformation campaigns across multiple platforms?\*\*

\*\*Answer:\*\* Disinformation often spreads from one platform to another (e.g., from Telegram to Twitter). CTI analysts can use \*\*cross-platform analysis\*\* tools such as \*\*ShadowDragon\*\* or \*\*DataMiner\*\* to correlate hashtags, URLs, and narratives. By monitoring these trends, analysts can identify the source and evolution of the campaign.

---

### \*\*On Case Study 2: Malware Delivery via Social Media (e.g., Malvertising)\*\*

6. \*\*What are the early signs of malvertising campaigns targeting social media users?\*\*

\*\*Answer:\*\* Indicators include unusual redirects when clicking on ads, advertisements linked to newly registered domains, and metadata suggesting domain connections to known malicious IPs.

Threat analysts can leverage \*\*domain reputation tools\*\* (like VirusTotal) to check the legitimacy of advertised URLs and monitor ad networks for unusual trends in click-through rates from compromised accounts.

7. \*\*What CTI practices are effective in tracking groups that use malvertising techniques?\*\*

\*\*Answer:\*\* Analysts can monitor underground forums where threat actors discuss purchasing or deploying ad services. They can also analyze common TTPs (Tactics, Techniques, and Procedures) used in malvertising, such as typosquatting domains. Combining OSINT with IOCs from previous campaigns helps narrow down potential suspects.

---

### \*\*On Case Study 3: Image Steganography and Advanced C2 Mechanisms\*\*

- 8. \*\*What CTI indicators help identify the use of image steganography as a C2 mechanism?\*\*

  \*\*Answer:\*\* Indicators include unusual outbound image uploads, consistent file sizes that don't match typical image compression ratios, and repetitive downloads of similar images.

  CTI teams can analyze traffic logs for anomalous HTTP requests and use \*\*steganography detection tools\*\* like StegExpose to evaluate suspicious images.
- 9. \*\*How can threat analysts use OSINT to detect and attribute image-based C2 activity to a specific threat group?\*\*

\*\*Answer:\*\* Analysts can track the distribution of images across platforms and analyze associated metadata, such as upload timestamps and geotags. Attribution may involve identifying linked user accounts or recurring domains/IPs tied to known groups. \*\*Threat actor profiling\*\*, based on historical campaigns, further aids attribution.

---

### \*\*On Disinformation Trends in Southeast Asia\*\*

10. \*\*How can CTI analysts detect state-sponsored disinformation campaigns targeting Southeast Asian nations?\*\*

\*\*Answer:\*\* State-sponsored campaigns often involve professional content production, dissemination through official-sounding accounts, and narratives aligned with geopolitical interests. Analysts can monitor \*\*fake news patterns\*\* using NLP techniques for repeated phrasing and sentiment. Correlating these patterns with timing (e.g., during elections or crises) can reveal state sponsorship.

Platforms like \*\*Graphika\*\* and \*\*Recorded Future\*\* specialize in uncovering such campaigns.

11. \*\*What methods can be used to differentiate state-sponsored campaigns from those driven by hacktivists or independent groups?\*\*

\*\*Answer:\*\* State-sponsored campaigns typically exhibit high levels of coordination, funding, and access to resources like advanced botnets or professional media outlets. Hacktivists, in contrast, rely on grassroots methods and tend to operate on a smaller scale. Analysts can compare infrastructure (e.g., server locations, hosting services) and TTPs to make distinctions.

---

### \*\*On Ransomware and Cybercriminal Trends\*\*

- 12. \*\*How can CTI analysts link a ransomware group's activities to trends on the dark web?\*\*

  \*\*Answer:\*\* Analysts can track forums where ransomware affiliates advertise their services, such as RaaS (Ransomware-as-a-Service) offerings, and correlate them with ongoing attacks. Following cryptocurrency wallets associated with ransom payments is another effective method. Tools like \*\*Chainalysis\*\* for blockchain analysis and \*\*dark web monitoring\*\* tools (e.g., DarkOwl, IntSights) can help identify trends and attribute activity.
- 13. \*\*What CTI insights can be gained from monitoring dark web discussions around ransomware negotiation strategies?\*\*

\*\*Answer:\*\* Threat actors often discuss successful negotiation tactics, share price thresholds they are willing to accept, and post victim profiles. Monitoring these discussions provides insights into how ransom demands evolve and what sectors are being targeted.

\_\_\_