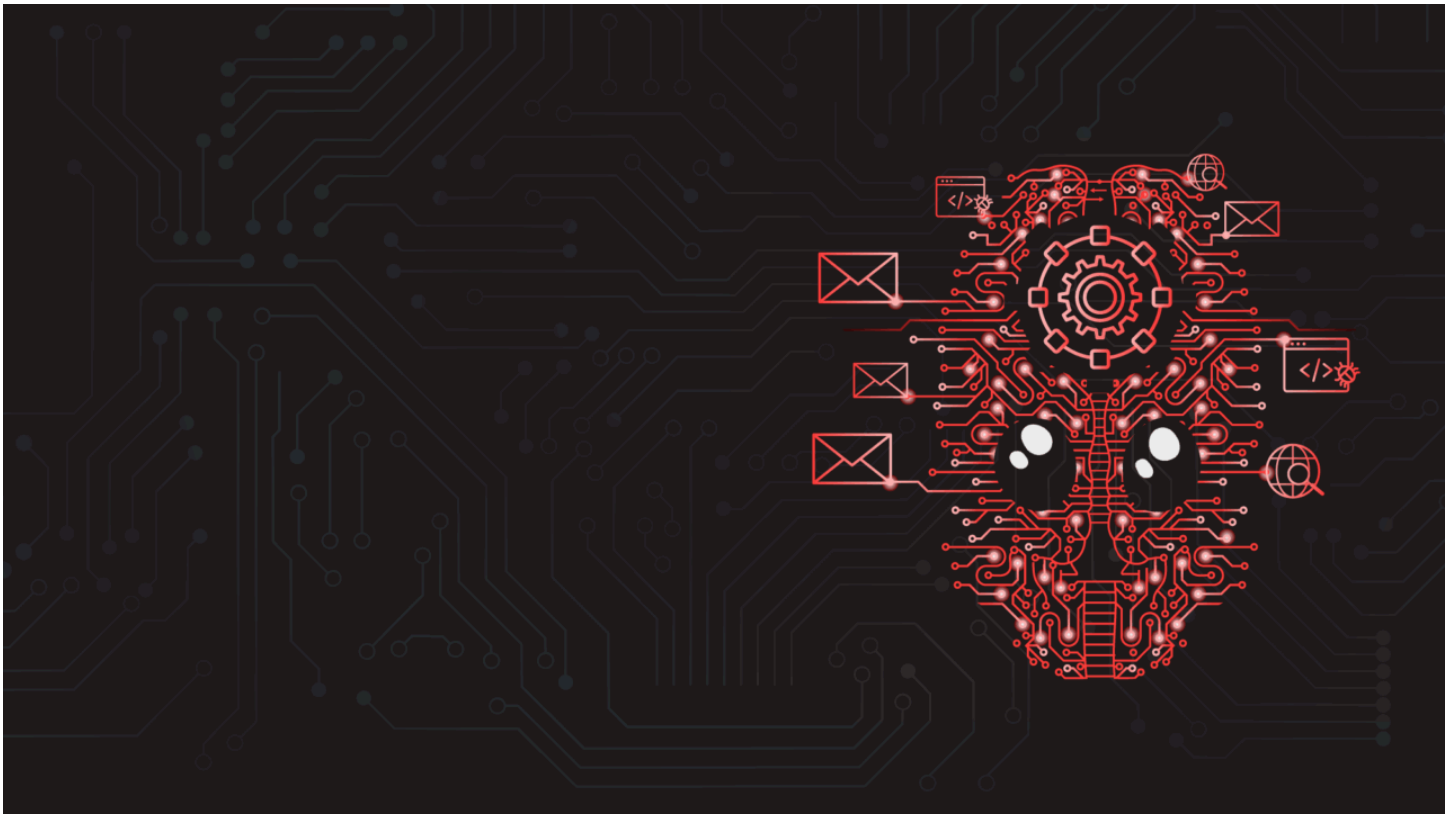
APT Meets GPT: Targeted Operations with Untamed LLMs

⋮ 10/8/2025



Starting in June 2025, Volexity detected a series of spear phishing campaigns targeting several customers and their users in North America, Asia, and Europe. The initially observed campaigns were tailored to the targets, and the messages purported to be sent by senior researchers and analysts from legitimate-sounding, completely fabricated organizations. The goal of these spear phishing campaigns was to socially engineer targets into clicking links that led to a remotely hosted archive containing a malicious payload. Volexity tracks the threat actor behind these campaigns under the alias **UTA0388** and assesses with a high degree of confidence that this is a China-aligned threat actor. This assessment is based both on technical artifacts and the targeting profile of the campaigns.

Over the course of three months, Volexity observed UTA0388 using various themes and fictional identities across dozens of spear phishing campaigns. As time passed, Volexity observed UTA0388 broaden their targeting and send emails in a variety of different languages, including English, Chinese, Japanese, French, and German. In most cases, the initial email sent by UTA0388 contained a link to phishing content hosted on a cloud-based service that would lead to malware. In a limited set of cases, Volexity observed UTA0388 hosting malware on their own servers. Once the initial and broader campaigns subsided, Volexity further observed multiple instances of highly tailored spear phishing against organizations where UTA0388 did not

send a link to malware, at first. Instead they engaged the target in conversation, and only after corresponding over the course of several emails would a malicious phishing link be sent. Volexity refers to this overall technique as "*rapport-building phishing*".

In all observed cases, UTA0388 sent a link leading to a ZIP or RAR archive file. Inside this file would be a legitimate executable that was given a filename relevant to the targeted organization or tied to the theme of the spear phish email.When executed, this legitimate executable would load a malicious payload in an included Dynamic Link Library (DLL), via search order hijacking which provided operators with the ability to remotely execute commands on infected devices.. Volexity tracks the deployed payload as **GOVERSHELL** and has observed five distinct variants of this malware family. Volexity assesses with high confidence that GOVERSHELL is used exclusively by UTA0388 and is still actively being developed at the time of writing.

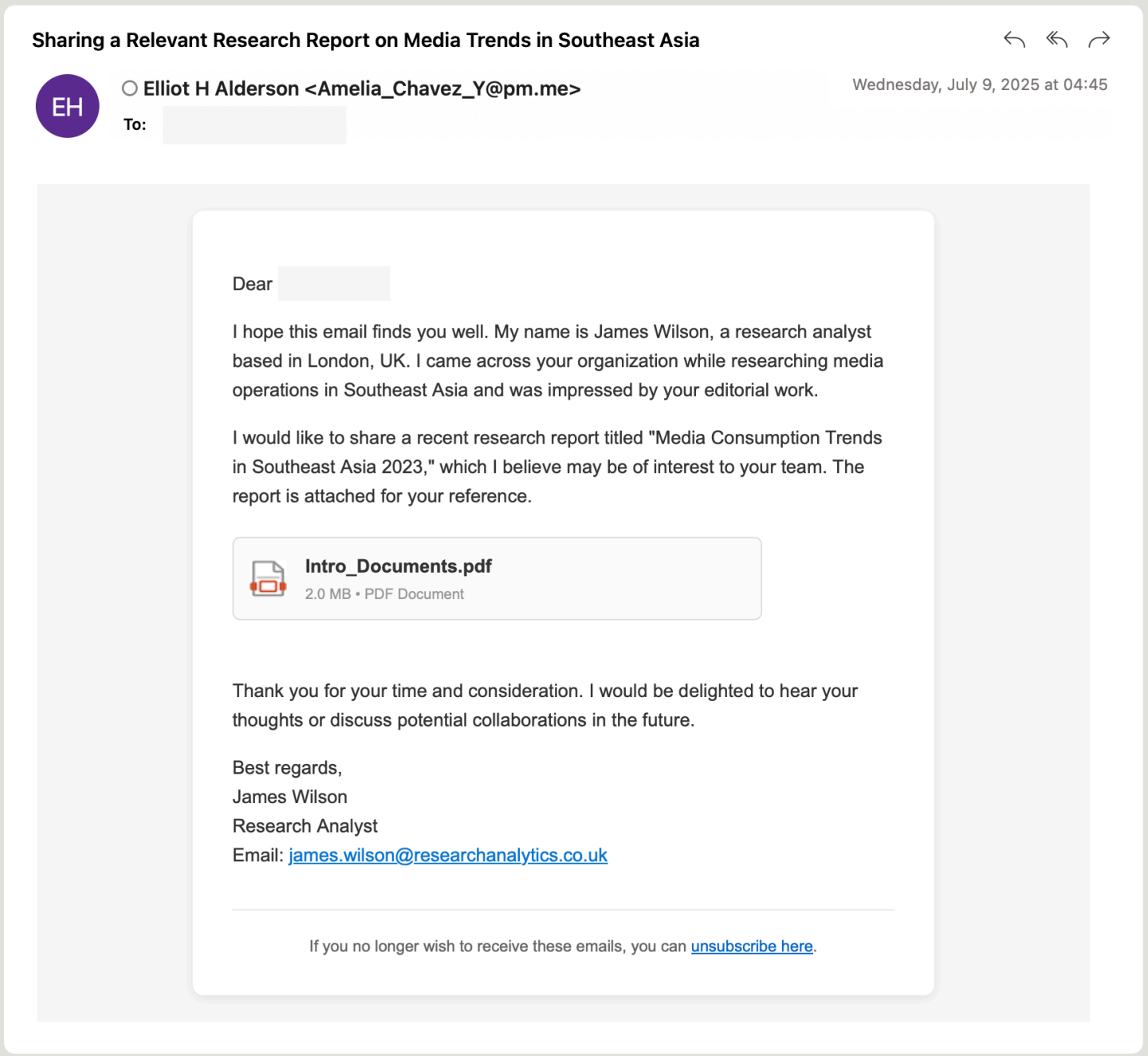
This blog post outlines technical details of various UTA0388 campaigns, and the evidence that led Volexity to assess with a high degree of confidence that UTA0388 employs Large Language Models (LLMs) to assist with their operations. Such an assessment can be difficult to credibly make, as no single data point can provide conclusive evidence of LLM usage. However, the aggregate evidence across UTA0388's campaigns, including nonsensical decisions made, the campaign tempo, and the sheer variety of campaigns supports this assessment. Since Volexity's initial research, and just prior to publishing this blog post, OpenAI published a report called *October 2025 Disrupting malicious uses of AI: an update*. That report confirms Volexity's suppositions that UTA0388 leveraged OpenAI's ChatGPT platform for several components of their spear phishing and malware development operations.

Volexity notes that UTA0388 is the same actor that Proofpoint tracks as "UNK_DropPitch", as described in a blog post published on July 16, 2025. That report described a malware family called "HealthKick", which is the earliest observed variant of what Volexity calls GOVERSHELL. Volexity also observed overlap in command-and-control (C2) infrastructure, and at least one sending email address from Proofpoint's reporting as well.

Technical Details

Spear Phishing Campaigns

Based on Volexity's visibility, UTA0388's primary and sole method for targeting organizations is by conducting spear phishing campaigns. Between June and August 2025, UTA0388 sent phishing emails containing HTML that included an image to make it appear a document was attached to the email. If the image were clicked, it led to the download of a remotely hosted archive file. Users would then need to open and execute the executable file within the archive in order to become infected. An example body from one such email is included below.



In this example, the email message body designed to look like a PDF was included with it. However, this was an image that was hyperlinked to the following URL:

https://aesthetic-donut-1af43s2.netlify[.]app/file/rar

Visiting the URL would result in a 302 redirect to a RAR file at the following URL:

https://aesthetic-donut-1af43s2.netlify[.]app/index/file/A_Introduction_Docs_v00546823.rar

Volexity observed several campaigns using different subdomains under the domain netlify.app, and the URL would vary between /file/rar and /file/zip. Each of the hostnames that UTA0388 used would support both endpoints and serve up both RAR and ZIP archives containing the same malicious files. UTA0388 would often use the same sending email address, but vary the “friendly name” and identity used in the actual email

body. In the above example, an immediate series of mismatches is obvious, with the left-hand side of the email containing “Amelia_Chavez_Y”, while the friendly name displayed “Elliot H Alderson” and the identity in the message body being “James Wilson”. These mismatches were consistently seen throughout the campaigns where phishing links were sent in the initial email.

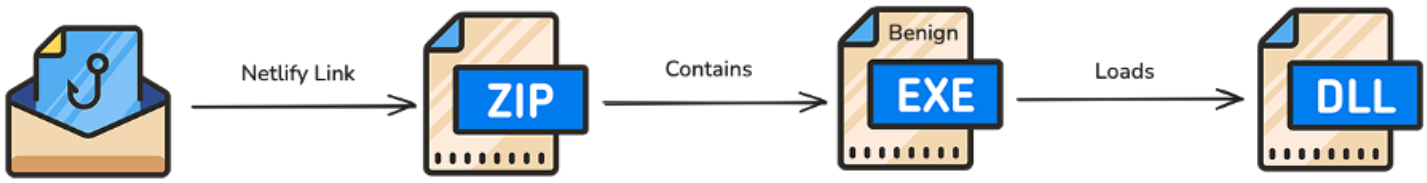
Beginning in August 2025, Volexity observed UTA0388 alter their approach, using rapport-building phishing. They would first contact targets with no malicious content, and only later deliver malicious content if the user replied. Based on Volexity’s visibility, this approach is increasingly popular amongst a variety of different threat actors, serving to only risk exposing the threat actor’s infrastructure and malware if a user has positively engaged in benign conversation first.

Phishing emails sent by UTA0388 that were observed by Volexity have all been sent from webmail providers that include ProtonMail, Outlook, and Gmail. Throughout June and July 2025, UTA0388 made use of Netlify to host their malicious RAR and ZIP archives, but they then diversified to use Sync, OneDrive, and their own domains not long after.

The remotely hosted archives are a mixture of RAR and ZIP formats containing at least one benign executable, a malicious Dynamic Link Library (DLL), and sometimes other superfluous files. The benign executable is always named to appear as though it were a legitimate document file. An example folder structure is shown below:

```
\Directory
|   2025 Important Documents and Materials.exe
|
└──lib (hidden directory)
        te64.dll
```

If the user runs the executable file, search order hijacking ensures the malicious DLL (\lib\te64.dll) placed alongside it is loaded. This would result in the end user’s system being compromised with a variant of UTA0388’s GOVERSHELL backdoor. This attack path is summarized in the image below.



GOVERSHELL

At the time of writing, Volexity has identified **five** distinct variants of the GOVERSHELL malware family. Throughout various campaigns Volexity observed active changes in the malware, with significant differences in how the malware communicated and functioned. All variants observed by Volexity make use of a *scheduled task* for persistence and provide the operator the ability to execute arbitrary commands on the target’s device. With the exception of the first variant, all GOVERSHELL implants were DLL files that were

loaded via search order hijacking from the legitimate version of either the 32- or 64-bit version of an open-source project called Tablacus Explorer.

Each variant of GOVERSHELL sets up persistence via a scheduled task on its first execution and includes a command-line flag in that persistence execution, which is required to execute the logic that includes C2 communication. If the flag is not present, the malware assumes it has been run for the first time, sets up persistence, and then exits. This has a potential side effect of evading sandbox dynamic detections, as the actual C2 traffic will not occur upon initial execution.

A high-level overview of some of the most notable differences in the GOVERSHELL variants is provided in the table below:

Variant	First Observed	C2 Comms	Command Execution	Language
1. Early (aka HealthKick)	April 2025	Fake TLS (Double header) to Port 465, XOR Encoded	CMD Reverse Shell	C++
2. TE32	June 2025	Fake TLS to Port 443, AES	PowerShell Reverse Shell	Golang
3. TE64	Early July 2025	HTTPS POST, Poll C2 for Commands, JSON Format	PowerShell + Predefined Commands	Golang
4. WebSocket	Mid July 2025	WebSocket's, AES, Command Task Queue Model	PowerShell + Partially Implemented Commands	Golang
5. Beacon	September 2025	HTTPS GET, B64 encoded, Jitter, Sleep	PowerShell + Predefined commands	Golang

The subsections that follow include more detail for each of the observed variants.

GOVERSHELL Variant 1 (Early)

Capabilities

- Can execute commands directly on the Windows command prompt (cmd.exe /c <command>).

Persistence

- Check for the presence of the -run command-line argument; if this is not present, it will set up persistence and exit. The malware is first copied to the following persistence location:
  - C:\ProgramData\{RANDOM_DIR_8_CHAR}.
- A scheduled task named SystemHealthMonitor is created and set to run every five minutes via the following command:
  - schtasks.exe /Create /TN "%hs" /TR "\"%s\" -run" /SC MINUTE /MO 5 /F

- This sample uses a legitimate binary named adobe_licensing_wf_helper.exe to perform search order hijacking in order to load malicious code via the file libcef.dll.

C2 Communication

- This GOVERSHELL variant attempts to blend in with legitimate network traffic by wrapping its C2 communication with a TLSv1.2 header. However, likely by mistake, this is done twice. The traffic is formatted as follows:
  - (17 03 03 [LEN WORD] (17 03 03 [LEN WORD] (Encoded payload)))
- There is no authentification with the C2 server for this variant. The payload, however, is encoded using a custom encoding function with the key mysecretkey.

GOVERSHELL Variant 2 (TE32)

Capabilities

- Can execute commands directly via a PowerShell reverse shell.

Persistence

- It checks for the presence of the **cuVn** command-line argument. If this is not present, it will set up persistence and exit. The malware is first copied to the following persistence location:
  - **C:\ProgramData\{RANDOM_DIR_8_CHAR}**.
- Persistence is then achieved through a scheduled task named MyGoTask , which is created through Windows's COM interface. The scheduled task is set to run the malware with the **cuVn** command-line argument every 15 minutes.
- This sample is loaded via search order hijacking through the 32-bit version of Tablacus Explorer via a file named **te32.dll** located in a directory named **lib**.

C2 Communication

- This variant attempts to blend in with legitimate network traffic by wrapping its C2 communications with a TLSv1.2 header and encrypts its content with the AES (CFB) cypher. The AES key used is supersecretkey16. This sample eliminates the second TLSv1.2 header found in the Variant 1.
- The malware's AES cypher is initialized with a randomly generated 16-byte IV. This value is sent to the C2 in plain text along with an Authentication packet consisting of the byte 0x01 followed by the C2 password your_secure_password. This is done such that the C2 can authenticate the malware and set up its own AES cipher.

GOVERSHELL Variant 3 (TE64)

Capabilities

- Supports multiple native and dynamic command execution via PowerShell. The malware expects the following commands:

- **builtin**
  - Data that follows is passed to **powershell.exe -NoProfile -Command <command>**
- **time**
  - Get the formatted current time on the victim's machine
  - Example: "2006-01-02 15:04:05"
- **sysinfo**
  - Retrieve the following information about the victim's machine:
    - OS
    - CPU Architecture
    - Number of CPU cores
    - Hostname
- **interval**
  - Set the malware polling rate in seconds

Persistence

- It checks for the presence of the **cuVn** command-line argument. If this not present, it will set up persistence and exit. The malware is first copied to the following persistence location:
  - **C:\ProgramData\{RANDOM_DIR_8_CHAR}**
- Persistence is then achieved through a scheduled task named MyGoTask, which is created through Windows's COM interface. The scheduled task is set to run the malware with the **cuVn** command-line argument every 15 minutes.
- This sample is loaded via search order hijacking through the 64-bit version of Tablacus Explorer via a file named **te64.dll** located in a directory named **lib**.

C2 Communication

- The malware polls the C2 over HTTPS at a regular interval (default is 10 seconds) to obtain the command to execute. Once the configured interval has elapsed, the malware sends an HTTP GET request to the following URL:
  - /fgbwwezskdfbeadgalidegsdfhfhaWhatHappenedTask?fgbwwezskdfbeadgalidegsdfhfhaclient_id=%s
- If the C2 accepts the malware's command request, it returns JSON data (application/json) which will be parsed as a te64_payload_Task structure:

```
struct te64_payload_Task // sizeof=0x48

{
  string ID; // The ID of the current task.
  string Command; // The command to execute.
  string ClientID; // The victim's IP address.
  time_Time Time; // The timestamp when the task was sent.
};
```

- The contents of Command are parsed and run based on what is described in the Capabilities section above. The result of the command's execution is then formatted as JSON data (application/json) from the following structure:

```
struct te64_payload_Result // sizeof=0x58
{
  string TaskID; // The ID of the current task.
  string ClientID; // The victim's IP address.
  string Output; // The command's output.
  string Error; // The commands's error output.
  time_Time Timestamp; // The timestamp when the task was run.
};
```

- The resulting JSON data is sent to the C2 with an HTTP POST request to the following URL:
  - /fgbwwezskdfbeadgalidegsdfhfhaWhatCanIdoResult

GOVERSHELL Variant 4 (WebSocket)

Capabilities

- Supports two native commands that allow for execution via PowerShell or a further command subsystem which was not fully implemented in the samples Volexity analyzed. The following commands were supported:
  - **system**
    - Supports the sub-command **update** , which is unimplemented
  - **command**
    - Run a PowerShell command via **powershell.exe -NoProfile -Command <command>**

Persistence

- It checks for the presence of the **cuVn** command-line argument. If this is not present, it will set up persistence and exit. The malware is first copied to the following persistence location:
  - **C:\ProgramData\{RANDOM_DIR_8_CHAR}**.
- Persistence is then achieved through a scheduled task named MyGoTask or UPnPHostUpdater, which is created through Windows's COM interface. The scheduled task is set to run the malware with the **cuVn** command-line argument every 15 minutes.
- This sample is loaded via search order hijacking through the 64-bit version of Tablacus Explorer via a file named **te64.dll** located in a directory named **lib**.

C2 Communication

- The malware connects to the C2 over WebSocket and communicates with the C2 using JSON encoded data that is encrypted with the AES (GCM) cypher using a dynamically established session key.

- When connecting to the C2, the malware waits for a message of type key_exch that contains the session key in the session field , which is encrypted by the malware's master key AES (GCM). The master key used varie per sample. Further exchanges are then encrypted using the established session key.
    - The following is a list of the observed master keys:
        - topIBApru76wra8REBrIb1it52H6B9Ap
        - 626bmcGzKuKfRvk4hW4pM3g70Q8XyBsq
        - nO3esWO4ucaCHLxayeblswO5iTRL37Ab
    - The AgentID
        - MD5 of { hostname}-{mac address}-{os name}-{CPU architecture} with format FFFFFFFF-FFFF-FFFF-FFFF-FFFFFFFFFFFF.
    - The HostID
        - Base64-encoded BCrypt hash of {HostName}{MAC Addr}{Boot time} which is then sent to the C2, along with the device's metadata and a Base64-encoded 16-byte nonce. The malware registers the victim's device with the C2. To do so it generates two IDs:
        - This registration message has the following JSON format:
          {"id": "<AgentID>","host_id": "<HostId>","metadata": {"Hostname": "…","OS": "…","Arch": "…","Username": "…","IP": "…",},"nonce": "<Base64 16 Byte nonce>"}

GOVERSHELL Variant 5 (Beacon)

Capabilities

- Supports multiple native and dynamic command execution via PowerShell. The malware expects the following commands:
    - **builtin**
        - Data that follows is passed to **powershell.exe -NoProfile -Command <command>**
    - **checkin**
        - Triggers "instant checking"
        - Affects te64_payload_immediateCheckin
        - Effect: Next sleep delay is zero (0) seconds
    - **jitter**
        - Usage: jitter <0-100> %
        - Sets the jitter percentage
        - Affects te64_payload_jitterPercent
        - Effect: Set how much the base sleep delay should be randomized
    - **sleep**
        - Usage: sleep <seconds>
        - The value ranges from 0 to 599800
        - Affects te64_payload_sleepTime
        - Effect: Set the base sleep delay (the delay between two client C2 request, jittered for stealth)

Additional GOVERSHELL Observations

One GOVERSHELL sample obtained from a mid-June 2025 phishing campaign contains a string referencing a folder path from the developer's device that contains Simplified Chinese characters: **C:\Users\Dev\Desktop\20250608新码\lib\te64**. The machine translation of these characters is "new code". Other developer paths observed in additional GOVERSHELL samples suggest there is more than one system being used to develop the malware. However, this was the only sample whose path included Simplified Chinese characters. Similarly, the WebSocket variant (Variant 4) of GOVERSHELL contains log statements in Chinese characters, while other variants have these statements in English.

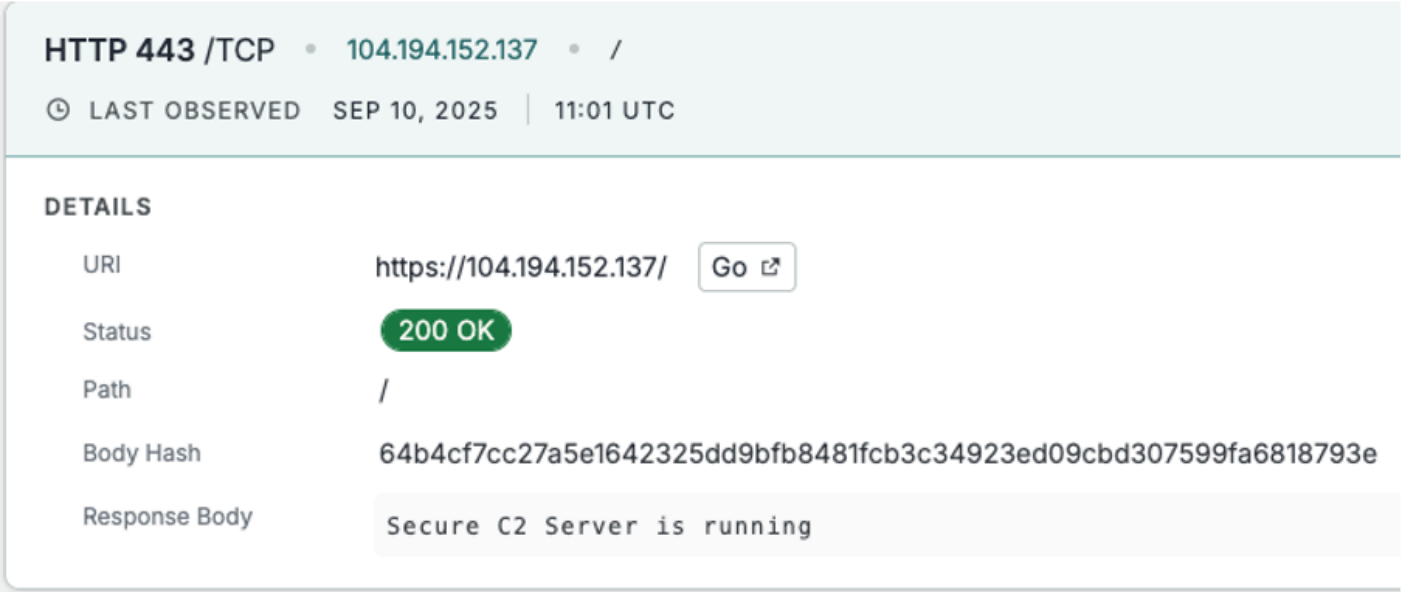
The number of rewrites of the network stack of the GOVERSHELL malware family could be a data point that supports the case for LLM usage, a case made later in this blog post. The development of GOVERSHELL does not appear to have been iterative, which is a commonly observed pattern for human development. Instead, each variant implements a new communication method, new capabilities, and rewrites of how basic functionality works (such as how a command should be executed).

Infrastructure

As the GOVERSHELL network stack has changed over time, so has the related C2 infrastructure. C2 was typically direct-to-IP until mid-July 2025, after which the threat actor switched to using DNS names and domains UTA0388 had registered. Domains used by UTA0388 for GOVERSHELL are named in the following ways:

- References to Taiwan, such as **moctw[.]info** and **twmoc[.]info**
- Impersonations of large organizations or legitimate-sounding services, such as **cdn-apple[.]info**, **azure-app[.]store**, **doccloude[.]info**, **sliddeshare[.]online**, and **windows-app[.]store**

UTA0388 domains are consistently registered and hosted behind Cloudflare. The C2 servers for the WebSocket variant have a default response showing *"Secure C2 Server is running"*. A screenshot of Censys Platform's record of this is shown below:



A Case for Identification of LLM Usage

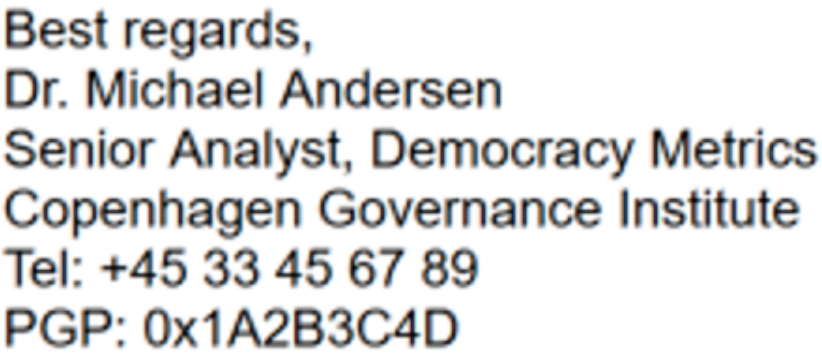
Before discussing UTA0388's potential LLM usage, it is important to understand how LLMs generate their output, and therefore how use of LLMs might be identified. A brief primer on this is provided in the Appendix .

Fabrications and Nonsensical Usage

This section gives examples of fabrications and nonsensical usage that may suggest LLM usage.

Emails

UTA0388 impersonates multiple entities in their phishing emails, but many of these entities are fabrications rather than impersonations of real-world personas. An example of this is shown in the following screenshot of the signature block from a spear phishing message sent by UTA0388. Neither "Copenhagen Governance Institute" nor "Dr. Michael Andersen" are real entities.



The phone number includes "**3 45 67 89**," a sequential pattern that suggests fabrication. The PGP key identifier intermingles "1234" and "ABCD" patterns, which is another clue that this was fabricated. Use of

predictable patterns for values like these are an inherent trait of LLM-generated output.

In several emails, the domains used in the email signatures were fabricated domains that do not exist or otherwise do not have MX records or any active DNS resolutions, and therefore would not be able to receive email. One example is **researchanalytics.co[.]uk**. It could be argued this is a human decision to add legitimacy to the phishing email; however, in Volexity's experience, the inclusion of a non-existent domain for this is extremely rare; more commonly, a legitimate and related domain is included. This tactic was used by UTA0388 in some of their most recent phishing emails, with the first email containing no malicious content and only subsequent replies containing the malicious links to archives. It is Volexity's view that the inclusion of a fabricated domain is reflective of an LLM's propensity to fabricate information.

Later, in one UTA0388's campaigns, the use of friendly names in other emails became more atypical and included pornographic references, like the following two examples:

- porndude2025 <LaurenBlackwell3278@proton.me>
- pornhublis <LaurenBlackwell3278@proton.me>

The use of fabricated personas and details is not necessarily proof of LLM usage; threat actors commonly fabricate details. However, it is usually observed as part of low-effort phishing operations that have similarly low-effort email body content, which was not the case with UTA0388. This contributes towards a picture of nonsensical usage, especially when factoring in details like the friendly names that seem implausible usage for a serious human operator.

Easter Eggs

UTA0388 campaigns were consistent in delivery of archive files containing a benign executable and a malicious GOVERSHELL payload in the form of a DLL found with a folder name lib. These files generally had names that would appear to be a legitimate document file and had themes focused on Asian geopolitical issues. However, in several instances, additional files were included in these archives. The additional files were not used as decoys to display to users, nor did they serve any purpose in the malware's operation.

The first odd file inclusion was a pornographic image that was modified to include brightly colored lines drawn over the image. The lines spelled out some text over the image: "TES", "XWX", and "NO". Volexity observed multiple other instances of pornographic media being included in the archives by UTA0388. In another instance, a file named Meeting-Cooperation introduction video by Jun 30.pdf" was included in the archive within the lib folder. This file was actually a 53 MB MPEG-4 file that contained pornographic content. In another case, a hidden file had a name that was a long Base64-encoded string that decoded to *"I am the sun from Korea, shining at 1000 degrees to burn you all!"*

In several archives, the threat actor included a waveform audio file (WAV) that is a recording of the reading of the Nīlakaṇṭha Dhāraṇī, a religious recitation that is popular in the Chinese form of Mahayana Buddhism. There were also several text files included in the archives, one of which just contained the *"text no!can i help me%"*. And finally, in some archives the threat actor included the same benign executable multiple times with

different names. There is no functional reason for this inclusion, as duplicate instances do not change the execution of the final payload.

Volexity cannot suggest a clear reason for a human to include the additional files mentioned in this section. Their inclusion appears to be nonsensical and counterproductive for the success of the campaign, as odd files are more likely to raise suspicion among targets. At the same time, these files were not designed to be seen by the recipient and were often, but not always, in the hidden lib folder. It is possible these files served as either intentional or unintentional “Easter eggs”.

Lack of Coherence

Volexity identified more than 50 unique phishing emails sent by UTA0388. In addition to emails written in English, Volexity also observed emails written in Chinese (Mandarin), German, French, and Japanese. Each contained target-specific text, were of reasonable length, did not follow a consistent template, and appeared to be in fluent natural-sounding language. Fluent crafting of emails in a wide variation of languages is unusual in a single campaign due to the language skills required by the humans crafting them, and the introduction of errors or awkward phrasing that can occur when using direct machine translations.

The fluency of the language in the emails was not reflected in the coherence of their use. For example, Volexity observed an email sent to an English-speaking target that was supposedly sent from an American persona, but the email had a subject line in Mandarin and a German message body. Another email was sent to a European target purportedly from a Spanish-speaking author but written in Japanese. It is not unheard of for threat actors to send poorly tailored phishes or emails to the wrong targets by accident; however, this campaign consistently lacked coherence in a way that is more suggestive of context-unaware automation.

This lack of coherence was also reflected in persona use, such as the example below; note the following:

- The email “friendly name” of **GeoffreyLewisMD3850**
- An email address of **ChristopherDelgado5328@proton.me**
- The introduction and email signature of **Michael Brown**

Proposal for Strategic Collaboration in Global Market Expansion

G

GeoffreyLewisMD3850 <ChristopherDelgado5328@proton.me>

Tuesday, July 15, 2025 at 02:29

To:

Dear ,

I hope this email finds you well. My name is Michael Brown, and I am the Director of Business Development at Global Solutions Inc., a leading firm specializing in innovative technology solutions and international market expansion. With over 15 years of experience in the industry, we have successfully partnered with organizations worldwide to drive growth and technological advancement.

I am reaching out to explore potential collaboration opportunities between Global Solutions Inc. and . Given your esteemed organization's focus on global peace and policy, we believe our expertise in technology transfer and market expansion could complement your initiatives, particularly in emerging markets.

Document.zip
5.1.mb • PDF Document

We have identified several areas of mutual interest, including:
1. Technology Transfer: Sharing cutting-edge solutions to enhance operational efficiency.
2. Market Expansion: Leveraging our global network to support your outreach
3. Investment Opportunities: Joint ventures to fund high-impact projects.

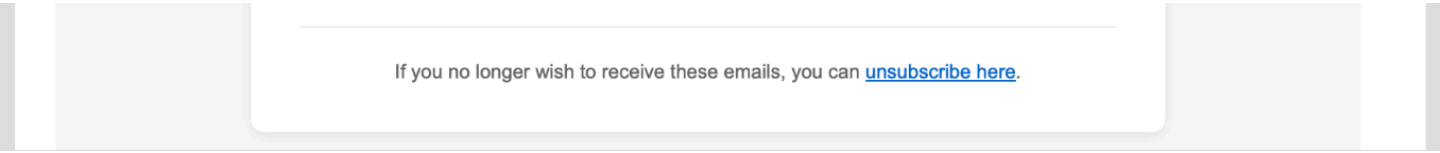
Attached, you will find our detailed proposal outlining the potential benefits and a roadmap for collaboration. We believe this partnership could yield significant commercial and societal value.

To discuss this further, I would be delighted to arrange a meeting at your convenience. Please let me know a suitable time, or feel free to contact me directly at +1 (212) 555-9510 or michael.brown@globalsolutionsinc.com.

Thank you for considering this proposal. I look forward to your positive response.

Best regards,

Michael Brown
Director of Business Development
Global Solutions Inc.
5864 Innovation Drive, Suite 567
New York, NY 10001, USA
Phone: +1 (212) 555-9510
Email: michael.brown@globalsolutionsinc.com
Website: www.globalsolutionsinc.com



The use of three personas in a single email does not align with human patterns, where a single persona is typically used in order to appear legitimate. Humans do make mistakes, and so more than one may be used accidentally when using a template or not creating new email accounts for different targets. However, the use of three different personas becomes less likely without some kind of automated input. This was also observed repeatedly, which suggests context-unaware automation rather than manual curation. The example email above also provides a fake phone number (starting with “555” after the area code) and invites the recipient to contact them using that number. The domain used in the email, globalsolutionsinc[.]com, is actually registered. However, it is just a parked website and does not have an MX record or accept email at the A records returned for the domain. In short, there would be no way to contact “Michael Brown” at the means specified in the email.

In terms of targeting, the majority of the phishing emails observed by Volexity were sent to addresses that were identified as being visible on publicly accessible webpages. This is not out of the ordinary; however, in several instances the attacker sent phishing emails to addresses that were clearly example data of email format and not real email addresses, e.g., *first.last@<domain>*. This could be a human mistake, but a known shortcoming of LLMs is their inability to understand the context of the data they are processing. Other targeted emails included a webmaster address, group contact addresses (such as *info@domain*), individuals no longer working at the target organization, and the email address of a podcast, all of which were available online. This pattern suggests automation, LLM or otherwise, that is not fully context aware.

Other incoherent details in the phishing emails included the following artifacts:

- Wrong day/date combinations
- Asking to be contacted by phone without providing a number
- Non-existent departments at real institutions
- Wrong names for targets

Overall, the phishing campaigns often lack coherence and contain multiple errors of a nature that leads Volexity assesses with high degree of confidence the threat actor used an LLM to craft the phishing emails in this campaign, with little oversight of whether the output was plausible or not.

Technical Artifacts

In a later phishing email the threat actor provided a link to an archive that contained the usual benign binary and GOVERSHELL DLL, as well as a benign Microsoft Word document file. This document contained metadata that indicated it had been created using python-docx, an open-source Python library. This library is documented online as being used by multiple LLMs to generate Word documents. This is not conclusive evidence of usage of these platforms but is another data point that supports the assessment that UTA0388 makes use of LLMs in their operations.

Conclusion

The evidence presented in this blog post provides insight into a persistent and active threat actor that conducts phishing campaigns using a single, if ever changing, malware family. The targeting profile of the campaign is consistent with a threat actor interested in Asian geopolitical issues, with a special focus on Taiwan. When combined with several technical artifacts that indicate the author of GOVERSHELL uses Simplified Chinese, this leads Volexity to assess with a high degree of confidence that UTA0388 operates in the interest of the Chinese state. The activity does not significantly overlap with any other existing threat actor that Volexity tracks.

Making the case for LLM usage by a threat actor can be difficult, as no single data point is conclusive enough to definitively prove its use. Volexity has detailed different aspects of UTA0388's campaign that illustrate an incoherent and nonsensical pattern of behavior that would align with LLM usage without oversight. This body of evidence leads Volexity to assess with a high degree of confidence that UTA0388 used LLMs to support its operations.

The emails and files used in this campaign leads Volexity to assess with medium confidence that UTA0388 made use of automation, LLM or otherwise, that generated and sent this content to targets with little to no human oversight in some cases. It is not clear if this is agentic AI usage, automation, or just a human operator that did not review and correct the outputs. The frequency of phishing emails sent through July 2025, where the threat actor was observed sending 26 emails in a three-day period to targets across Volexity's visibility, also supports the assessment.

Volexity does not have sufficient data to be able to say whether UTA0388's foray into LLM-powered campaigns has been a success, but the volume of tailored phishing output (even if sometimes in the wrong language) will yield a significant number of opportunities to successfully gain access to targets. UTA0388's activity appears to have slowed down from its peak in July 2025 but remains a consistent threat. The observed continued development of the GOVERSHELL malware family speaks to intended ongoing activity.

To detect UTA0388 related activity Volexity recommends the following:

- Use the IOCs listed here.
- Use the rules provided here.

KEY TAKEAWAYS

Volexity is tracking an unknown threat actor, UTA0388, and has observed the following:

- UTA0388 is a threat actor operating in the interests of the Chinese state that conducts spear phishing operations aimed at gaining access to the systems belonging to its targets
- The threat actor has leveraged OpenAI's ChatGPT to assist in crafting its spear phishing emails, obtaining targets of interest, and in the development of its malware.
- There is active, ongoing development and deployment of multiple variants of a custom malware family, GOVERSHELL, with five distinct variants identified at the time of writing.

- UTA0388 has been active since April 2025 and continues to be active, iterating its operational tradecraft through September 2025.