Threatray – Getting started

Session 2 – Intelligence capabilities

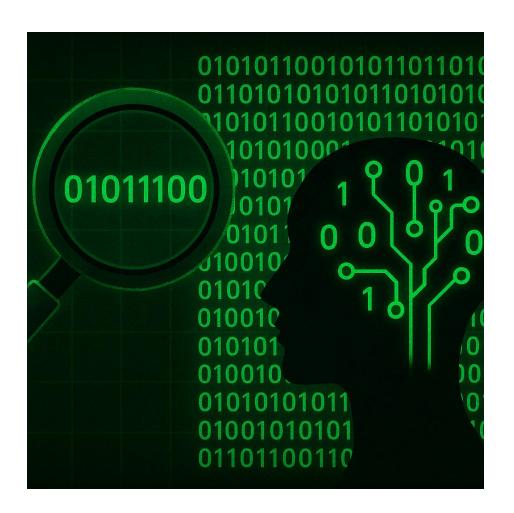
Content

- Background information on our intelligence capabilities
- Classical search, finding IOCs fast
- Code search (finding OSINT through code and retro-hunting for code)
- Code intelligence

Intro

Threatray intelligence capabilities

- Threatray not only has deep malware analysis capabilities to analyse a malware "in isolation" but also unique capabilities in malware intelligence.
- Intelligence analysis in Threatray has two aspects intelligence aided investigations of new malware, as well as intelligence production.



Threatray intelligence capabilities

- Intelligence analysis is malware data and search driven.
- For the data side, the Threatray platform contains a public and private malware repository.

The public repository is shared by all Threatray users and contains malware curated by Threatray:

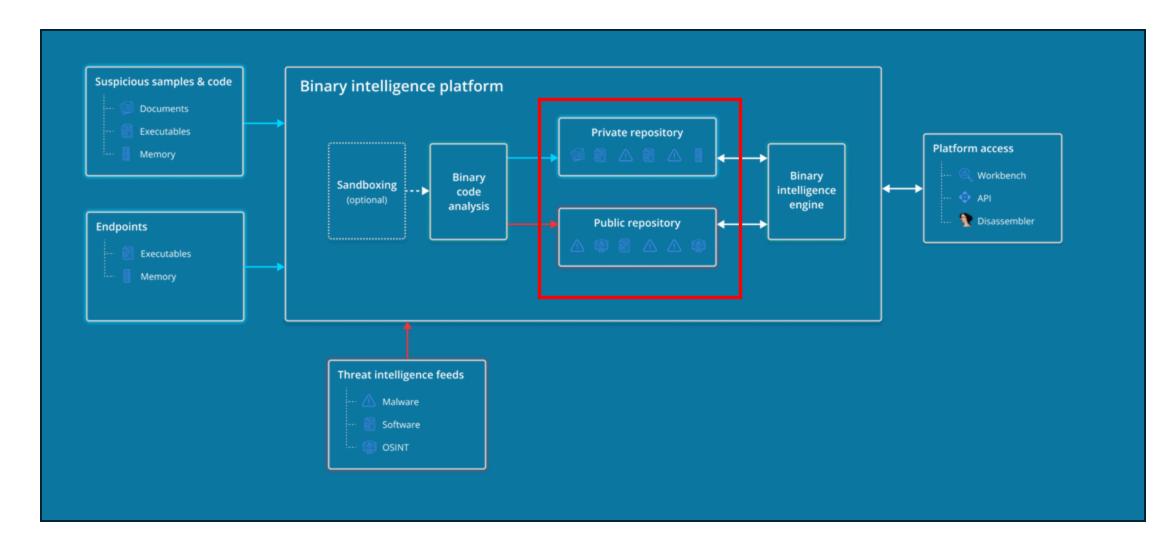
- Malware samples from partners, from harvesting OSINT sample,....
- NOT from our users!
- Contains cybercrime, APTs, C2s etc.

 For many of our users the private repository also serves as a malware repository for their own malware, making it analysable and turning their own malware collection into intelligence.

New Threatray users, can **import their existing malware** collection into Threatray.

- For the search side of things, we provide
 - **classical search** (e.g.,. for IPs, domains, process names, mutexes,...) and
 - unique code search capabilities that allow you to find code-wise similar samples to the one under investigation.

Threatray architecture – Two Data Repositories



Classical search capabilities

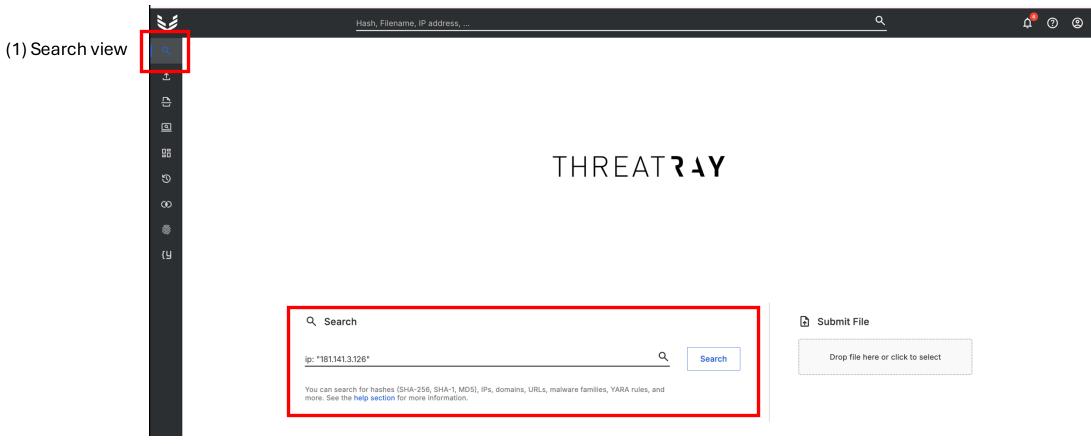
Classical search

- Classical search is like what you know from existing malware platforms, with the important difference, that you can search our public repository and your private repository ("we make your private analyses searchable")
- Search across public and private is seamless, so that you can relate what your own malware to malware seen globally and vice versa.
- What is unique in Threatray is the statistics view over the search results, allowing to identify cluster properties and IOCs (see later)

- ip: Contacting an IP or IP range (ex: 192.168.1.0/24)
- url: Contacting an URL
- domain: Contacting a domain
- file: Modifying a file (path)
- mutex: Modifying a mutex
- registry : Modifying a registry key
- process: Process name and command line
- signature: Malware family detection
- vara: YARA rule match
- verdict: Sample verdict (unknown, suspicious, malicious)
- label: Label assigned to a file during submission
- sample-name: Submitted file name
- analysis-id: Analysis ID
- file-hash: Submitted file hash (MD5, SHA1, SHA256)
- memory-hash: Memory region hash (MD5, SHA1, SHA256)

Entities we can search for

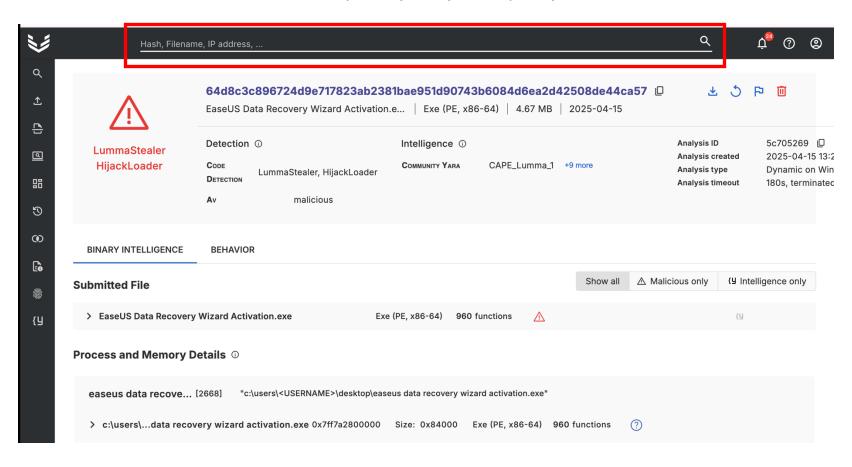
Where can you launch a search?



(2) Enter query here

Where can you launch a search?

There is a search bar on top of any analysis to quickly launch searches



Where can you launch a search?



In the behavior report, there is a search icon next to each dynamic artifact.

Search syntax

We support autocompletion, and wildcards "*", for AND you can simply use multiple queries in the same search bar, no OR at the moment.

ip: "181.141.3.12	26"	Contacting an IP or IP range
url: "*/Plugins	/cred64.dll"	Contacting an URL
domain: "geo.	netsupportsoftware.com"	Contacting a domain
file: "c:\users\	<username>\appdata\local\temp*\files_\system_info.txt"</username>	Modifying a file (path)
mutex: "xtrem	eupdate"	Modifying a mutex
registry: "HKE	Y_CURRENT_USER\SOFTWARE\Remcos*"	Modifying a registry key
process: "cmc	d /c set /a \"0x*^*\""	Process and command line
signature: "Co	baltstrike"	Malware family detection
yara: "CAPE_L	.umma_1"	YARA rule match

See here https://docs.threatray.com/docs/search

Search example - Finding IOC

• In this worked examples we show how we can use search to identify a mutex type IOC for the Latrodectus malware family.

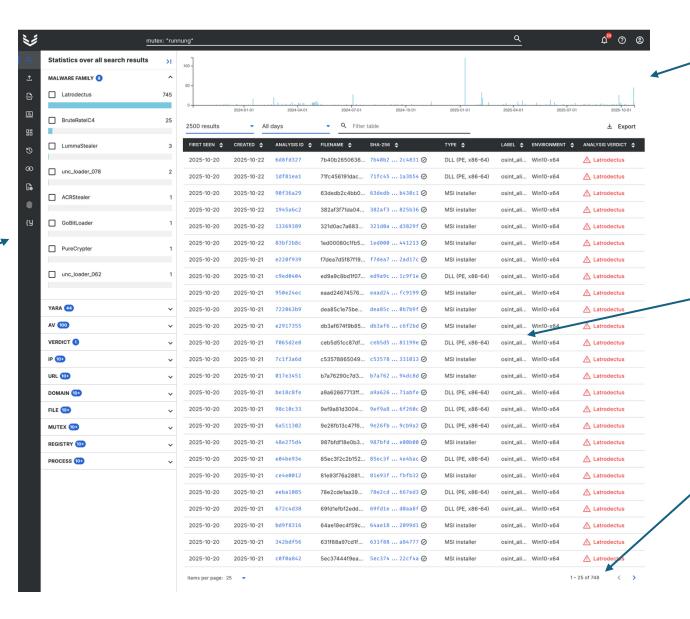
• We are investigating the dynamic artifacts in the analysis view and click on the "running" mutex to search for it.



Search example – Finding IOC

Static results over the search result set.

Think of it as "cluster statistics"



Distribution of first seen dates of samples

Matching samples / analyses.

Pivot to any of them

Total analyses found (748)

THREAT 7 4Y

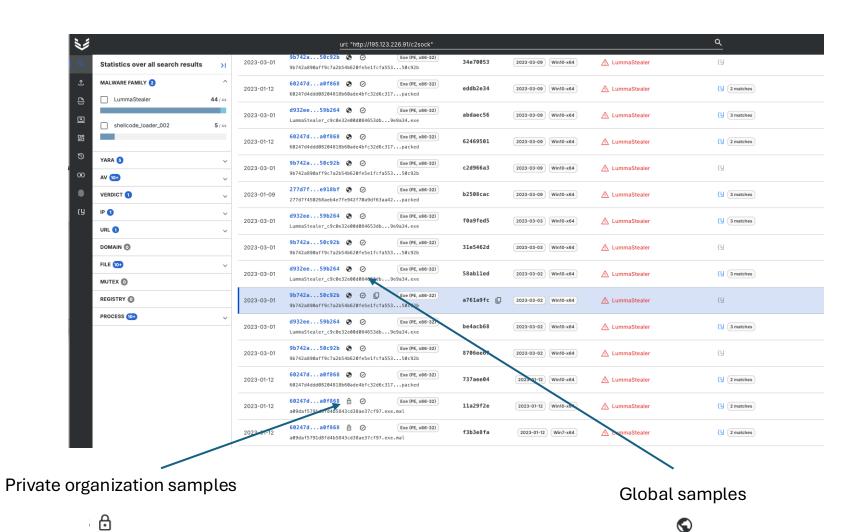
Search example - Finding IOC

The "cluster statistics" for the "runnung mutex" show the following (see previous slide):

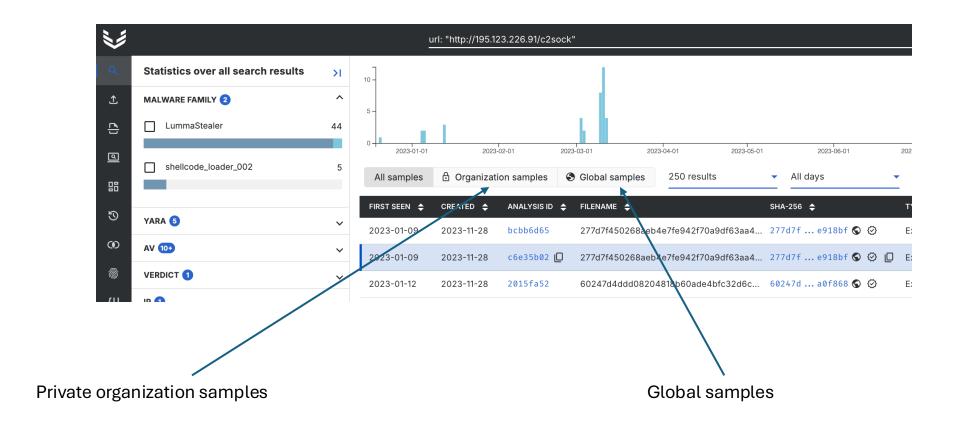
- We have found 839 files with that mutex
- Of those 836 files 839 are classified as Latrodectus

→ The mutex "runnung" and the "Latrodectus" family are strongly correlated, and thus very likely the "runnung" mutex is a good IOC for the family.

Private, global, samples from OSINT



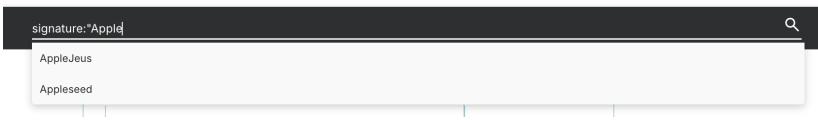
Filtering for private and global samples



Searching for malware families

 When investigating malware families, a searches by family name can be tremendously useful.

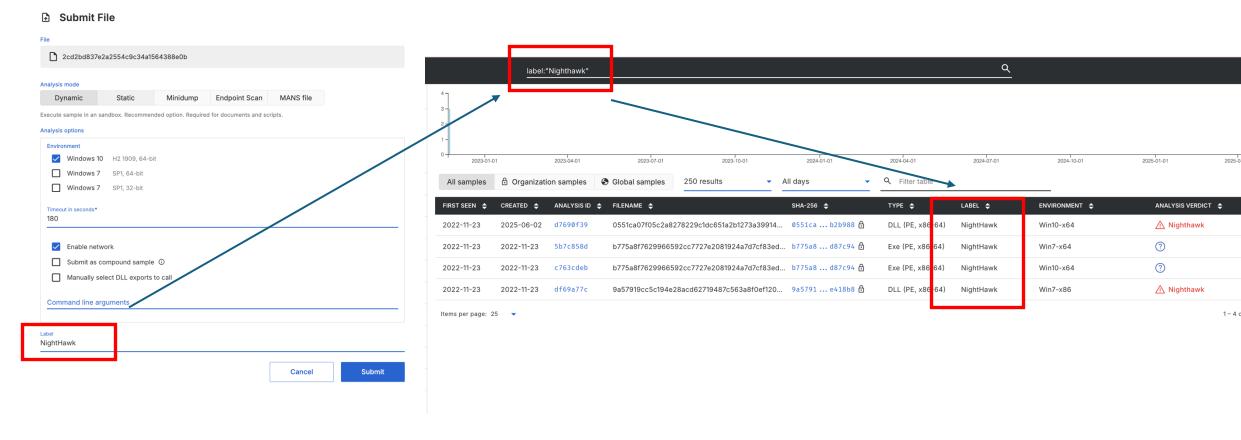
For family name searches we have auto completion, use the "signature:" search operator:



You can also search by YARA rules names:

• **Note:** The "signature:" search covers all samples, whereas "yara:" search only covers samples that were ingested after a YARA rule became active.

Search by label



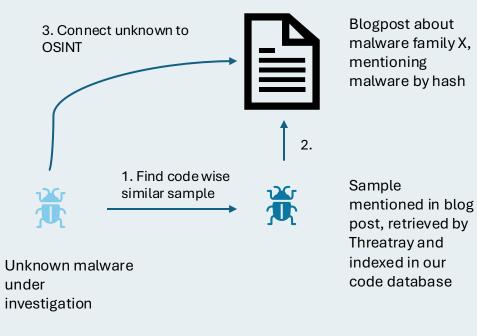
Search & find samples by using the label that was given upon samples submission

Code search

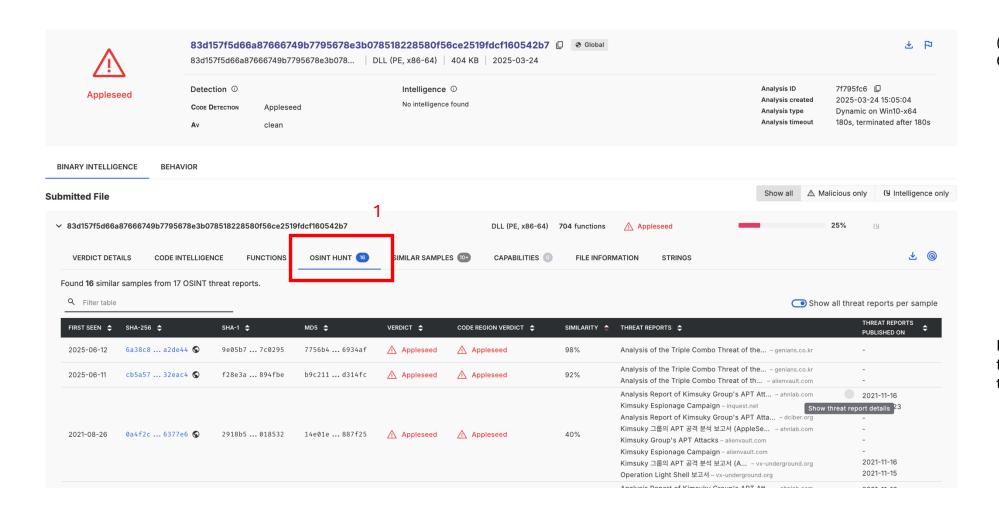
OSINT search

- What is OSINT search? The feature finds OSINT reports, like blogs, research articles, Tweets, that mention malware samples which are code-wise similar to the code region under investigation. In short: it finds relevant OSINT through code similarity. This capability significantly enhances our ability to discover relevant OSINT and accelerates investigation processes.
- **Discover OSINT in a Whole New Way.** This is a completely new and unique approach to searching for OSINT reports.
- Why is it useful?
 - **Uncover hidden insights:** Find OSINT you wouldn't discover through traditional search methods.
 - **Get highly relevant results:** Identify OSINT directly related to the code you're investigating often far more precise, as it's based on similar or related samples.
- OSINT search is available on a "code block level", i.e., for files and process memory regions.

Technically, we do this by continuously retrieving OSINT reports, extracting the hashes, downloading the samples that correspond to these hashes, and indexing the code of malware samples into our malware database. We then use our code search engine to match code regions under investigation to OSINT samples.



OSINT search example

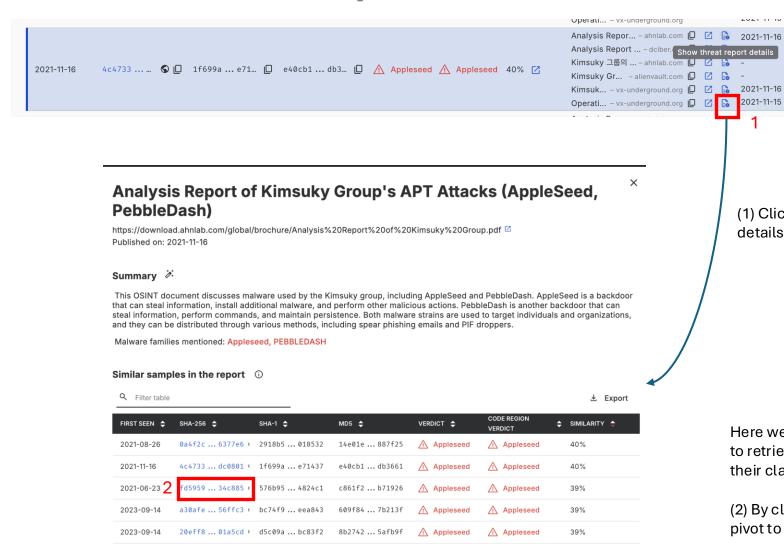


(1) The OSINT results are found in the OSINT tab of a code block

Each line line in the table is sample found in an OSINT report, that is similar to the code block under investigation

OSINT search example

Items per page: 5



(1) Click on the "Show threat report details icon to open a report summary.

2021-11-15

1 - 5 of 12 < >

Here we show all samples we were able to retrieve from the report, including their classification.

(2) By clicking on the blue hash, you can pivot to any of them.

OSINT search example

Summary of the Appleseed example:

- Through OSINT search we find a blog post by Ahnlabs on the Appleseed family, used by Kimsuky actor. The report contains rich information, we quickly learn important details about the actor and two malware families, PebbleDash and Appleseed. We also learn that there is HTTP and SMTP version of Appleseed.
- By searching for the hashes of the similar samples in the report (using normal text find), we quickly learn that our sample is similar to HTTP versions of Appleseed.
- We find in the report a common URL pattern used "/?m=a&p1=[PcID]&p2=[PcInfo]- [MalwareVersion]".
- We go back to our sample to the "Strings Tab" and find the the "/?m=" string in the sample under investigation, which definitely confirms that we're dealing with an Appleseed sample.
- We can also use the URL pattern search $\underline{ur1:"*/?m=b\&p1=*"}$ and pivot to many more sample in our global repository.

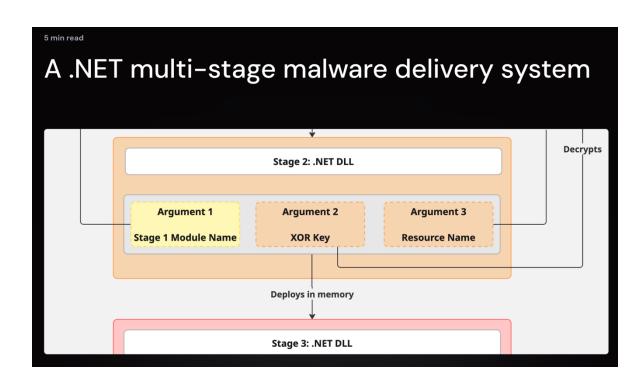
Code retro-hunts

- What is code retro-hunt? It is a unique search feature in Threatray, that allows to "search for similar code".
 The feature starts from any code-block in the sample under investigation and finds samples that are code wise similar (more precisely: it finds analyses that contain a similar code block).
- Fast code pivots. Traditional tech like VT use YARA to search for similar code. This approach has its merits, but it takes time and skill. Threatray is point and click, thanks to our native code search capabilities.

- Why is it useful? It is feature that has many applications, but also and advanced feature, since depending on what code you hunt for the results, and their interpretation are quite different. It takes time to learn and appreciate this feature. Some examples:
 - **Pivot to similar samples,** e.g., when you need multiple samples from same family for YARA rule development, to find common IPs used, etc.
 - Relate samples from different investigations in your private repository
 - Intelligence investigations of droppers and downloaders
 - Finding that an APT is always using the same libraries or other code components
 -
- **Retro-hunt is available on a "code block level"**, i.e., for files and process memory.

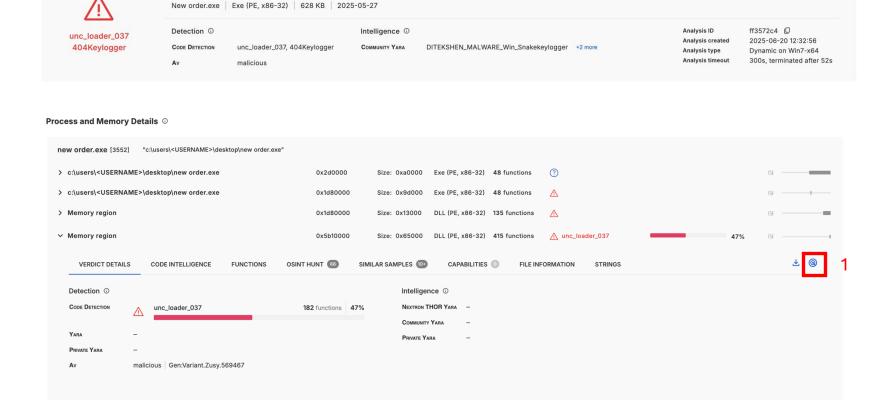
Code retro-hunt example

- In this example we are using the code-retro hunt investigate a loader and find out what malware it is dropping.
- It is a real-world example in the sense that we have discovered a previously unknown malware loader, see blog post here: https://www.threatray.com/blog/a-net-multi-stage-malware-delivery-system



Code retro-hunt example

ffe925e06253c9e33952ee8e425dc85b5d42889de8eb6a4bbfdc59b85cbbe604 D Siobal



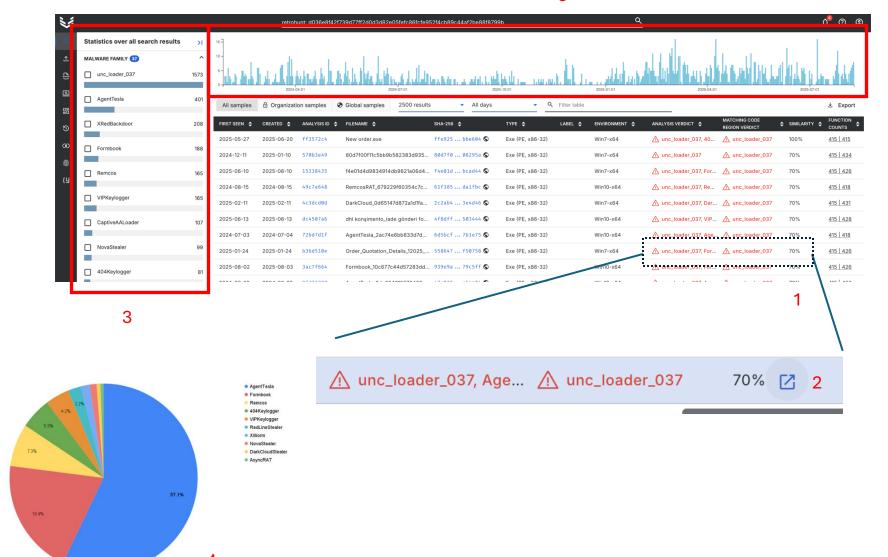
We want to find samples that use a code wise similar loader to the "unc_loader_037" to better understand that loader.

平日

(1) We click on the retro-hunt icon next to the code region (a)

Code retro-hunt example





The search results look the same as when you use classical search – after all, retro-hunt is a search, a search for code.

(1) Unlike with classical search, we also show how code-wise similar the samples in the result set are.

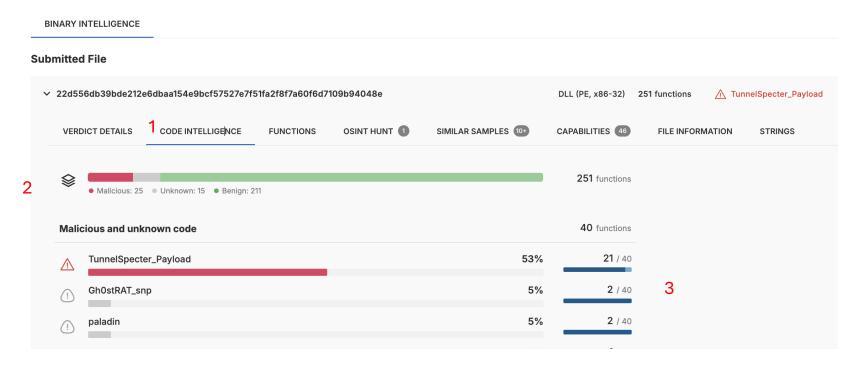
Matches of 80%+ similarity are typically reliable, lower similarity can be helpful or not.

- (2) You can pivot to any sample in the search results. The click will take you you to the *matching code block (!)* in the analysis of the samples.
- (3) The cluster stats show us which families are deployed by this loader: AgentTesla, XRedBackdoor, Formbook,.... This is how we produced the stats (4) in our blog.
- (5) The timeline shows us that this loader has been active since beginning of 2024 → relevant loader to track.

THREAT 7 4Y

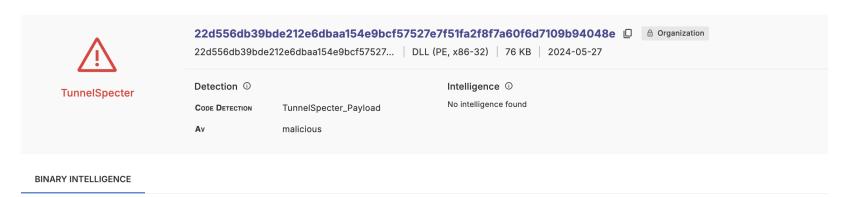
Code intelligence

Code intelligence (code DNA)

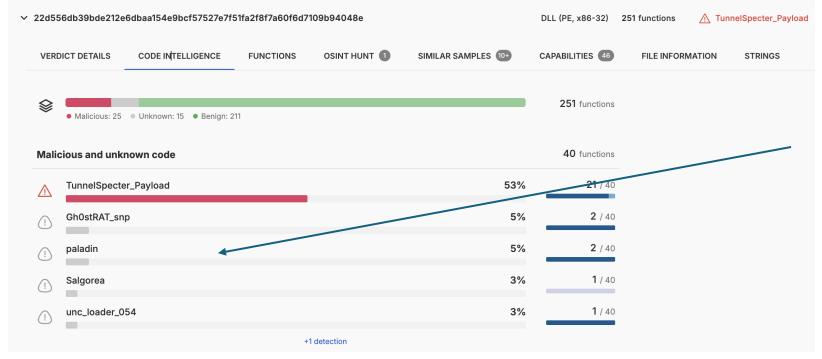


- (1) The code intelligence tab for each code block shows in detail the code composition of the that code block.
- (2) We distinguish malware code (red), unknown (grey) and green (benign). The stats are on a per function granularity, i.e., we attribute functions to families.
- (3) We also show how confident we are we the matches. Moreover, since we match by similarity, a few matches are often noise / useless, the more functions match the more reliable.
- This is an advanced feature, which is useful in some cases that matter, but less useful in others. Also, it might require further analysis in IDA pro (see next session)

Code intelligence (code DNA) – Example Chinese APT



Submitted File



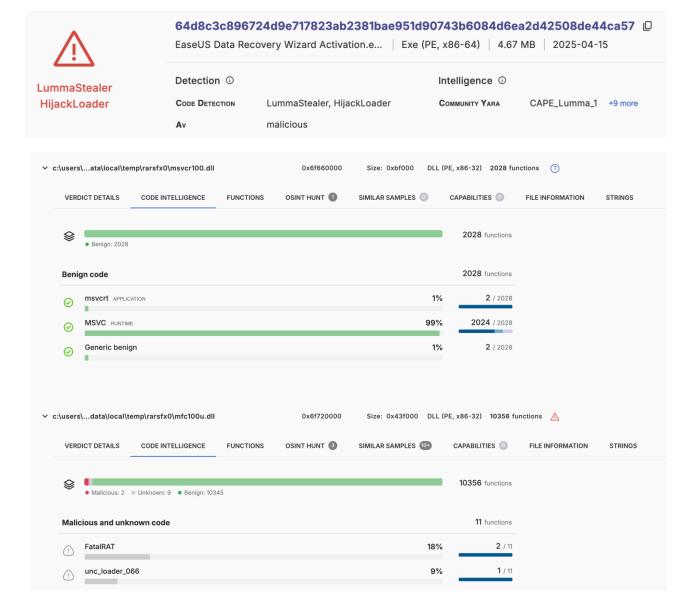
This is a sample we were investigating. Initially, we didn't know that it is TunnelSpectre.

But the code fragments from Gh0stRAT, paladin etc. put us on the China track.

Also it seems to indicate that the developers of these tools share some code.

THREAT ? AY

Code intelligence (code DNA) – Backdoored DLL



Sample drops multiple of DLLs (can be seen in dynamic analysis file operations).

In the code intelligence tab, we see that some of the DLLs are 100% benign (1) and others are mostly benign but contain 9 unknown functions (2).

Using manual analysis in IDA pro we found that these 9 functions are malware code.

How to replay in your Threatray instance

Here are the analysis IDs for the examples shown:

- For the Latrodectus "running" mutex
 1a3ee1f5-94b9-457b-bc50-a8ce85e2ceb1
- For the Appleseed APT OSINT example, we have used analysis 7f795fc6-0892-40d5-a5c2-5e1e38528b0e
- For the 404Keylogger loader retro-hunt we have used ff3572c4-1893-4a35-818f-ebcbde33b092
- For the code intelligence example on Chinese APT we have used c31bb52d-7a29-4c27-8e57-eb69ac1e1bb5
- For the code of the LummaStealer backdoored DLL we have used 5c705269-e076-4ad2-bfa8-1a514ce08dbf