Following the LNK metadata trail

By Guilherme Venere

Published: 2023-01-19 · Archived: 2026-04-05 23:17:54 UTC

Thursday, January 19, 2023 08:00

- Adversaries' shift toward Shell Link (LNK) files, likely sparked by Microsoft's decision to block macros, provides the opportunity to capitalize on information that can be provided by LNK metadata.
- Cisco Talos analyzed metadata in LNK files and correlated it with threat actors tactics techniques and procedures, to identify and track threat actor activity. This report outlines our research on Qakbot and Gamaredon as examples.
- Talos also used LNK file metadata to identify relationships among different threat actors. In this report we demonstrate this by using metadata to connect Bumblebee with IcedID and Qakbot respectively.

Executive Summary

Microsoft announced at the beginning of 2022 that they would soon start to disable macros by default in Office documents downloaded from the Internet. They implemented the changes around June, only to remove the feature later that month. The feature was finally re-enabled by the end of July. Cisco Talos observed threat actors reacting to these changes by moving away from malicious macros as an initial access method in favor of other types of executable attachments.

While tracking some prevalent commodity malware threat actors, Talos observed the popularization of malicious LNK files as their initial access method to download and execute payloads. A closer look at the LNK files illustrates how their metadata could be used to identify and track new campaigns.

An overview of LNK file format

Microsoft describes the Shell Link binary file format - the format used by Windows files with the extension "LNK" - as a file which contains information that can be used by the Operating System or an application to access another data object on the system. Although the format itself has applications supporting Object Linking and Embedding (OLE) object access, it is more commonly used as "shortcuts" to applications and file locations on the file system.

The LNK structure stores information about the target object as well as other related information about the application behavior and metadata from the machine where the LNK file was created. These metadata sections can contain optional data about various attributes of the target file. Among these attributes, a few provide valuable information to identify the exact system where the file was created. Other researchers have published some good information on the basic features present in LNK files so we won't delve too deep in the details, but the most important fields for our research are the following:

- **Modification/Access/Creation (MAC) timestamps**: A FILETIME structure that specifies the MAC time of the LNK target in UTC.
- **VolumeID**: An optional VolumeID structure that specifies information about the volume where the link target was on when the link was created. Of special interest to analysts is the field **DriveSerialNumber** which is unique and can be associated with a specific **Disk Device.**
- **MachineID (16 bytes)**: A NULL–terminated character string, as defined by the system default code page, which specifies the NetBIOS name of the machine where the link target was last known to reside.
- **DROID (CDomainRelativeObjId) GUID**: Two values in GUID packet representation which are used to indicate the location of the link target. There are different values for the Volume identifier as well as the File identifier. The last section of the File identifier is generated based on the **MAC Address** of the machine where the file resides, which is another important correlation data.
- **DROID Birth GUID**: The same as the field above, but stores the location where the LNK file originally pointed to. Together with DROID GUID these two fields could indicate if a LNK file was moved to a different system after creation.
- **Metadata Store structure**: Extra structures of attributes used to describe additional details about the link target. Sometimes, when the target information is not present on the LNK file, we might need to look at the Metadata Store structure for additional data. The fields "**ItemFolderPathDisplayNarrow**", "**ParsingPath**" and security identifier ("**SID**") in this structure might provide valuable data to identify the LNK target as well as embedded payloads.

In order to extract these attributes, several tools are available in the public domain to parse and analyze the LNK structure. Google provides a free command line tool called "LNK Parser", which we are going to use in this blog to demonstrate the examples, but there are other options like LeCMD or the Python library "LnkParse3". Some of these tools are able to parse even malformed LNK files, which is a characteristic of some files created by threat actors.

LNK Builder Tools

With the increasing usage of LNK files in attack chains, it's logical that threat actors have started developing and using tools to create such files. Several tools have been documented before and many are available publicly or through a paid subscription:

- MLNK Builder
- Quantum Builder
- Macropack
- LNKUp
- Lnk2pwn
- SharPersist
- RustLnkBuilder

These different tools may sometimes leave traces in the LNK metadata, which can be useful for detecting or tracking malicious actors. By examining the output payload of some of these builders, we can see that most of them wipe out most metadata from the file. This could be used as a good indicator of suspicious behavior, as these fields normally are present by default when the shortcut is properly created.

SharPersist Payload

Figure 1: SharPersist payload

In Figure 1 we can see the MAC timestamps are empty, but the LNK file still contains information about the relative path where the file was created, as well as the SID of the user who created the file.

Quantum Builder Payload

Figure 2: Quantum builder payload

In Figure 2, we can see the Quantum Builder does not attempt to wipe out any metadata and the generated LNK file contains a trove of metadata about the system where it was created. In this example, we also see the Disk Serial Number, user SID and DROID information including the MAC address of the machine used to create the file. A quick search for that MAC address indicates it is from a VMWare virtual network card.

Meterpreter Payload

Figure 3: Meterpreter payload generated for MS15-020 (CVE-2015-0096)

In the LNK files created by Meterpreter, as well as other exploit frameworks like Cobalt Strike, the metadata are completely wiped, with the malicious code present in fields not normally parsed by LnkParser. In Figure 3, the LNK target is pointing to a DLL file on a remote share but since this is a malformed file created to exploit a bug, the parsing tools cannot see the payload.

Figure 4: Meterpreter payload viewed in an Hex editor showing the payload

Comparing the data present in these examples we can start to see patterns that may be useful to detect these malicious files. Most of the tools tend to have wiped out data simply because the APIs used to create such files don't require all fields to be present, so they implement the bare minimum to have the malicious code running. A simple YARA rule could be used to detect such samples:

```
rule lnk_wiped {
    meta:
      author="gvenere"
      description="LNK with wiped metadata"

    strings:
        $lnk_magic = { 4c 00 00 00 }
        $ext1 = ".js"  // additional strings to search
        $ext2 = ".bat" // in the LNK target area
```

```
        $ext3 = ".cmd" // These are for Qakbot


    condition:
        $lnk_magic at 0x0 and
        uint16(0x1c) == 0x0 and // CreationTime == 0x0
        uint16(0x24) == 0x0 and // AccessTime == 0x0
        uint16(0x2c) == 0x0 and // WriteTime == 0x0
        // To target specific families we can add additional checks here
        ( any of ($ext*) in (0xa0..0x100) )
}
```

Other tools use the proper methods to create these files, but that opens an opportunity to identify information specific to the machine where the sample was created. Looking at the Quantum Builder example in Figure 2, it's possible to see attributes which identify the Disk and Machine where the LNK was created.

The information present in the LNK files can prove extremely valuable when it comes to tracking specific threat actors in the wild.

LNK files as Initial Access Tool

When Microsoft announced the changes to macro behavior in Office at the end of 2021, very few of the most prevalent malware families used LNK files as part of their initial infection chain. In general, LNK files are used by worm type malware like Raspberry Robin in order to spread to removable disks or network shares.

However, Talos observed a steady increase in LNK file usage by main malware families starting at the beginning of the year, with a big spike by the time Microsoft implemented the changes in Office 365. Looking for VirusTotal (VT) data for the past year, and searching exclusively for files related to prevalent malware families, we can see the following trend (Figure 5):

Figure 5: LNK files used as initial access mechanism for prevalent malware families. Source: VirusTotal

When we look at specific families using these files, we can observe Qakbot as the main source of files in their last two campaigns, one starting in May and ending in July, and the other starting around the beginning of August and ending in November.

Figure 6: LNK file telemetry mapped to malware families. Source: VirusTotal

Interestingly, many malicious LNK files submitted to VT during this period had all the metadata removed from the file. Looking at Qakbot data, we see the group started to use LNK files with wiped metadata during the August campaign. Wiped metadata could also be explained by the increasing usage of toolkits to generate such files, as explained before.

The data in Figure 6 also depicts a decrease in activity around July and August, followed by another spike around October and November. This could be explained by the announcement of vulnerabilities which allowed malware to bypass the Mark-of-the-Web flag used by Microsoft Defender and other Antivirus products to decide whether to scan or not a file. As reported by other sources, this bug was exploited by many malware families and could explain the second spike in LNK files usage as delivery mechanism.

Threat Actor Tracking

Qakbot

Qakbot (also known as Qbot or Pinkslipbot) is one of the oldest malware operations still in activity. First observed in the wild around 2007, it is still one of the most active malware families today, as we recently reported in Talos's Q2 Quarterly Report.

Qakbot is known to evolve and adapt their operation according to the current popular delivery methods and defense techniques. As recently as May 2022, their preferred method of distribution was to hijack email threads

gathered from compromised machines, and insert attachments containing Office XLSB documents embedded with malicious macros.

However, after Microsoft announced changes to how macros were executed by default on internet downloaded content, Talos found Qakbot increasingly moving away from the XLSB files in favor of ISO files containing a LNK file, which would download and execute the payload.

While examining the content of these LNK files used in the last 6 months of campaigns, we observed some interesting characteristics. Looking at the information present in the metadata about where these samples are created, we see that there is no overlapping metadata between the different campaigns. Additionally, looking at the Top 200 samples in VT known to be part of a Qakbot campaign, we see what machines were the most active in generating these LNK files (Figure 7):

Figure 7: Distribution of LNK files related to main Qakbot campaigns

Metadata information can also help detect correlations between these actors and other malware. For example, we analyzed the samples below, which are all part of the “AA” Qakbot campaign from June and July 2022:

- 8fda14f91e27afec5c1b1f71d708775c9b6e2af31e8331bbf26751bc0583dc7e
- 2f9da7145056a4217552a5a536ceb8365e853fbd04d28ae2d494afb20e9c021f
- 52458b4aaddbcb04048be963ea7d669c2ff7a69642d027f88812a5c6c1ade955
- 6a980d7659efb8bfb997dec3259d6eb090d4e6a4609e4c0666e04ad612151d71
- 67bbffb2ff5f724a201445f26018cb09fbf0588689f98f90fd82082aae7c6eec
- da2a0d9a6b5dd2123c4c2cbd55d81fd22ab72bf7ceb1489a5a770e10bcf67137
- 54681cbb4c61dd4fe03341cfd8d2b796366a0372b53dd3e1d52c9e6ff98692d1
- a7f31c98147d98ac08f4b8afe7faa2f2b4aab821655717f4bde519fcd87300ac

- c5c0daaa26815bb6528332dd4f56f7eb72db4456d5a84b8bc69239c45079a1c4
- efdb91497fe213e8f696065c2fe81f64cbaa219da16e2b3f8e1e146d098652b5
- c9dfafd3536977289b4bfda1369fbd113a778cf06ac0c01cdc8e00e1c300e774
- e818b0115a9a877a9517c99b16e5a2df9cf7c5eb1fb249d9153b68e8fa94e60b
- 7ba3eaee591cc73ab85aeb09d8c02b1e569b9dcaffcbc7c4473f504f939697d2

The metadata in these samples indicate they were created on a machine with Drive Serial # “0x2848e8a8”. Looking at VT for this serial we found samples that were both related to the AA campaign, and linked to other malware.

For example the sample beacb63904c2624ae02601f283671b3ef61650109aea3259b63a0aeefe4133fa, which was submitted on August 15th, 2022, contains Powershell code to download and execute a binary from hxxp://88.198.148[.]231/u.exe.

According to VT information, this sample with hash 6161c01fd590c98c6dee4e510ba9be4f574c9cc5c89283dbff6bb79cd9383d70 is detected as a Redline variant (Figure 8):

Figure 8: VirusTotal details for Redline sample

Starting in August 2022, Qakbot resumed activity, including their two main campaigns, “Obama” and “AA”. Interestingly, at this point the “Obama” Qakbot campaign began wiping out the metadata on their LNK files, but the Target field always pointed to the JS, BAT or CMD file used to start the malicious DLL part of the initial infection chain (Figure 9):

Figure 9: Metadata for LNK files part of "Obama" campaign

The same behavior started to happen on "BB" campaign's LNK files starting on September 13th. We can see in Figure 10 that samples mapped to the "BB" campaign in September still had Drive Serial Number information:

Figure 10: Metadata for LNK files part of "BB" campaign

We also observe that the Drive Serial Number matches the one used by the "AA" campaign from the May/July timeframe, which could indicate "AA" and "BB" are probably managed by the same group. More recently, samples from "BB" started to have their metadata wiped too.

Gamaredon

In June 2022, Inquest published a report about a new threat actor called Glowsand that was targeting Ukranian entities using phishing emails with malicious documents and LNK files to download and execute second stage payloads.

By analyzing the metadata content of the LNK file in the report, Talos associated the machine IDs where the files were generated, to files associated with the Gamaredon APT. Furthermore, based on this metadata, Talos identified a new campaign targeting Ukrainian organizations that started around August 8th, 2022, which we wrote about in a separate blog post. In fact, Gamaredon files reported as far back as Feb 2017 contained the same LNK metadata as the files found in our research.

On September 6th a new set of samples was identified during our hunts, which we also connected to previous Gamaredon campaigns via metadata. But the samples this time had an interesting feature: an embedded digital signature (digisig) pertaining to a Microsoft development unit in Puerto Rico.

- 7f66f4411983001d29236c5d3fb4ff26f01b5742badca1db8d49264c01ba506c
- 1b2ed05f488f8439688a02cc6ef84f939d16169117b489219b688a3ea482e5ed
- 6ce64dedbe81c36aef38fd2d567f6ab9737df708591dc2f0cafa56db26a1d043
- 1e0b92485e09ac970ae38214fb5c7407f73027ada47ea697017e49cacb576908

The samples had very few or no detections at all in VT, even though some of them are related to a Gamaredon campaign from 2016/2017, which may indicate the embedded digisig might have been added in an attempt to bypass AV detections.

Following the lead on the digisig name "Microsoft Operations Puerto Rico1" we can find a multitude of LNK files using the same technique. The interesting point here is that there is no provision in the LNK file format for a digisig, which means the digisig is probably present only as garbage data to confuse AV scanners and is a common technique used by many malware families.

Bumblebee relationship with IcedID and Qakbot

In addition to tracking malware groups' activity over time, LNK file metadata can help identify relationships among different threat actors. Researchers have identified relationships between Bumblebee and other malware families before, which we independently confirmed by looking at the LNK metadata.

In August 2022, a user submitted a test link file to VT, with a file path containing the individual's username: "Lamar". This LNK file was created on the same machine responsible for the LNK files used in a previous IcedID campaign. A few hours later the same user submitted another file with hash e89cd1999517b47805106111e14de4a03669cac30adb3b3304655febce25955f, this time with user information sanitized, and packaged in a Zip file containing a BAT and DLL file. The DLL file was an IcedID bot.

We also found a correlation between LNK files leading to IcedID infection, with LNK files used in a Bumblebee campaign, which both use the same Drive Serial Number:

Figure 11: Bumblebee and IcedID samples sharing the same metadata

By examining the relationships between these hashes in VT, we can see that the LNK file 9c7e01c2c39dadc020a0cf8dc74b62e6453b56413f09705b4ad4d391981f5a3f seen in Figure 11 leads to a Bumblebee DLL while other hashes like 3cca8d1b4cfe0ebcf105621700454d0285ef1b44dfed3e3abf70060bb62aa5b4 lead to an IcedID DLL. In both cases the same username is present in the ParsingPath field.

A similar approach was used to find a relationship between Bumblebee and Qakbot. During our research on Qakbot we found samples associated with the “Obama” campaign which still had the LNK information and used the Drive Serial Number “300D-05E9”. This same serial number was found later in LNK samples leading to Bumblebee infections:

Figure 12: Bumblebee and Qakbot samples sharing the same metadata

Just as an example, the sample 2738ee3f181994cca5d9ea19359b8142981583d17563934ab3212eefe13af3ff in Figure 12 leads to a Bumblebee DLL.

Conclusion

In the cyber threat landscape, any new information on the adversary could be critical toward improving defenses. In this blog Talos demonstrated metadata’s value using LNK files, but the same concept is applicable to other file formats that include metadata, or attack tools that leave signature traces about their use in their payloads.

By analyzing and tracking information leaked through metadata, and correlating this information with other actor’s tactics, techniques and procedures, defenders can develop better detections and even predict future behavior, to prepare for an attack.

Coverage

Cisco Secure Endpoint (AMP for Endpoints)	Cloudlock	Cisco Secure Email	Cisco Secure Firewall/Secure IPS (Network Security)
✓	N/A	✓	✓
Cisco Secure Malware Analytics (Threat Grid)	**Cisco Umbrella DNS Security**	**Cisco Umbrella SIG**	**Cisco Secure Web Appliance (Web Security Appliance)**
✓	N/A	N/A	✓

Cisco Secure Endpoint (formerly AMP for Endpoints) is ideally suited to prevent the execution of the malware detailed in this post. Try Secure Endpoint for free here.

Cisco Secure Web Appliance web scanning prevents access to malicious websites and detects malware used in these attacks.

Cisco Secure Email (formerly Cisco Email Security) can block malicious emails sent by threat actors as part of their campaign. You can try Secure Email for free here.

Cisco Secure Firewall (formerly Next-Generation Firewall and Firepower NGFW) appliances such as Threat Defense Virtual, Adaptive Security Appliance and Meraki MX can detect malicious activity associated with this threat.

Cisco Secure Malware Analytics (Threat Grid) identifies malicious binaries and builds protection into all Cisco Secure products.

Umbrella, Cisco's secure internet gateway (SIG), blocks users from connecting to malicious domains, IPs and URLs, whether users are on or off the corporate network. Sign up for a free trial of Umbrella here.

Cisco Secure Web Appliance (formerly Web Security Appliance) automatically blocks potentially dangerous sites and tests suspicious sites before users access them.

Additional protections with context to your specific environment and threat data are available from the Firewall Management Center.

Cisco Duo provides multi-factor authentication for users to ensure only those authorized are accessing your network.

Open-source Snort Subscriber Rule Set customers can stay up to date by downloading the latest rule pack available for purchase on Snort.org. Snort SIDs for this threat are: 61099-61102, 300367-300368

The following ClamAV detections are also available for this threat:

- Lnk.Dropper.Agent
- Lnk.Trojan.Qakbot
- Lnk.Trojan.BazarLoader
- Lnk.Downloader.Agent
- Win.Dropper.Agent

Indicators of Compromise

Indicators of Compromise associated with this threat can be found here.
