Examining Emotet's Activities, Infrastructure

blog.trendmicro.com/trendlabs-security-intelligence/exploring-emotet-examining-emotets-activities-infrastructure/

November 16, 2018

Malware

We did comprehensive research on Emotet's artifacts — 8,528 unique URLs, 5,849 document droppers, and 571 executables collected between June 1, 2018 and September 15, 2018 — to discover it's infrastructure as well as possible attribution information.

By: Trend Micro November 16, 2018 Read time: (words)

Discovered by Trend Micro in 2014, the banking Trojan Emotet has been brought back to life by malware authors last year with its own spamming module that has allowed it to spread, target new industries and regions, and evade sandbox and malware analysis techniques. This year, we examined Emotet's activities to learn more about how this modular malware wreaks havoc: We did a comprehensive research on Emotet's artifacts — 8,528 unique URLs, 5,849 document droppers, and 571 executables collected between June 1, 2018 and September 15, 2018 — to discover Emotet's infrastructure as well as possible attribution information.

Some of the highlights of our research include the following:

1. **We discovered that there are at least two infrastructures running parallel to one another that support the Emotet botnet.** By grouping the C&C servers and the RSA keys of the malware, we were able to see two distinct groups of infrastructures. We also saw that the threat actors switched RSA keys on a monthly basis. While the next-stage payloads each group pushed does not show any major difference in terms of purpose or targets, the differing infrastructures of both groups may be designed to make it more difficult to track Emotet and minimize the possibility of failure.
2. **Multilayer operating mechanisms might have been adopted in the creation of Emotet's artifacts.** The inconsistency between the activity patterns show that the infrastructure used to create and spread document droppers are different from those used to pack and deploy Emotet executables. The creation of document droppers stops during the non-working hours between 1:00 to 6:00 (UTC). Meanwhile, there might be three sets of machines that are used to pack and deploy Emotet's executable payloads, two of which are probably set to the time zones UTC +0 and UTC +7, respectively.
3. **The author of the Emotet malware may live somewhere in the UTC+10 time zone, or further east.** After we grouped the executable samples by their unpacked payloads' compilation timestamps, we found two sample groups that showed an inconsistency between the compilation timestamps and the corresponding first-seen records in the wild. This might lead to the possibility that the compilation timestamps are pointing to the local time on the malware author's machine. The conclusion of the malware author probably staying in UTC +10 time zone or further east can therefore be drawn if the local time is accurate.

Emotet's two infrastructures

We have collected and analyzed 571 executable samples of Emotet. The configuration inside an executable includes a list of C&C servers and an RSA key for connection encryption.

There were only six unique RSA public keys extracted from the executable samples. Each RSA key has a 768-bit modulus and uses the public exponent 65537. We calculated the CRC32 of each RSA key blob and gave each key a name for easier identification.

Key Name	CRC32	Emotet Group
A	fcb2fb3b	1
B	86e9acef	1
C	ceff5362	1
D	fc8e8aaa	2
E	8f1eb5e	2

F	aef0def8	2

Table 1. The RSA keys extracted from Emotet executables

Meanwhile, Emotet's C&C server is an IP/port pair on top of its HTTP protocol. We extracted 721 unique C&C servers in total. On average, one Emotet sample contains 39 C&C servers, with a maximum number of 44 and a minimum of 14. Based on our observation, only a few C&C servers embedded in a single Emotet sample are actually active.

We found that most of the C&C servers are located in the United States, Mexico, and Canada. The top 3 ASN connected to Emotet are ASN7922, ASN8151 and ASN22773.

Figure 1. Countries wherein Emotet C&C servers are distributed

Figure 2. Distribution of Emotet C&C servers' ports

We visualized the relationship between each RSA key and its set of C&C servers and discovered that there were two RSA key groups. Keys A, B, and C were in one group (Group 1), and keys D, E, and F were in another (Group 2).

Figure 3. Relationships between RSA keys and C&C servers. Each blue dot represents a unique C&C server, while the red ones indicate RSA public keys.

As Figure 3 shows, these two distinct groups do not share C&C servers.

Figure 4. Timestamps when the RSA keys were received. The green dots represent the keys used by Group 1, while the orange dots represent the keys used by Group 2. Each dot represents the timestamp when each RSA key was found in the configuration of a new sample.

Month	June	July	August	September
Keys used by Group 1	fcb2fb3b (A)	86e9acef (B)	86e9acef (B)	ceff5362 (C)
Keys used by Group 2	fc8e8aaa (D)	8f1eb5e (E)		aef0def8 (F)

Table 2. Two groups of RSA keys and the corresponding months they have shown activity

Our analysis shows a link between the dates the RSA keys were received and the two groups' activities: each RSA key was observed to have been used for one month before threat actors switched to another RSA key on the first working day of the succeeding month (i.e. Jul. 2, 2018 and Sep. 3, 2018, both fall on a Monday).

We also observed that there were more artifacts belonging to Group 1 compared to those in Group 2. Based on our data, we received 469 unpacked Emotet samples for Group 1 and 102 for Group 2, respectively. We also did not find any activity for Group 2 for the month of August, as shown in Figure 4.

Two different Emotet groups, two different agendas?

Our initial assumption was that the two Emotet groups were created for different purposes or are being utilized by different operators. To prove this assumption, we referred to data from @malware_traffic and categorized the IoCs respectively. However, we did not find any major difference between the IoCs under these two groups. For instance, TrickBot with gtag arz1 was found to have been sent by Group 1 on September 20 and by Group 2 the next day. Without any strong evidence, we can only tell that the two groups might be different infrastructures designed to make tracking Emotet more difficult and help minimize the possibility of failure.

Date	Emotet Group	RSA Key	Next-stage Payload
2018-07-03	2	E	Panda Banker
2018-07-09	1	B	Panda Banker
2018-07-16	2	E	Panda Banker
2018-07-19	2	E	Panda Banker
2018-07-30	1	B	Panda Banker
2018-07-31	1	B	Panda Banker
2018-08-08	1	B	Trickbot
2018-08-10	1	B	Panda Banker
2018-08-13	1	B	Panda Banker
2018-08-14	1	B	Panda Banker
2018-08-15	1	B	Panda Banker
2018-08-16	1	B	Panda Banker
2018-08-22	1	B	Panda Banker
2018-08-24	1	B	Panda Banker
2018-08-26	1	B	Panda Banker
2018-09-04	2	F	IcedID, TrickBot

2018-09-05	2	F	IcedID, AZORult
2018-09-06	1	C	IcedID, AZORult
2018-09-14	1	C	TrickBot gtag: del72
2018-09-20	1	C	TrickBot gtag: arz1
2018-09-21	2	F	TrickBot gtag: arz1, del77, jim316, lib316

Table 3. The next-stage payload delivered by Emotet's two infrastructures between July and September 2018

Compiling Emotet's Source Code for Each Infrastructure

Emotet payloads are protected by customized packers and obfuscators. We studied the compilation of timestamps against each sample before and after packing and saw that some of the timestamps in packed samples were forged, while some seemed legitimate. The samples with legitimate timestamps show just a few minutes difference from being compiled to that of when they were found in the wild. For example, sample SHA256: 648dce03ac4c32217ce5c0b279bc3775faf030cafb313c74009fe60ffde3c924 (Detected by Trend Micro as TSPY_EMOTET.NSFOGAH) was compiled at 2018-06-06 05:40:17 and was found in the wild four minutes later. However, sample SHA256: 07deb1b8a86d2a4c7a3015899383dcc4c15dfadcfafc3f2b8d1e3aa89a6c7ac4 (Detected by Trend Micro as TSPY_EMOTET.TTIBBJD) was compiled at 2035-07-30 21:36:11, which is obviously a fake timestamp. Since it is difficult to distinguish legitimate timestamps from forged ones, research on the packed files' timestamps may prove to be fruitless.

Even though the compilation timestamp might be bogus, we decided to analyze the unpacked Emotet samples and saw that their timestamps seem legitimate. Out of 571 unpacked Emotet samples, only 11 distinct compilation timestamps were found. If the timestamp is forged during every compilation, the samples compiled with the same pieces of code should contain identical code sections but with different compilation timestamps. However, we found that the unpacked samples with the same timestamp share the identical code section, while differences can be found among those with different timestamps. The changes between the different timestamps also seem to be new-version updates.

Data in Table 4 show that the actor might have used automatic tools or scripts to compile Emotet's source code for each infrastructure, since a number of unique samples share the same compilation timestamp. The data also shows that the actors prepared the payload for Groups 1 and 2 sequentially. For example, on June 3, 2018, 46 Emotet samples were generated at 20:08 UTC by using Group 1's RSA public key and C&C servers. Two minutes later, 37 other Emotet samples were generated.

We noticed that the attackers tended to update Emotet samples on Monday or Wednesday (UTC). We also observed that the code section is exactly the same among the samples that had the same compilation timestamp. The only difference is the C&C servers embedded in the data section. It is possible that each time a source code is compiled, several C&C servers on the attacker's control list were chosen for generating a new sample.

Emotet Group	RSA Key	Compilation Timestamp in Epoch	Payload's Compilation Timestamp (UTC)	Unique Sample Count
1	A	1528056487	2018-06-03 20:08:07	56
2	D	1528056680	2018-06-03 20:11:20	38
1	B	1530547690	2018-07-02 16:08:10	28
2	E	1530547815	2018-07-02 16:10:15	25
1	B	1531161666	2018-07-09 18:41:06	31
2	E	1531161732	2018-07-09 18:42:12	18
2	E	1531899206	2018-07-18 07:33:26	57
2	E	1531906587	2018-07-18 09:36:27	5
1	B	1532502303	2018-07-25 07:05:03	276
1	C	1536011873	2018-09-03 21:57:53	21
2	F	1536011945	2018-09-03 21:59:05	16

Table 4. Unique samples collected in the wild with corresponding compilation timestamps

We will release more information about Emotet's technical details and also possible attribution-related intelligence at a later time.