# COM-301 Computer Security
# Exercise 11: Malware

## Malware

1. Are the following statements True or False:

   (a) Eliminating buffer overflows would completely prevent the problem of Internet worms

   (b) Some viruses add their code to that of existing executables residing on disk.

   (c) Having a centralized botnet with one command & control (C&C) is a good idea. You do not need to manage many computers since all report to the same C&C.

   (d) We cannot trust that executing a program will not do damage just by inspecting the program's source code

   (e) Trojans are self-contained programs that replicate themselves to infect other machines

   (f) Ransomware is a type of malware focused on network denial of service

   (g) Signature-based antiviruses are very effective against known viruses but not new

   **Solution**:

   (a) False - The worms can exploit other vulnerabilities to gain control over systems.

   (b) True - This is one of the modus operandi of viruses. They can also add themselves in macros or in the booting files

   (c) False - Having a unique point of failure is never a good idea. The more spread is the C&C the more difficult it is to shut down the botnet.

   (d) True - The source code needs to be compiled in order to get executed. That there is no problem in the source code does not mean that the compiler will not introduce vulnerabilities or malicious code in the executable.

(e) False - Trojans are not self-contained, and they do not replicate themselves. Only when combined with a worm they spread on their own.

(f) False - Ransomware does not aim at shutting down a network. It aims at temporarily locking a computer to obtain money for releasing it.

(g) True - The antivirus does not have the signature of the new virus, so the antivirus won't be able to detect it. This is the reason why signature based antiviruses require frequent updates.

2. You got access to an encrypted version of the final exam of COM-301 which is a file of 100KB encrypted with AES-128 (symmetric encryption with a key length of 128 bits). You really want to decrypt the file to get a very good score. It turns out that you have become a millionaire. For some reason instead of going to Hawaii, you decide that you are going to finish your degree and need to pass this course. Therefore, you really, really want to decrypt the exam. Assume you have three options:

   (a) Buy a botnet that costs you 100 CHF for every $2^{10}$ bots. Each bot can brute force $2^{40}$ keys a day (fictitious number!). You can buy at most $2^{24}$ bots.

   (b) Pay one of the assistants to leak you the key using a covert channel so that the professor does not realize. The covert channel leaks 4 bits of the key per day. The assistant does want to go to Hawaii for vacation every summer of his life, and since he has learned that you have became a millionaire, asks you for 25,000 CHF per bit.

   (c) Buy the latest coolest ever technology by Dr Nefario that can use a side channel attack from your seat in the classroom to get the cleartext version of the exam from the professor computer that can extract 1KB of the file every week. This coolest tech costs 1M CHF, but you are a millionaire and you do not care.

   Which option is the fastest? And which is the cheapest? Justify.

   **Solution**:
   First, we analyze each approach to approximate the cost and delay of decrypting the exam.

   (a) The attacker can buy between $2^{10}$ and $2^{24}$ bot nodes. We study the two extremes:

   $2^{10}$ nodes: Cost = 100 CHF. Computation power = $2^{50}$ keys per day. Expected time to break the decryption: $2^{128}/2^{50} = 2^{78}$, aprox $3*10^{23}$ days.

   $2^{24}$ nodes: Cost = $2^{14}*100$ CHF = 1,638,400 CHF. Computation power = $2^{64}$ keys per day. Expected time to break the decryption: $2^{128}/2^{64} = 2^{64}$, aprox $2*10^{19}$ days.

(b) Cost = 128 * 25,000 = 3.2 Million CHF. Time = 128/4 = 32 days.

(c) Cost = 1 Million CHF. Time = 100KB / 1KB = 100 days.

Getting a small 1024 machine botnet with 100 CHF is the cheapest way of getting the exam, but it takes 221 years.

Asking TAs is the fastest way of getting good grades!!! It only requires 32 days.

3. Malware can use different approaches to choose to which machines in the network to spread. Compare each of the following approaches for the following points: i) how well they work to spread faster, ii) how effective they are if the malware has a pre-defined target, iii) how well they work for remaining undetected, iv) how well they work to obtain the widest coverage.

   (a) Random (i.e., try random nodes in the network) vs Targeted (i.e., select new targets from a pre-define set)

   (b) Full (send to every reachable node every time it spreads) vs Limited (restrict the number of nodes in each spreading step)

   **Solution**:

   (a) Random vs targeted.

   The speed of random choice depends on the number of nodes tried every time. If the malware uses full, this is the fastest way. Targeted is fast if the attack is targeted (i.e., the set of victims is predefined). This mode is not good if the target is pre-defined. There is no guarantee that the random search will find it fast. Targeted is optimal for this task.
   Random search is more likely to be detected, since it will try many options increasing the possibility of hitting a monitored victim. Target is more likely to not be detected as it only touches a limited amount of nodes.
   Random is good for widest coverage. It ensures that, over time, all possible victims will be infected. The targeted mode will likely not obtain large coverage, but it does not matter since it is usually selected when the victim set is pre-defined.
   *Stuxnet used a targeted spreading to target Iran's nuclear infrastructure while WannaCry used a random spreading since they didn't care about the identity of the victim.*

(b) Full vs limited

Full spread is likely the fastest. Using this every node will try all of the possible available targets, getting to them faster. Limited nodes is not a fast way to cover all the network.
Full mode should be better than limited if the target is pre-defined, since one would find it faster even using random search.
A limited spreading tries to keep a constant number of infected machines while full spreading aims to infect as many machines as possible. Security experts and antimalware softwares inspect systems to find malwares. Having less infected nodes makes the detection of the malware harder.
Full spreading is likely to provide wider coverage, giving the attacker more computational power. Limited spread can also obtain wide coverage, but at a much slower pace.

*Slammer did a full spreading to infect a large set of machines. Stuxnet restricted itself to a small number of infected machine to remain undetected.*

4. Explain what type of Intrusion Detection System (IDS) would you choose for (justify your answer):

   (a) Detecting a well-known worm that rarely mutates

   (b) Detecting unknown worm

   (c) Detecting a known worm that randomizes its operation all the time (i.e., it never repeats the same pattern).

   (d) Detecting this particular worm that acts by sending 5 SYN-ACK at the beginning of the attack

   **Solution**:

   (a) Signature based. If the worm is known and never mutates it must be possible to learn a signature that can be used to identify it.

   (b) Anomaly-based. If you do not know how the threat is going to look like, best try to detect anything that is not in the universe of good things.

   (c) Anomaly-based. Even being known, it is not possible to find a signature. Only anomaly detection is possible.

   (d) The 5 SYN-ACK are a signature. Just add it to your Signature-based IDS.