Data and Network Security

(Master Degree in Computer Science and Cybersecurity)

Lecture 2

Google Classroom

4b6bekh

Outline for today

- Recap last lecture
- Malware types
- Emerging threats

Outline for today

- **Recap last lecture**
- Malware types
- Emerging threats

What is this course about?

This course provides a comprehensive introduction to fundamental concepts, principles, and practices in cybersecurity focusing into emerging trends and future directions in the field.

The goals of DNS

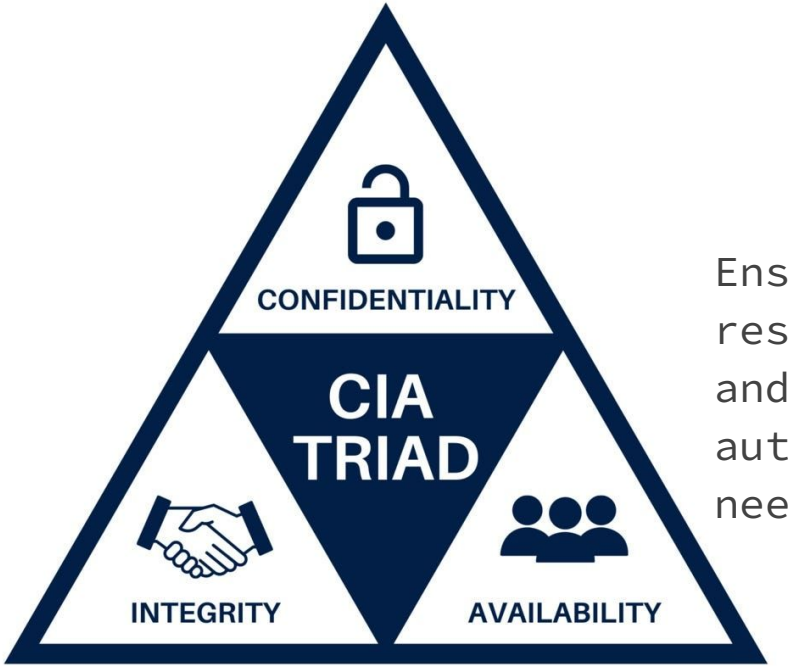
Confidentiality

Protecting sensitive information from unauthorized disclosure.

Integrity

Ensuring the accuracy and trustworthiness of data by preventing unauthorized modifications

Availability

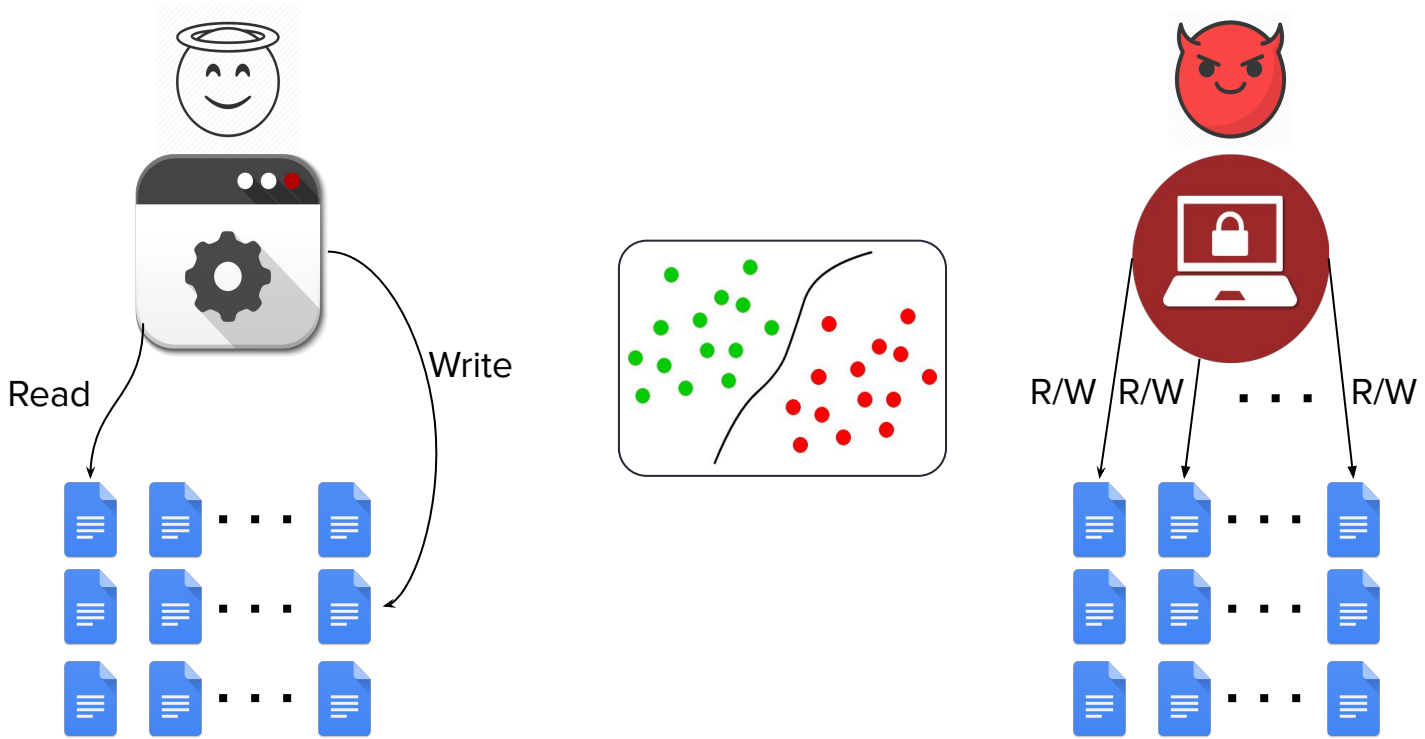


Ensuring that data and resources are available and accessible to authorized users when needed.

Malware Threats - Ransomware

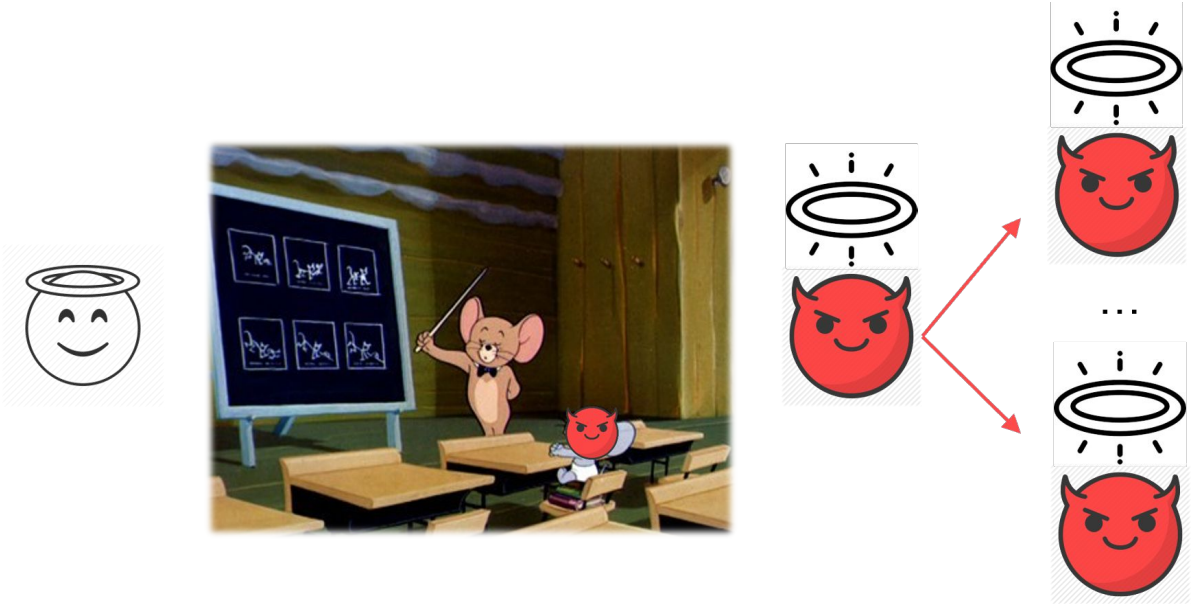
Type of malware from cryptovirology that threatens to publish the victim's data or perpetually block access to it unless a ransom is paid.

Ransomware detectors observing at process behaviour



Fooling behavioral-based detectors

Mimicry:



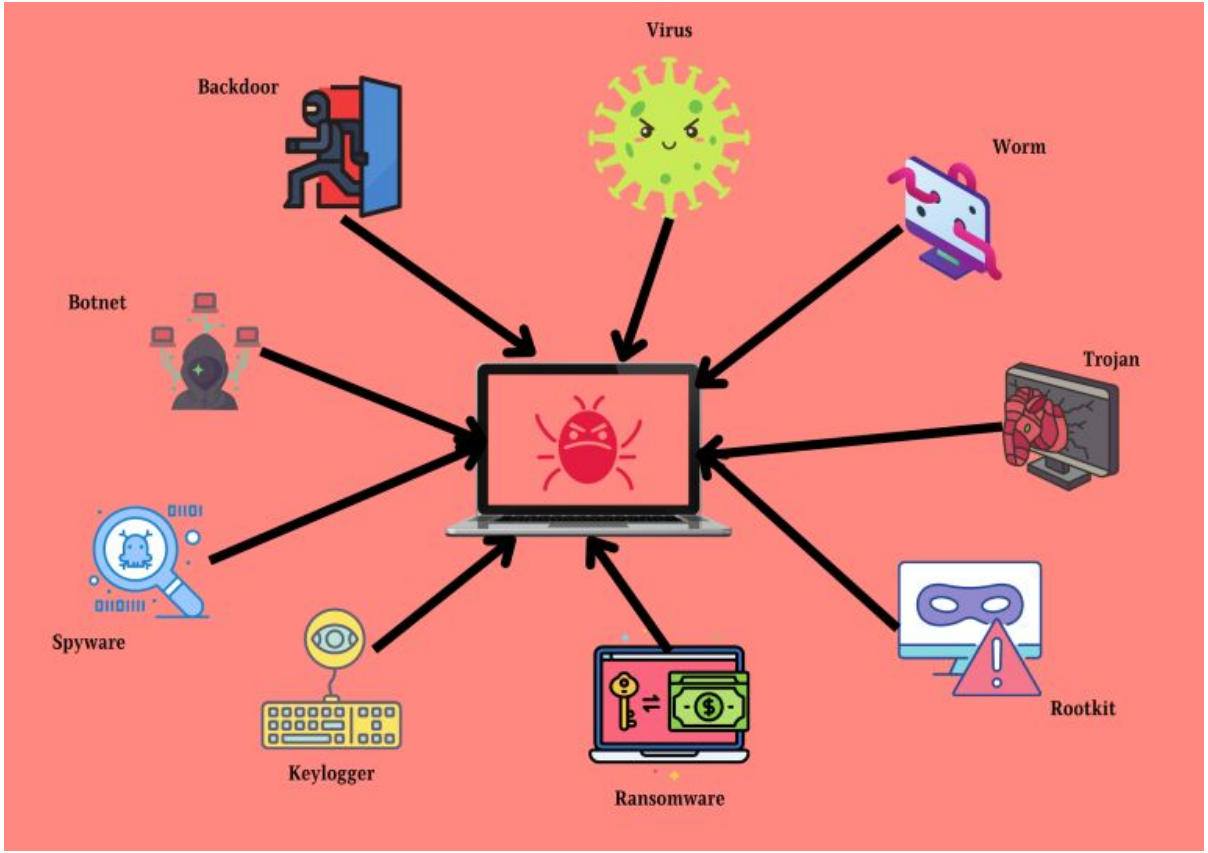
Outline for today

- Recap last lecture
- **Malware types**
- Emerging threats

Malware

Malicious software designed to infiltrate, damage, or disrupt computer systems, networks, or devices, often with harmful intent.

Malware Types



Malware Threats - knowing them

1. Identification and Detection
2. Prevention and mitigation
3. Remediation
4. Risk (assessment and management)
5. General user education
6. Adapting towards evolving threats

Malware Threats - knowing them

1. **Identification and Detection**

By recognizing the signs of malicious activity, such as unusual network activity or unauthorized system changes, security teams can respond promptly to mitigate potential risks.

Malware Threats - knowing them

1. Identification and Detection
2. **Prevention and mitigation**
3. Remediation
4. Risk (assessment and management)
5. General user education
6. Adapting towards evolving threats

Malware Threats - knowing them

2. Prevention and mitigation

Understanding how malware operates enables allows organizations to implement **proactive** measures to prevent and mitigate malware.

Employing security controls such as antivirus, firewalls, IDS, and secure configuration practices, organizations can reduce the likelihood of malware attacks and limit their impact.

Malware Threats - knowing them

1. Identification and Detection
2. Prevention and mitigation
3. **Remediation**
4. Risk (assessment and management)
5. General user education
6. Adapting towards evolving threats

Malware Threats - knowing them

3. Remediation

Knowing malware types and how they operate facilitates an effective response and remediation process.

Security teams can leverage their understanding of specific malware behaviors to contain infections, remove malicious code, and restore affected systems to a secure state.

Malware Threats - knowing them

1. Identification and Detection
2. Prevention and mitigation
3. Remediation
4. **Risk (assessment and management)**
5. General user education
6. Adapting towards evolving threats

Malware Threats - knowing them

4. Risk (assessment and management)

Identifying potential threats and vulnerabilities associated with different malware, entities can prioritize security measures and allocate resources to address the most significant risks.

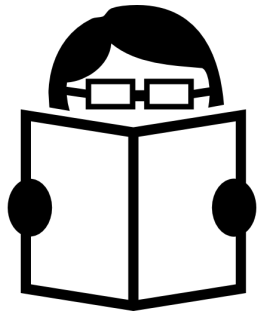
Malware Threats - knowing them

1. Identification and Detection
2. Prevention and mitigation
3. Remediation
4. Risk (assessment and management)
5. **General user education**
6. Adapting towards evolving threats

Malware Threats - knowing them

5. General user education

Teaching users to recognize common signs of malware attacks, such as suspicious emails or unexpected pop-up messages, organizations can assist them in taking proactive measures to protect themselves and the organization.



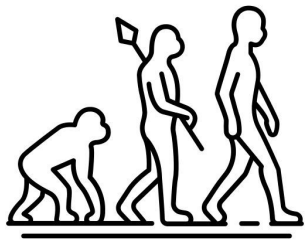
Malware Threats - knowing them

1. Identification and Detection
2. Prevention and mitigation
3. Remediation
4. Risk (assessment and management)
5. General user education
6. **Adapting towards evolving threats**

Malware Threats - knowing them

6. Adapting towards evolving threats

Continuously monitoring and analyzing emerging malware trends, organizations can enhance their cybersecurity position and try to stay one step ahead of adversaries.

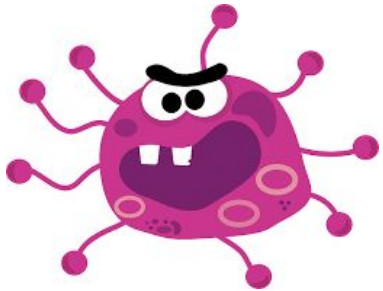


Malware Types - Common

- Viruses
- Worms
- Trojans
- Ransomware
- Spyware
- Adware
- Rootkits
- Scareware

Virus

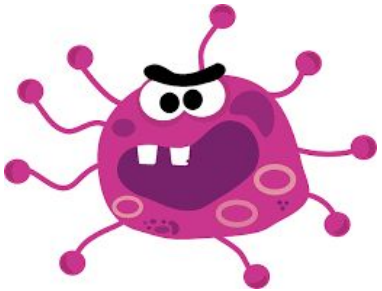
Program that can infect other programs by modifying them to include a, possibly evolved, version of itself, with intent to cause damage.



Virus

Program that can infect other programs by modifying them to include a, possibly evolved, version of itself, with intent to cause damage.

ILOVEYOU (discovered in 2000). The malware was delivered to millions of users as an email attachment with the subject line “ILOVEYOU.” Once opened, it spread to every contact in a user’s Microsoft Outlook address book and overwrote certain files (e.g., JPEG and MP3 files) from the hard drive.

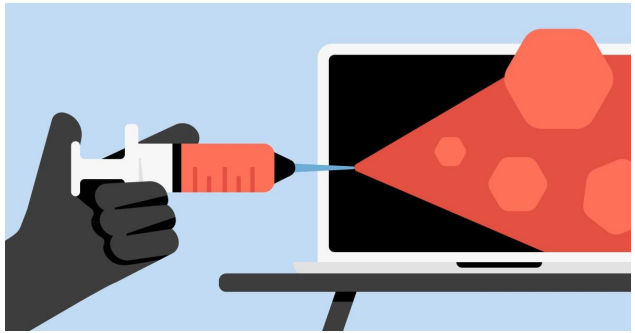


Trojan

Class of malware that appears to perform a desirable function but in fact performs undisclosed malicious functions that usually allow unauthorized access to the victim computer.

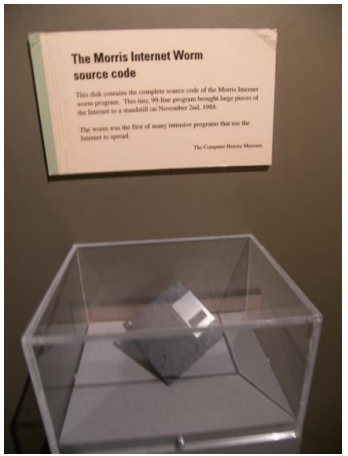
Rootkit

A rootkit is a program or a collection of malicious software tools that give an adversary remote access to and control over a computer.



Worms

A computer worm is a type of malware that can automatically propagate or self-replicate without human interaction, enabling its spread to other computers across a network.



Spyware

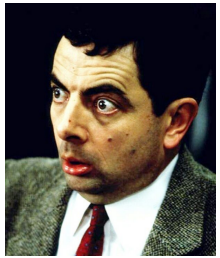
Malicious software that infects PCs and mobile devices in order to collect information on users and data on browsing habits, internet use, etc.

Adware

Adware is a type of malicious software that secretly installs itself on your device and displays unwanted advertisements and pop-ups. In some cases, adware can even track your online behavior and display personalized ads.

Scareware

Malware that scares people into visiting spoofed or infected websites or downloading malicious software (malware). Scareware can come in the form of pop-up ads that appear on a user's computer or spread through spam email attacks.



How malware is spread?

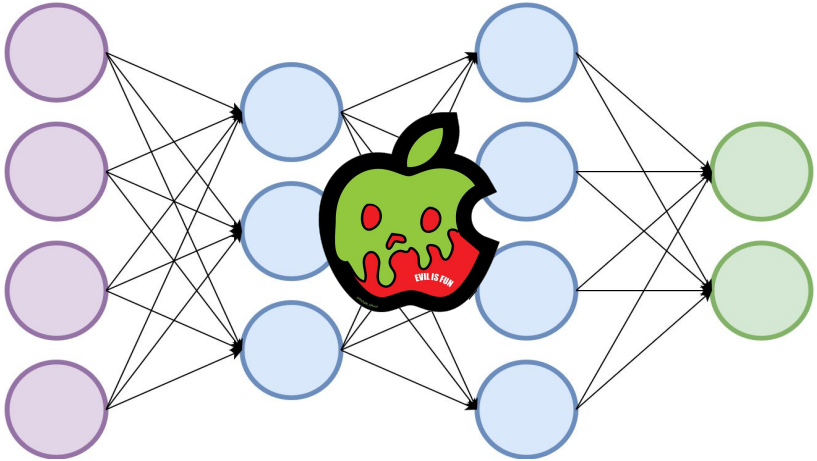
- Email attachments
- Internet downloads
- File sharing networks
- Removable media
- …

Outline for today

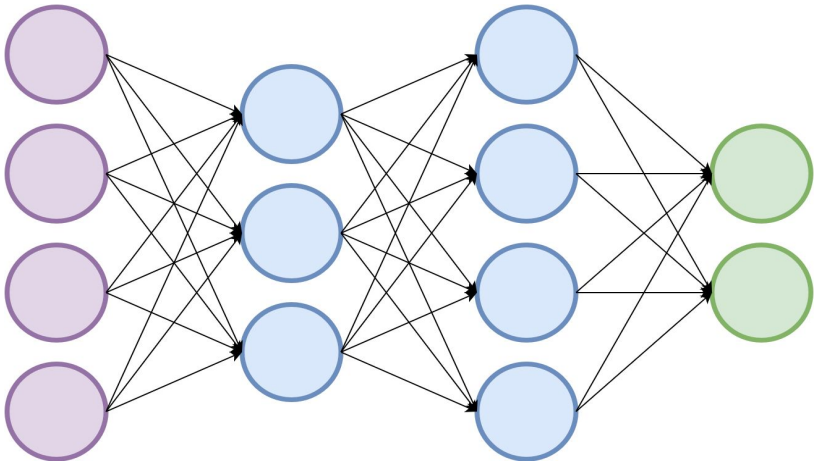
- Recap last lecture
- Malware types
- **Emerging threats**

How malware is spread again?

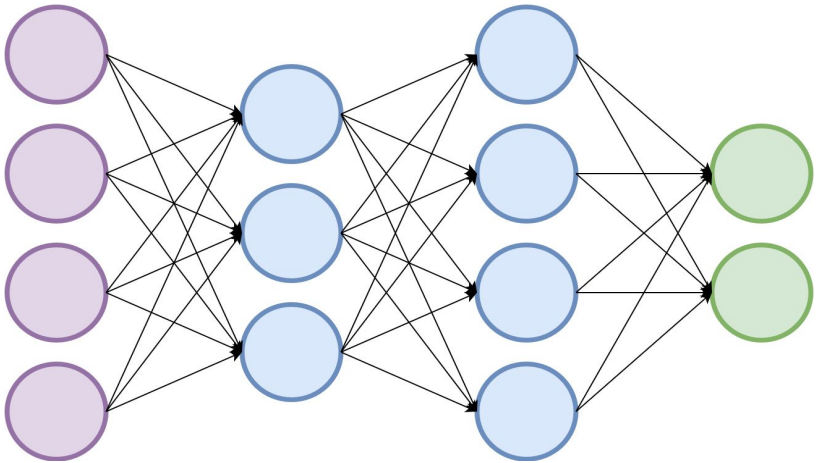
GIST - Malware can also hide inside a deep neural network



Preliminaries - Neural Network



A neural network is a computational model inspired by the structure and functioning of the human brain. It consists of interconnected nodes, or "neurons," organized in layers.



Information flows through the network, undergoing transformations at each layer, to perform tasks such as pattern recognition, classification, and prediction.

The (Deep) Neural Network ecosystem

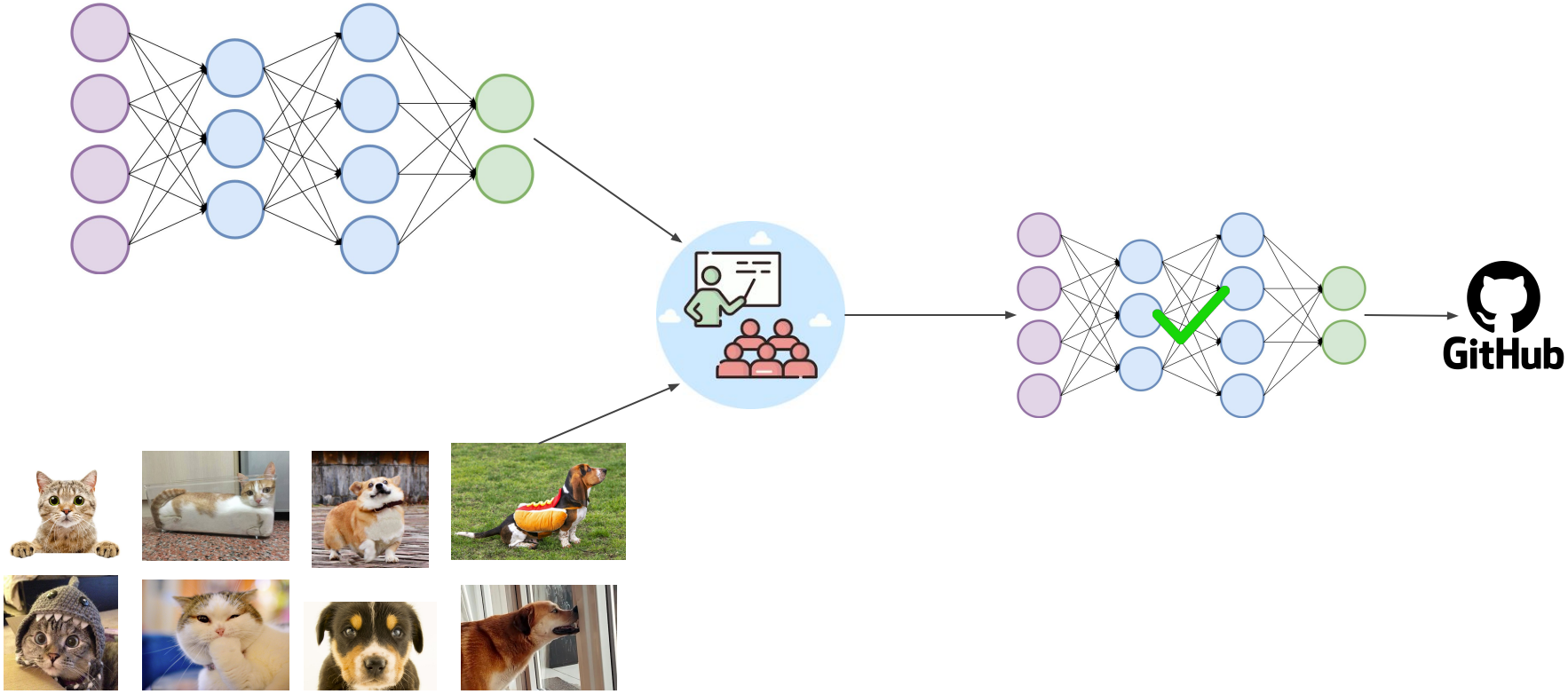
Hugging Face

kaggle

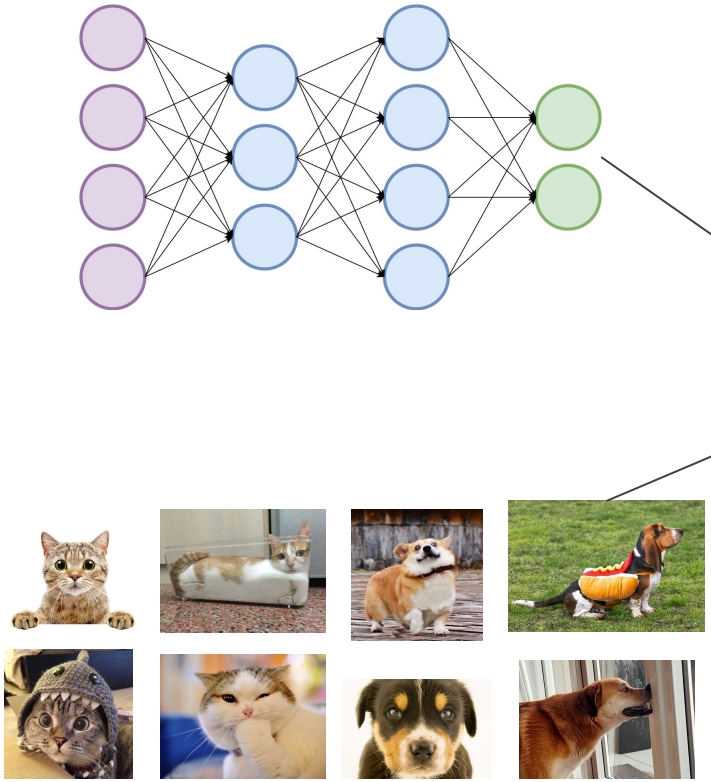
Model Zoo

GitHub

Do you trust your model?



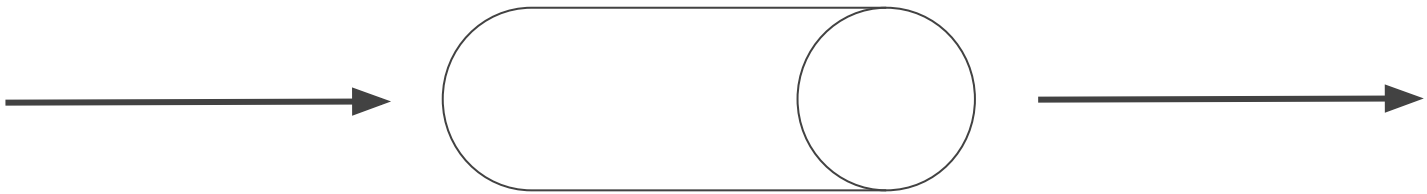
Do you trust your model?



Preliminaries - Digital (stealthy) communication

Digital Communication

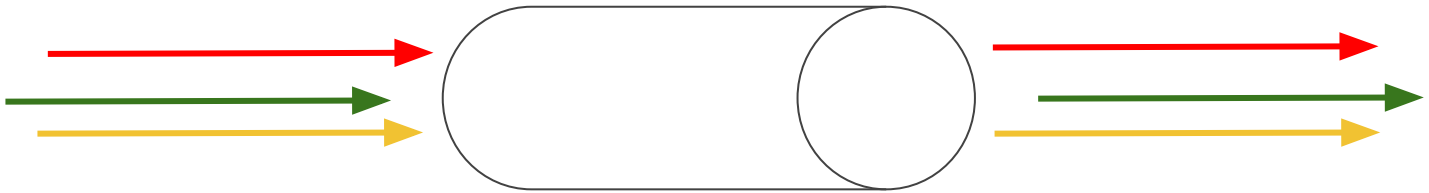
- Consists of the transfer of data or information using digital signals over a point-to-point channel.
- Data or information are digitally encoded as discrete signals, and are transferred through a communication channel.



Communication channel

What is CDMA?

- Channel coding technique by which a narrowband signal is spread into a wider bandwidth.
- Codes the data at a higher frequency by using pseudorandomly generated codes that the receiver knows.
- developed in the 1950s for military communications:
  - resist the enemy efforts to jam the communication channel.
  - hide the fact that the communication is taking place.



Communication channel

CDMA 101

Binary sequence: **[0, 1, 1]**

Assuming ω = 0:

PSK **[-1, 1, 1]**

Using Phase-Shift Keying:

- 0 → -cos(ωt)
- 1 → cos(ωt)
  - ω is the transmitting frequency
  - 0 ≤ t < T

CDMA 101

Binary sequence: **[0, 1, 1]**

PSK **[-1, 1, 1]**

Spreading code: **[-1, 1, -1, -1, 1]**

CDMA 101

Binary sequence: **[0, 1, 1]**

PSK **[-1, 1, 1]**

Spreading code: **[-1, 1, -1, -1, 1]**

Chip sequence: **[1, -1, 1, 1, -1, -1, 1, -1, -1, 1, -1, 1, -1, -1, 1]**

CDMA 101

Binary sequence: **[0, 1, 1]**

PSK **[-1, 1, 1]**

Spreading code: **[-1, 1, -1, -1, 1]**

Chip sequence: **[1, -1, 1, 1, -1, -1, 1, -1, -1, 1, -1, 1, -1, -1, 1]**

-5 **+5** **+5**

CDMA example communication - 2 users

Binary sequence: **[0, 1]**

PSK **[-1, 1]**

Spreading code: **[-1, -1, 1, 1]**

Chip sequence: **[1, 1, -1, -1, -1, -1 ,1 , 1]**

Binary sequence: **[0, 0]**

[-1, -1]

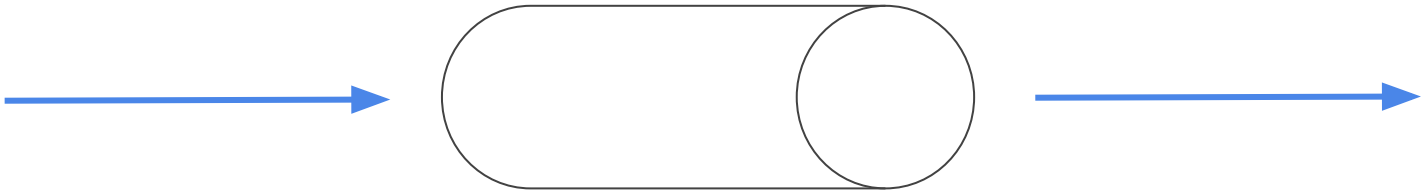
[1, -1, 1, -1]

[-1, 1, -1, 1, -1, 1, -1, 1]

CDMA example in communication

User #1 **[1, 1, -1, -1, -1, -1, 1, 1]** User #2 **[-1, 1, -1, 1, -1, 1, -1, 1]**

Combined signal: [0, 2, -2, 0, -2, 0, 0, 2] - two users worth of data



CDMA example - Decode user #1

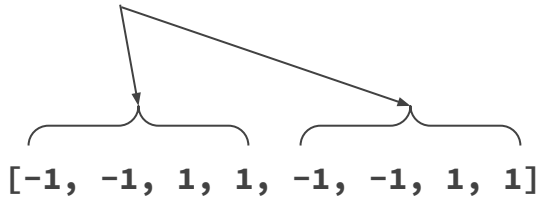
Combined signal: `[0, 2, -2, 0, -2, 0, 0, 2]` **- two users worth of data**

Spreading code: `[-1, -1, 1, 1]`

CDMA example - Decode user #1

Combined signal: **[0, 2, -2, 0, -2, 0, 0, 2] - two users worth of data**

Spreading code: **[-1, -1, 1, 1]**



CDMA example - Decode user #1

Combined signal: `[0, 2, -2, 0, -2, 0, 0, 2]` - two users worth of data

✖

`[-1, -1, 1, 1, -1, -1, 1, 1]`

`[0, -2, -2, 0, 2, 0, 0, 2]`

Add up Add up

CDMA example - Decode user #1

Combined signal: `[0, 2, -2, 0, -2, 0, 0, 2]` - two users worth of data

`[-1, -1, 1, 1, -1, -1, 1, 1]`

`[0, -2, -2, 0, 2, 0, 0, 2]`

−4 +4

CDMA example - Decode user #1

Combined signal: **[0, 2, -2, 0, -2, 0, 0, 2] - two users worth of data**

[-1, -1, 1, 1, -1, -1, 1, 1]

[0, -2, -2, 0, 2, 0, 0, 2]

_____ _____

−4 +4

What was the length of the spreading code?

CDMA example - Decode user #1

Combined signal: **[0, 2, -2, 0, -2, 0, 0, 2] - two users worth of data**

[-1, -1, 1, 1, -1, -1, 1, 1]

[0, -2, -2, 0, 2, 0, 0, 2]

We have to divide by 4

−4 +4

−1 +1

CDMA example - Decode user #1

Combined signal: **[0, 2, -2, 0, -2, 0, 0, 2] - two users worth of data**

✖

[-1, -1, 1, 1, -1, -1, 1, 1]

[0, -2, -2, 0, 2, 0, 0, 2]

−4 +4

−1 +1

Remember PSK mapping
-1 was 0
and
+1 was 1

CDMA example - Decode user #1

Combined signal: **[0, 2, -2, 0, -2, 0, 0, 2] - two users worth of data**

✖

[-1, -1, 1, 1, -1, -1, 1, 1]

[0, -2, -2, 0, 2, 0, 0, 2]

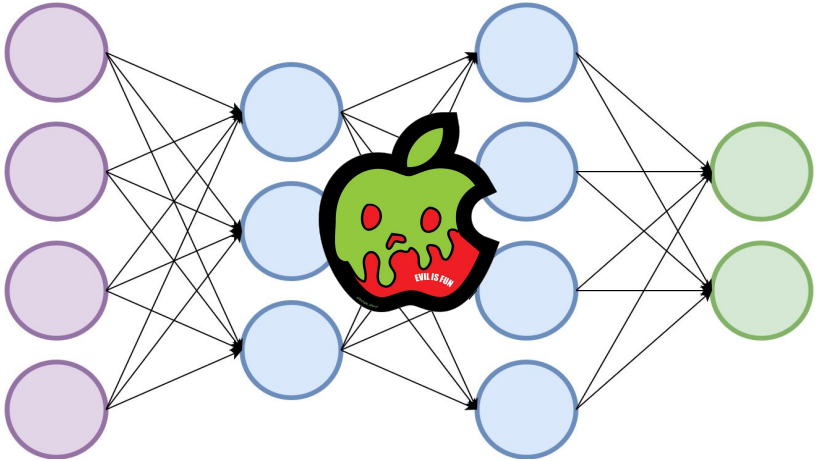
User #1 message was **01**

−4 +4

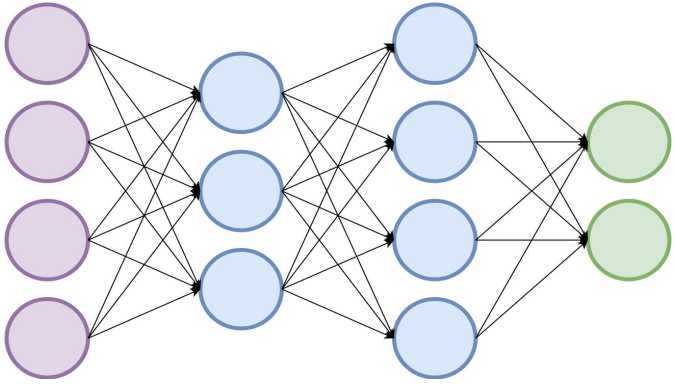
0 1

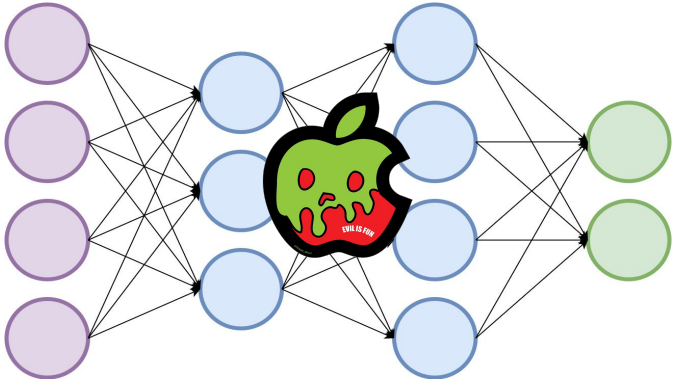
Why CDMA is stealthy?

- Spreading codes are pseudorandomly generated and are in the orders of **tens of thousands**


EVIL IS FUN

Malware embedding into DNNs: HowTo





Malware embedding into DNNs: HowTo

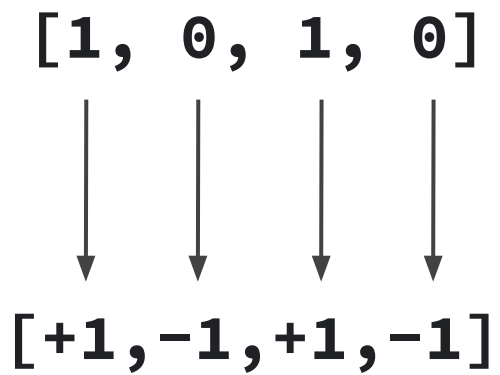
Malicious Payload -> **P** bits **b** = $[b_0, \ldots, b_{P-1}]$

[1, 0, 1, 0]

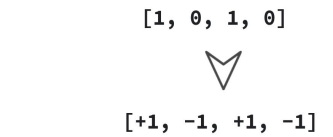
Malware embedding into DNNs: HowTo (cont.)

Malicious Payload -> **P** bits **b** = [b_0, …,b_{P-1}]

- Each bit is encoded as **±1**.



Malware embedding into DNNs: HowTo (cont.)



Malicious Payload -> **P** bits $\mathbf{b = [b_0, \ldots, b_{P-1}]}$

- The spreading code (c_i) of each bit is a vector of ±1 that is of the same length of vector W, namely R.

$$\mathbf{[+1,-1,+1,-1]}$$

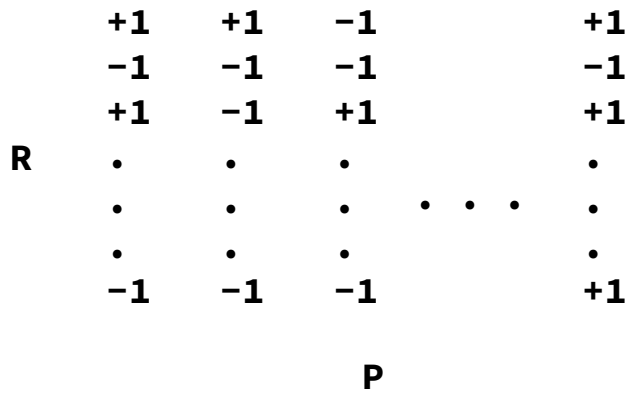
$$\mathbf{[+1, +1, -1, +1, \ldots, -1]}\ \mathbf{c_i}$$

R elements

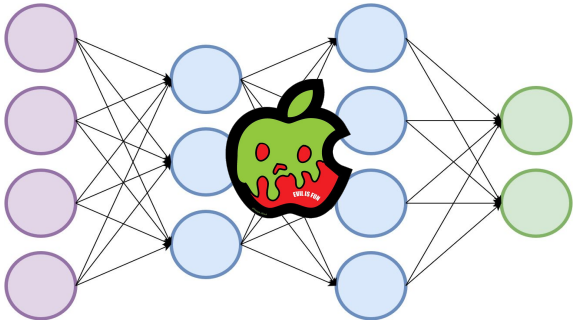
Malware embedding into DNNs: HowTo (cont.)

Malicious Payload -> **P** bits $\mathbf{b} = [b_0, \ldots, b_{P-1}]$

- **C** is an **R** by **P** matrix that collects all the codes.



Malware embedding into DNNs: HowTo (cont.)



$$W^{Malicious} = W + \gamma Cb$$

γ – gain factor to control the power of the signal

Payload Extraction

Each bit b_i of the malicious payload $\mathbf{b} = [b_0, \ldots, b_{P-1}]$ can be recovered by performing:

$$\hat{b}_i = c_i^T W^{Malicious}$$

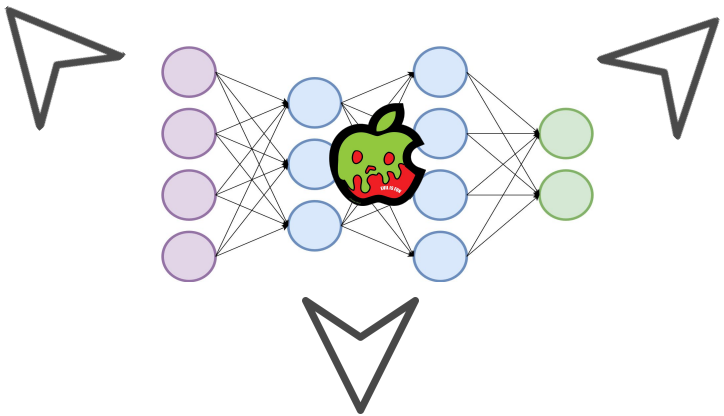
$$\downarrow$$

$$b_i = sign(\hat{b}_i)$$

How does it perform?

Stealthiness

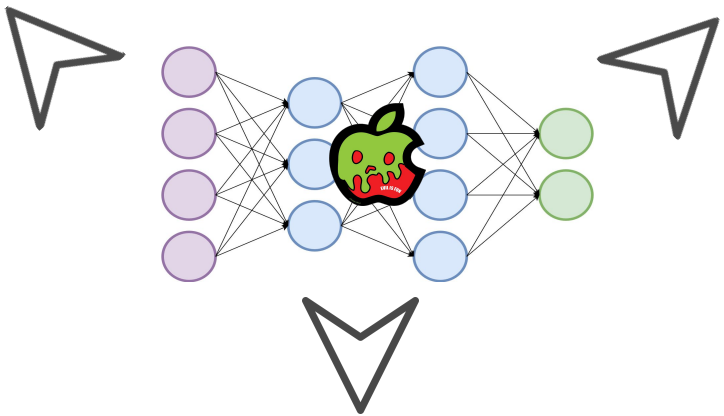
Impact on model performance



Robustness

Stealthiness

Stealthiness



Impact on model performance

Robustness

Stealthiness

1. Evaluation against anti-malware software.
2. Statistical analysis

Stealthiness

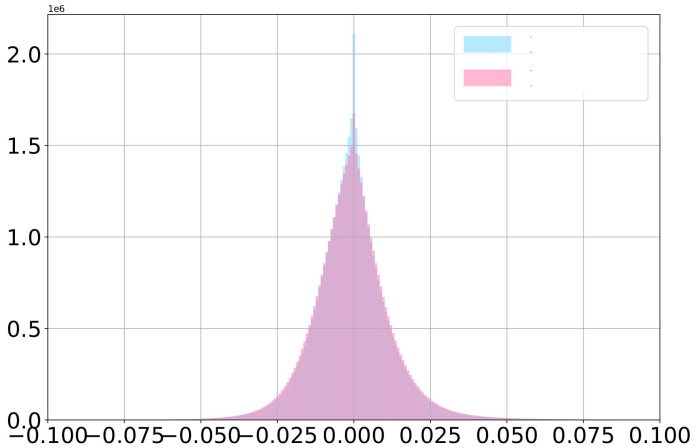
1. **Evaluation against anti-malware software.**
2. Statistical analysis

	Malware samples				
	Stuxnet	Destover	Bladabindi	Zeus	Kovter
Plain malware	89.19%	83.78%	72.97%	91.89%	62.16%
Stegomalware	0.00%	13.51%	8.11%	10.81%	5.41%
Malicious-DNN	**0.00%**	**0.00%**	**0.00%**	**0.00%**	**0.00%**

Detection rate on **Metadefender** metascan feature

Stealthiness

1. Evaluation against anti-malware software.
2. **Statistical analysis**

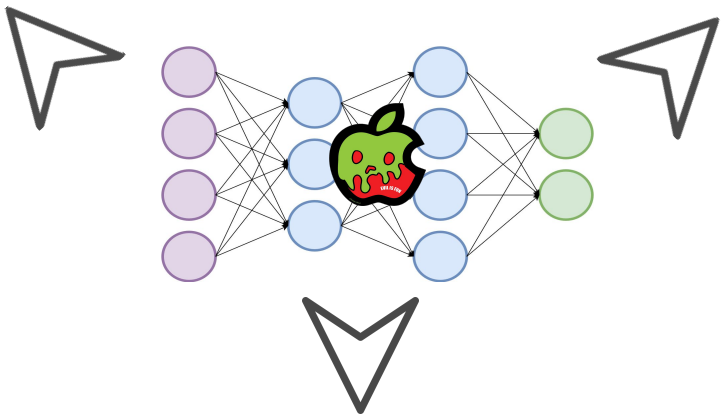


ResNet-101 architecture
Cerber malware

Comparison between the distribution of the weight parameters of the baseline regular model and the **one containing the malware.**

Stealthiness

Stealthiness

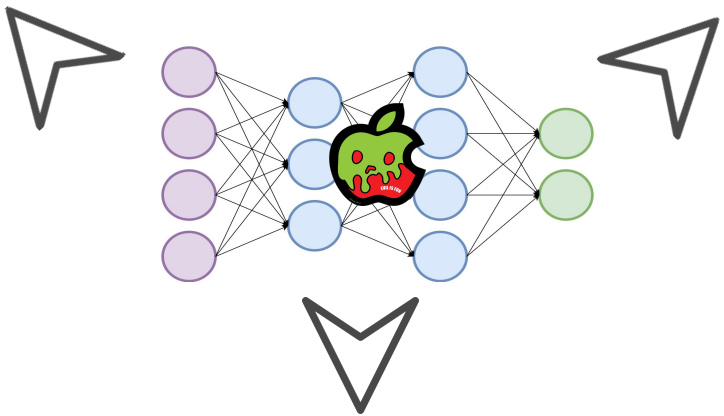


Impact on model performance

Robustness

Impact on model performance

Stealthiness



Impact on model performance

Robustness

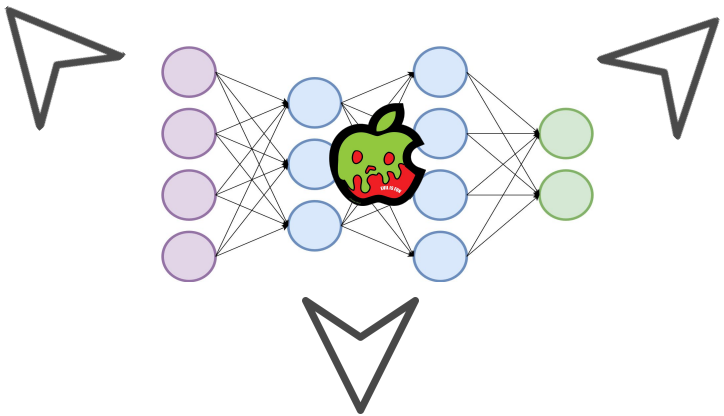
Impact on model performance

	DenseNet		ResNet50		ResNet101		VGG11		VGG16	
Malware	Bas.	Mal.	Bas.	Mal.	Bas.	Mal.	Bas.	Mal.	Bas.	Mal.
Stuxnet	62.13	61.22	75.69	75.34	76.96	76.87	70.13	70.09	73.37	73.34
Destover	62.13	52.36	75.69	74.89	76.96	76.79	70.13	70.05	73.37	73.28

Classification accuracy (%) on the ImageNet task before and after embedding different sized malware into different sized neural networks.

Impact on model performance

Stealthiness



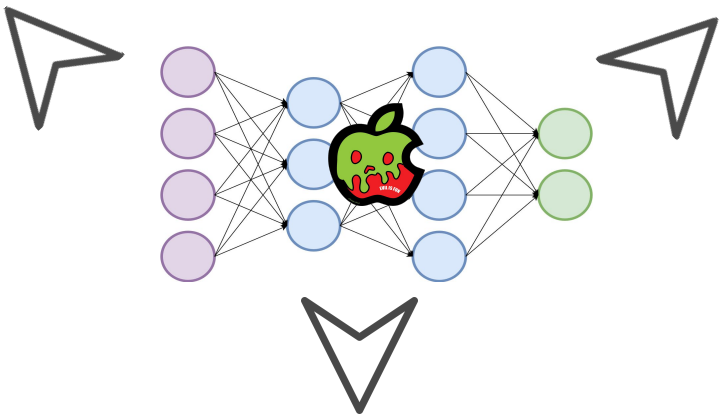
Impact on model performance

Robustness

Robustness

Stealthiness

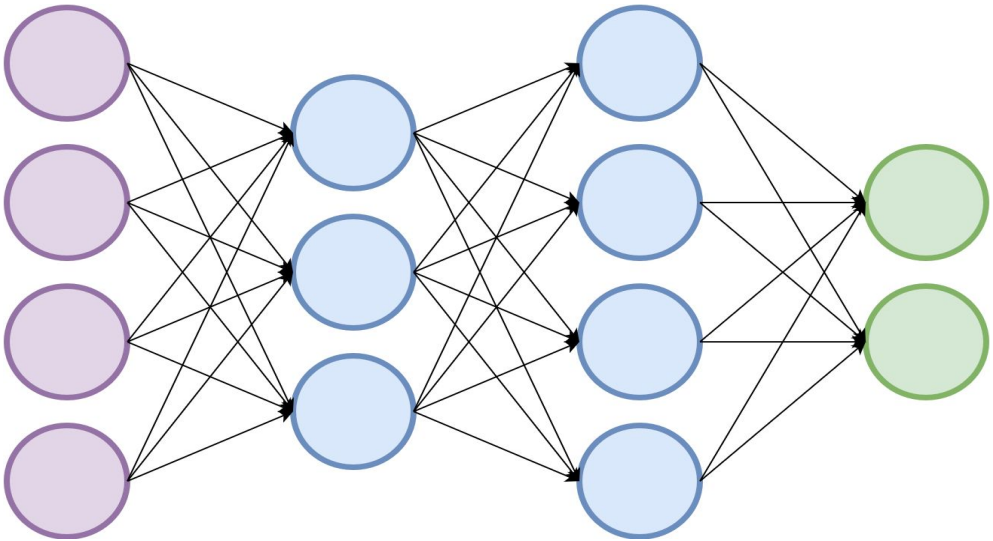
Impact on model performance



Robustness

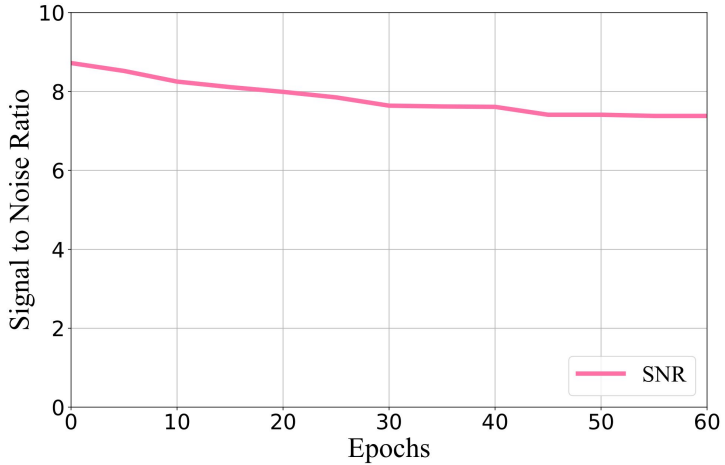
Robustness

1. Robustness towards fine-tuning
2. Robustness towards parameter pruning



Robustness

1. **Robustness towards fine-tuning**
2. Robustness towards parameter pruning



The change in signal-to-noise ratio of the malicious payload when the VGG11 architecture is trained on CIFAR10 and repurposed to solve the CIFAR100 task.

Robustness

1. Robustness towards fine-tuning
2. **Robustness towards parameter pruning**

Pruning Ratio	Does the payload survive?
25.00%	YES
50.00%	YES
75.00%	YES
90.00%	NO
99.00%	NO

Model being pruned - **ResNet50** containing **Stuxnet** malware payload embedded in the network

Why so robust?

Binary sequence: **[0, 1, 1]**

PSK **[-1, 1, 1]**

Spreading code: **[-1, 1, -1, -1, 1]**

Chip sequence: **[1, -1, 1, 1, -1, -1, 1, -1, -1, 1, -1, 1, -1, -1, 1]**

-5 **+5** **+5**


That's all Folks!

Reading Material

1. Common types of malware by Crowdstrike (web article).
2. Digital communication and CDMA: Link-1, Link-2 (research papers).
3. Introduction to Deep Neural Networks. (book chapter)
4. Hiding malware into Deep Neural Networks (research paper).