Chen Eliyahou | Alon Shats | Noam Simon | Mentor: Assaf Barak

Detecting Malicious Images using Machine Learning

General Description

Developing solution to detect malicious images attacks using Machine Learning algorithms

Web-Based Attacks

As introduction to the project we Implemented two web applications (one for XSS and one for CSRF) and illustrated the following web-based attacks:

XSS

The hacker takes advantage of the trust that a user has for a certain website.

A hacker injects a malicious client side script in a website. This script is added to cause some form of vulnerability to a victim.

CSRF

The hacker takes advantage of a website's trust for a certain user's browser.

A malicious attack is designed in such a way that a user sends malicious requests to the target website without having knowledge of the attack.

What Is Steganography?

Unlike encryption, where it's obvious that a message is being hidden, steganography hides data in plain view, inside a file such as a picture.
As far as images are concerned, to anyone who isn't aware that it contains hidden data, it looks like just a normal, innocent picture.

Stenography Image Creation

By LSB Technique

This technique changes the last few bits in a byte to encode a message

CIFAR-10 Database

The **CIFAR-10** database consists of **60000** **32x32** color images in **10** classes, with **6000** images per class

XGBoost

XGBoost is an implementation of gradient boosted decision trees designed for **speed** and **performance** that is dominative competitive machine learning.

The Dataset

Malicious images

30K images of the CIFAR-10 database was injected by 384B malicious JavaScript code. The forth bit of each pixel has changed similarly to LSB technique.

Clear images

The other 30K images of the CIFAR-10 database.

The Machine Leaning Model

Features : pixels of the image

Labels : 1 - malicious image and 0 - clear image

XGBoost model with depth of 256

Accuracy results **85.56%**

Future Product

Development of protection against attacks through machine learning

Check another types of learning

Adding option for predicting on different image formats and sizes, or even videos