Project Plan

AMAP: Automated Malware Analysis Platform

The Capstone Experience

Team Accenture

Teng Xu, Hefei
Griffin Metevia
Julian Ellis
Andrew Mitchell
Sam Kling

Department of Computer Science and Engineering
Michigan State University

Spring 2018
Functional Specifications

• Automate the process of practically analyzing malware samples
• Perform analysis on a large volume of malware samples
• Focus on basic static and basic dynamic analysis
• Record results of analysis and display information to a dashboard
Design Specifications

• Series of modules used to analysis each malware sample
• Status dashboard displays information on the state of the AMAP system
• Malware search page allows users to see specific malware sample information
• Wizard-style UI to add, edit, or remove modules
Screen Mockup: Malware Search

Malware Sample Search

- Sample ID
- Strong Password
  - ID: 000000
  - Custom
  - 14
- File Test
  - ID: 000000
  - Custom
  - 0
- Slack Install
  - ID: 000000
  - Custom
  - Required Software
  - 1
- Custom Foo
  - ID: 000000
  - Custom
  - 32
- api-manual
  - ID: 000000
  - Patch
  - 22

File Information:
- File Name: 040d71e56512ao786bcs.xls
- File Size: 77824 bytes
- File Mime: application/vnd.ms-excel
- File Type: Composite Document File V2 Document
- MD5: 040d71e56512ao786b
- SHA1: a4jn43kn34mk34mk343mk343mk343mk
- Notes:
Screen Mockup: AMAP Dashboard

AMAP: Automated Malware Analysis Platform

Cost Per Hour

- AWS Storage: 71%
- AWS Containers: 9%

$55

Malware Type

- Windows: 69%
- Web: 15%
- Mac: 9%
- Linux: 7%

Container Status

- Starting Up: 4%
- Running: 18%
- Idle: 3%

Malware Processed Per Hour

New Samples Per Hour
Screen Mockup: AMAP Dashboard

Cost Per Hour

AWS Storage: 71%
AWS Containers: 9%

$55

Malware Type

Windows: 69%
Web: 15%
Mac: 9%
Linux: 7%
Screen Mockup: AMAP Dashboard

Malware Processed Per Hour

Container Status

- Starting Up: 4%
- Running: 18%
- Idle: 2%
Technical Specifications

• Basic static analysis
  ▪ Provides information about functionality
  ▪ Produce simple network signatures
  ▪ Ineffective against sophisticated malware

• Basic dynamic analysis
  ▪ Observe malware behavior after executing on system such as encrypting files or changing file names
  ▪ Takes place in a controlled environment such as VM or sandbox
System Architecture

AMAP Modules

iDefense
Part of Accenture Security

Malware Database

Data Server

AMAP Dashboard

Python

mongoDB

django

AWS

HTML

CSS
System Components

• Software Platforms / Technologies
  ▪ iDefense IntelGraph API
  ▪ iDefense Malware Repository
  ▪ MongoDB hosted on AWS
  ▪ Python/Django -- PyCharm
  ▪ HTML, CSS, JavaScript
Risks

• Processing a large quantity of samples
  ▪ System needs to handle an average of 300 thousand per day
  ▪ Using multithreading to allow many modules to be run concurrently

• Categorizing malware based on type
  ▪ Malware must be classified based on detection signatures, byte patterns, and other information
  ▪ Undergoing training from the client to learn how to categorize malware based on these criteria

• Getting information from dynamic analysis
  ▪ Malware samples are executed in a VM or sandbox environment and information about their effects must be recorded
  ▪ The client has extensive knowledge about how to perform this method of malware analysis

• Determining when a sample is finished processing
  ▪ Malware analysis can sometimes produce as a result encoded payloads that require further analysis
  ▪ Client can provide information about when this situation occurs and small scale testing can be used to determine what kinds of samples might cause this
Questions?