Malware Lab4

Chen Zhuge IVCM132116

(This report was regenerated based on historical materials on 16 June, 2020.)

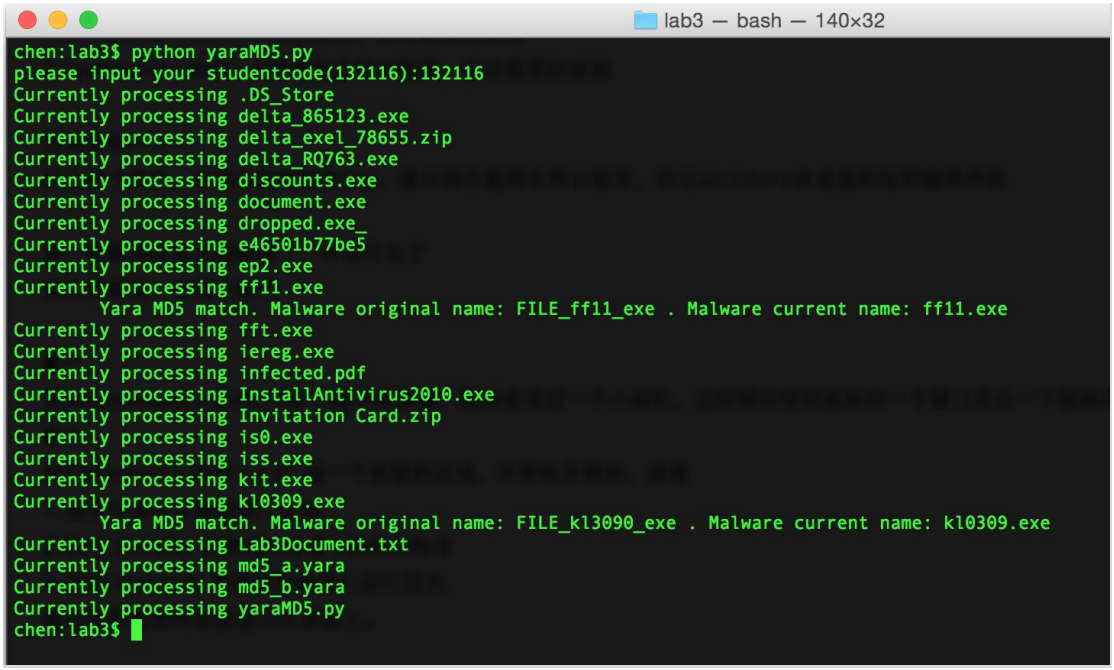
My work is largely based on the scripts on http://hooked-on-mnemonics.blogspot.com/2011/12/yara-md5.html. I changed the original scripts a bit to suit the lab's requirement.

Based on the MD5 scanning, the malwares could still be found if their names were changed intentionally. The following logs prove that Yara can still find ff11.exe , after its name was changed to ChangedName.txt .

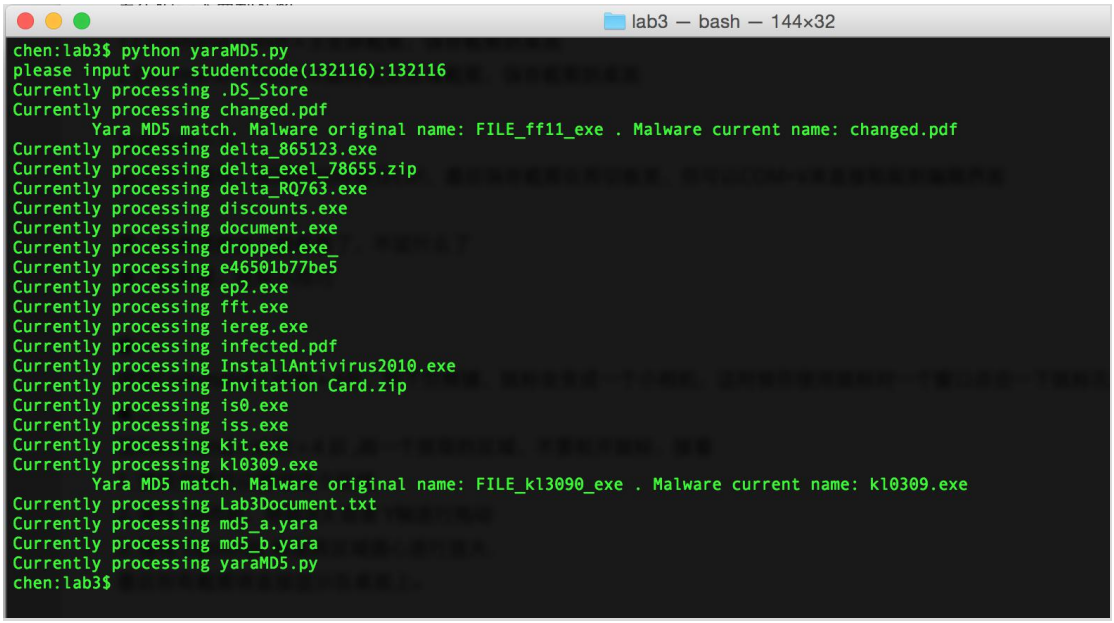
```
chen:yaratest$ python yaraMD5.py
    Yara MD5 match. Malware original name: FILE_ff11_exe . Malware current name: ff11.exe
    Yara MD5 match. Malware original name: FILE_kl3090_exe . Malware current name: kl0309.exe
chen:yaratest$ python yaraMD5.py
    Yara MD5 match. Malware original name: FILE_ff11_exe . Malware current name: ChangeName.txt
    Yara MD5 match. Malware original name: FILE_kl3090_exe . Malware current name: kl0309.exe
```

Appendix:

Before changing the file name:



After changing the file name:



Source code:

```
import hashlib
import sys
import imp

import yara
from StringIO import StringIO
from os import listdir

def MD5(d):
    # d = buffer of the read file
    # This function hashes the buffer
    # source: http://stackoverflow.com/q/5853830
    if type(d) is str:
      d = StringIO(d)
    md5 = hashlib.md5()
    while True:
        data = d.read(128)
        if not data:
            break
        md5.update(data)
    return md5.hexdigest()


def yaraScan(d, n, st):
    # d = buffer of the read file
    # Scans SWF using Yara
    # test if yara module is installed
    # if not Yara can be downloaded from http://code.google.com/p/yara-project/
    if st != 132116:
        print "Warning: Studentcode doesn't match. Yara will run but result will not be as expected."
    try:
        imp.find_module('yara')
        import yara
    except ImportError:
        print '\t[ERROR] Yara module not installed - aborting scan'
        return
    # test for yara compile errors
    try:
        r = yara.compile(r'md5_a.yara', externals={"var1":st})
    except:
        pass
        print '\t[ERROR] Yara compile error - aborting scan'
        return
    # get matches
    m = r.match(data=d)
    # print matches
    for X in m:
        print '\tYara MD5 match. Malware original name:', X, '. Malware current name:', n

    try:
        r = yara.compile(r'md5_b.yara', externals={"var1":st})
    except:
        pass
        print '\t[ERROR] Yara compile error - aborting scan'
        return
    # get matches
    m = r.match(data=d)
    # print matches
    for X in m:
        print '\tYara MD5 match. Malware original name:', X, '. Malware current name:', n

def main():
    studentcode=input("please input your studentcode(132116):")
    for each_file in listdir("/Users/chen/study/currentcourses/Malware2/lab3"):
        try:
        #f = open(sys.argv[len(sys.argv)-1],'rb+')
            f = open(each_file)
            na = each_file
            print "Currently processing", na
        except Exception:
            print '[ERROR] File can not be opended/accessed'
            #return
        yaraScan(MD5(f), na, studentcode)

if __name__ == '__main__':
   main()
```