Advanced Topics in Malware Analysis

Brendan Saltaformaggio, PhD
Assistant Professor
School of Electrical and Computer Engineering

Course Introduction

Georgia Tech

1

Today's Agenda

- Introduction to the Class and Syllabus
- Basics of x86 Assembly Language
- Basics of 16/32/64bit assembly code and registers
- Executable files
- Difference between Intel and AT&T syntax

Georgia Tech

2

Before We Begin

Professor Brendan Saltaformaggio

"Salt" – "uh" – "for" – "mah" – "gee" – "oh"

Assistant Professor, School of ECE and CS

Director, Cyber Forensics Innovation Laboratory (CyFI Lab)

Research Interests

- Cyber Forensics & Computer Systems Security
- Binary Analysis & Instrumentation
- Vetting Of Untrusted Software
- Memory Image Forensics
- Mobile/IoT Security

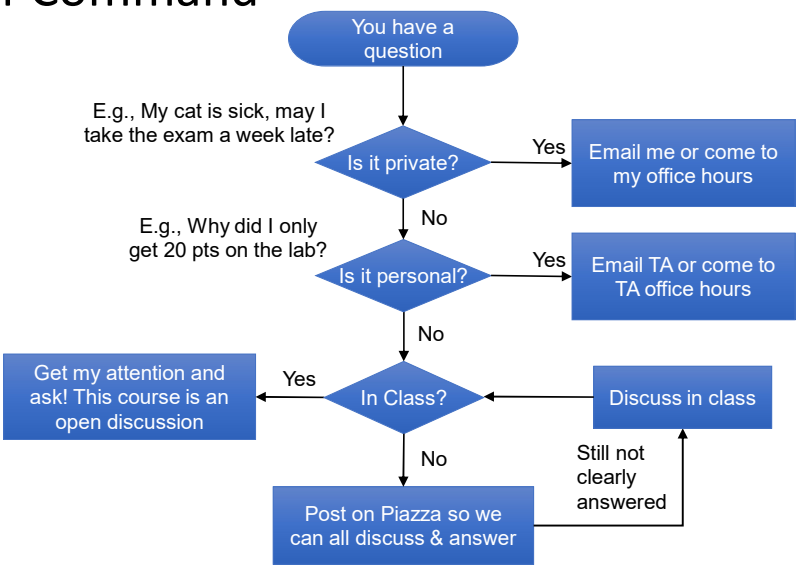
brendan@ece.gatech.edu

http://saltaformaggio.ece.gatech.edu

Georgia Tech

3

Chain of Command



Georgia Tech

4

The Course in One Slide

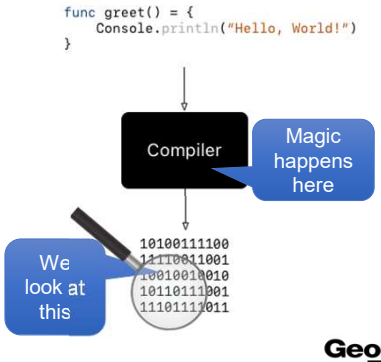
- **Advanced approaches for detecting vulnerabilities/malware within binary software**

Software security is a rapidly changing field!

- NO textbook can keep up
- Instead, we will study published papers from top academic venues

There are a few principle techniques for software analysis

- We will cover these "building-blocks" in the lecture
- You will apply this knowledge in labs



Georgia Tech

5

Warning!

- This is **a time consuming and challenging** course! This is not a "requirement-filler"!
- Malware analysis is "spy versus spy"
  - You are trying to reverse engineer what someone else (the malware author) is trying to hide from you!
- This course will require a significant amount of work!
  - Each assignment will require very careful time allocation to complete by the deadline
  - Start each assignment **as soon as it opens!**
  - Malware analysis will become your new hobby!
- If you do not **LOVE** malware analysis and software security, it will be very hard!
  - Lots and lots of assembly language and C
  - You can learn assembly language as we go

Georgia Tech

6

Testimonials

- “The class doesn't aim to mollycoddle you, and I appreciated that. It encouraged you to aspire for more and push your limits. Only in that extreme can one learn so much so well. “
- “The labs were long, and incredibly time consuming, but nothing we weren't fairly warned about. ”
- “One of the most rewarding and challenging courses I have taken at Georgia Tech.”
- “The amount of sleep I lost over this class was enormous, but we were warned so I can't complain about it. This course was great.”
- “Professor Brendan is a boss.”
- “Great professor great course Would malware again.”

Georgia Tech

7

Course Agenda (Two Halves)

“In-Class” Half Of The Course

- Binary program analysis principles (building blocks of this research field)
- Traditional lecture format
- You will complete 6 binary program analysis labs out of class
  - 4 will be static analysis with Ghidra
  - https://ghidra-sre.org
  - 2 will be dynamic analysis with Pin
  - https://software.intel.com/en-us/articles/pin-a-dynamic-binary-instrumentation-tool
  - More on these tools is later videos…
  - Each lab will require careful time allocation to complete on time!!
    - 1 or 2 week deadlines

1st HALF
OF
THE COURSE

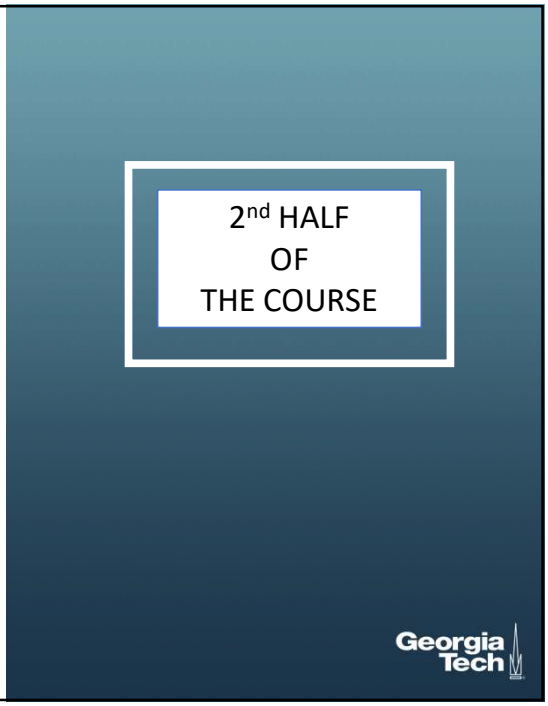
Georgia Tech

8

Course Agenda (Cont.)

"Research" Half of the Course

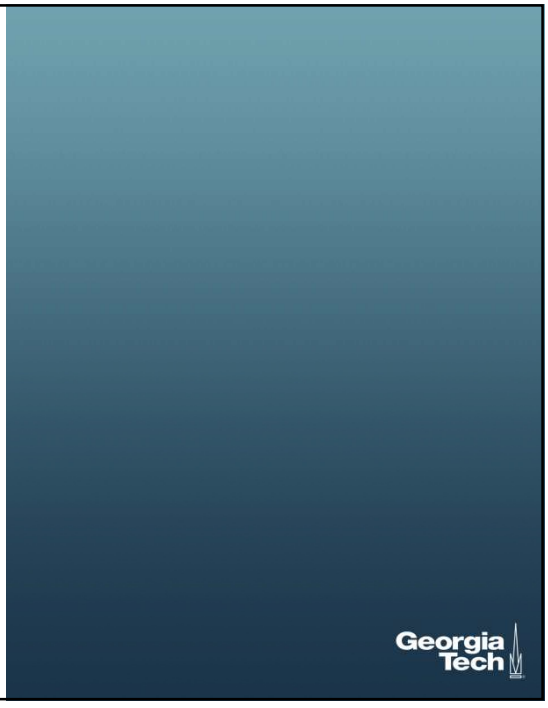
- How to conduct cutting-edge research in software security and cyber forensics
- Study published research papers
  - The building blocks we learn are applied in nearly all malware analysis research.
  - We will learn the limitations of these techniques and see how researchers are trying to solve them (or failing to solve them?).
  - Pay attention to the following in each paper:
    1. The Problem
    2. Previous Solutions/Techniques
    3. Novel Solution Presented In This Paper
    4. Limitations Of Their Approach
    5. Future Research Opportunities



9

Grading

- No Midterm, No Final Exam
- 6 Labs, 15% each
  - Grade based on the results produced by your tool
  - For some labs, we may schedule demos during office hours
- 10% for class participation
  - Piazza will become your best friend!
  - Help out others on there!
- Small extra credit assignments are likely to be announced in class



10

Zero-Tolerance Cheating Policy

- Labs are individual or teams of 2
  - Please discuss ideas with other students/teams (class participation points ☺)
  - DO NOT share code (that includes comments in code!)
- I reserve the right to use MOSS to detect cases of substantial overlap
  - http://theory.stanford.edu/~aiken/moss/
- Zero tolerance towards violation of the GT honor code
  - http://www.honor.gatech.edu/
- If you are caught cheating:
  Zero on lab assignment + One grade drop + Report to dean (Academic Warning in file)

Georgia Tech

11

Our Goal This Semester

- Learn and apply the fundamental principles of dissecting malware, vulnerability finding/defense, and cyber attack triage
- Become aware of limitations of existing defense mechanisms and how to avoid them
- Read cutting-edge research publications on these topics
- Engage in critical discussion around key research topics in software security and forensics
- Propose solutions to open-ended research problems
  - Projects which align with your thesis research are encouraged as long as it still has an interesting security/forensics component
  - There is ample scope to publish in this area: If the results from your course project look promising, we can write a paper on it and I will fund your travel to go present it

Georgia Tech

12

Programming Requirements

This course requires heavy programming

- It is a 3-credit course but can feel like a 4-5 credit course
- I said this before: Each project will require careful time allocation to complete on time!

You MUST be proficient in C

- You will be happier if you know some python and Assembler
  - It is ok if you do not
  - Everyone will be masters of them after this course ☺

And with that…

- On to an introduction to Assembly Language!

Georgia Tech

13