CASE WESTERN RESERVE UNIVERSITY EST. 1826
think beyond the possible

EECS 349/444 – Fall 2019

Computer Security

Prof. Fanny Ye

yanfang.ye@case.edu

Gain Deep Insights into the Online Underground Ecosystem

Group Project (40%)

- **Group project:** In the project, you are required to use cutting-edge techniques to solve the proposed cybersecurity research problems.
  - ✓ 3-4 students per group
  - ✓ Select a seed idea for your group project
  - ✓ Fully motivate the problem and survey related work (10%)
  - ✓ Project preparation (e.g., data collection, annotation, preprocessing, surveys, etc.) (20%)
  - ✓ Develop your own solutions – substantial novel technique development and implementation (20%)
  - ✓ A thorough empirical evaluation and comparing with baseline methods (15%)
  - ✓ A fully developed project report (20%): **You should NOT copy anything from anywhere!!**

    EECS 349: 8-page in ACM Master article template for LaTeX

    EECS 444: 12-page in ACM Master article template for LaTeX

    https://www.acm.org/publications/proceedings-template

    (Note: if your team is a mix of EECS 349 & 444, then your team needs to submit 12-page report)
  - ✓ Project presentation (15%): 12-min presentation + 3-min Q/A

T1: Gain Deep Insights into the Online Underground Ecosystem

❑ Motivation and Background

Nowadays, many sophisticated underground markets (e.g., underground forums, social media groups) have emerged in the cyberspace, where cybercriminals exchange information with fellow criminals on abusive tactics and engage in the sale of illicit goods and services. The function for these markets is not only for social contact within users, but also to support criminal activities, such as buying or selling crimeware such as exploits and CaaS such as hacking services. In order to allow law enforcement communities to devise effective disruptive strategies, there's an urgent need for novel techniques and tools to gain valuable insights into the online underground ecosystem.

- **Crimeware** is any computer program designed expressly to facilitate illegal activities online, such as malware (e.g., Trojans, ransomware) and crypters that encrypt malware making it undetectable to security programs;
- **Crimeware-as-a-Service (CaaS)** is a do-it-for-me service such as hacking, black SEO and DDoS attack.

T1: Gain Deep Insights into the Online Underground Ecosystem

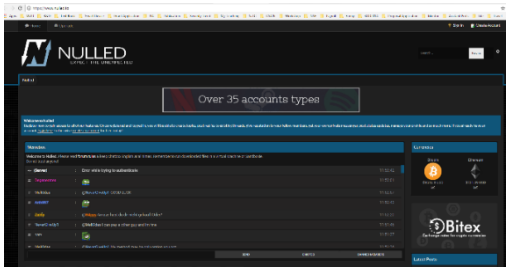
❑ Motivation and Background

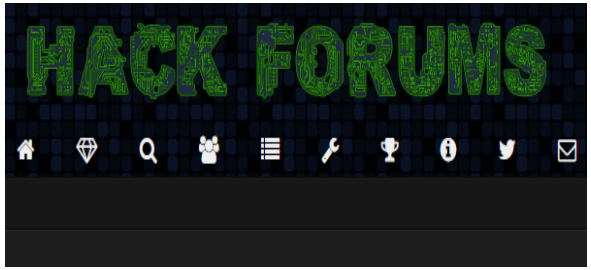
Nowadays, many sophisticated underground markets (e.g., underground forums, social media groups) have emerged in the cyberspace, where cybercriminals exchange information with fellow criminals on abusive tactics and engage in the sale of illicit goods and services. The function for these markets is not only for social contact within users, but also to support criminal activities, such as buying or selling crimeware such as exploits and CaaS such as hacking services. In order to allow law enforcement communities to devise effective disruptive strategies, there's an urgent need for novel techniques and tools to gain valuable insights into the online underground ecosystem.

- The emerging online underground markets have enabled cybercriminals to realize considerable profits. For example, the estimated annual revenue for an individual credit card steal organization was $300 millions; it's also revealed that a group of cybercriminals profited $864 millions per year by renting out the DDoS attacks.

T1: Gain Deep Insights into the Online Underground Ecosystem

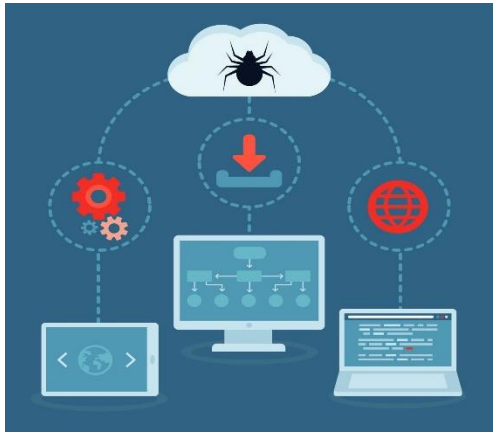
- To gain deep insights into online underground markets and better understand the cybercrime ecosystem, in this project, you are asked to first find at least five active underground markets, such as underground forums (e.g., Nulled: https://www.nulled.to/#!Marketplace, Hack Forums: https://hackforums.net/forumdisplay.php?fid=107).You need to describe and summarize how you find these underground markets and explain why these markets are significant for gaining insights into the online underground ecosystem.





T1: Gain Deep Insights into the Online Underground Ecosystem

1. Select one online underground market you explore, based on which each group needs to focus on **one particular kind of crimeware** (e.g., exploit, botnet) **and one kind of CaaS** (e.g., malware attack, hacking service) traded in the market. For each type of crimeware (denoted as P1) or CaaS (denoted as S1), you need to first develop your own solutions/tools to collect a number of threads (>50 threads for P1 and >50 threads for S1) and their related comments (>30 comments/thread) for further analysis (see steps 2-3). Describe and summarize how you collect the data. *Note that you only collect and analyze those threads whose comments > 30.*



T1: Gain Deep Insights into the Online Underground Ecosystem

2. Base on the collected data, you are asked to develop your own solutions/tools to analyze:

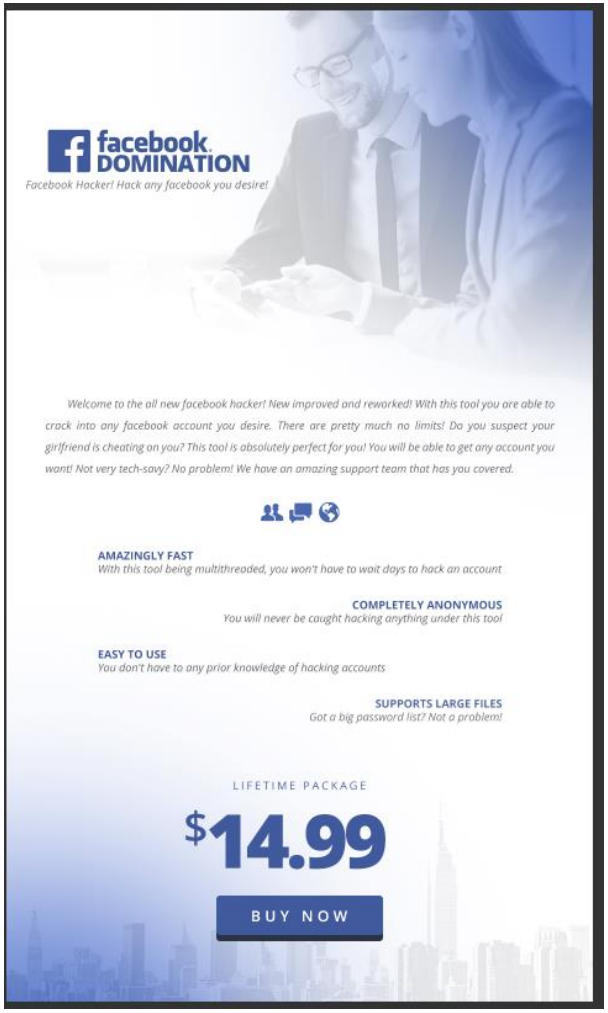
(1) Crimeware/CaaS trading **threads**:
extract use name and profile of vendor, product/service name of each thread, price, payment method, # of comments (i.e., replies), # of reviews;

(2) **Comments**: classify each comment

- username and profile of commenter
- trading [Yes/No/Uncertain]
- contracted customer [Yes/No]
- review [Positive/Negative/Neutral]
- Q&A [Yes/No]
- other.

You need to submit the analysis results using the required template shared in Canvas. Your analysis/annotation will be validated by cross-validation during grading.

https://hackforums.net/showthread.php?tid=5873970



(1) Crimeware/CaaS trading **threads**:

extract username and profile of vendor, product/service name of each thread, price, payment method, # of comments (i.e., replies), # of reviews;

Example

https://hackforums.net/showthread.php?tid=5873970

Underground Market: Hack Forums

Object: Threads

Link: https://hackforums.net/showthread.php?tid=5873970

Username: Mysterium

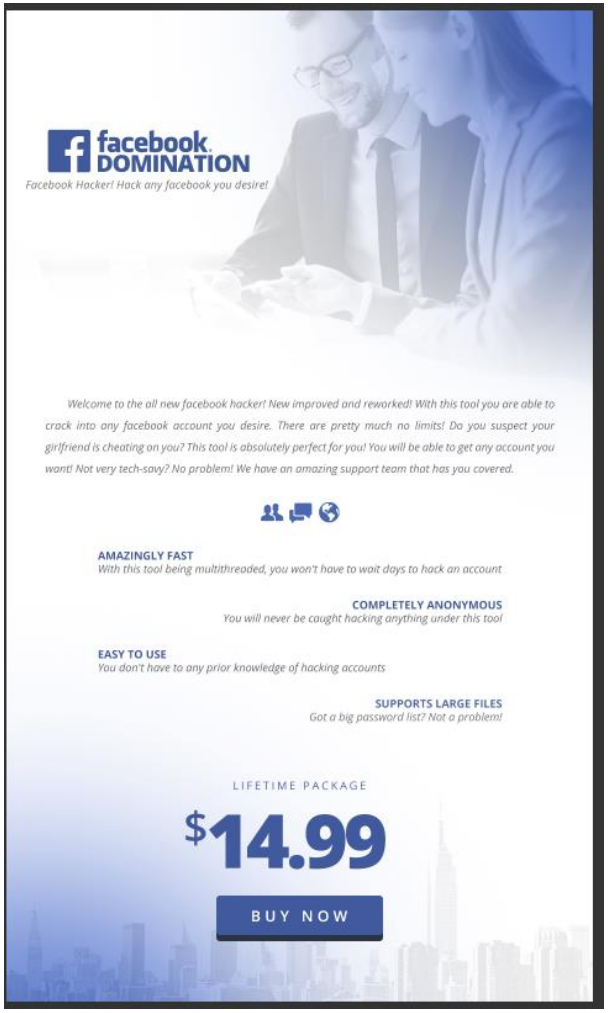
Profile-posts: 2,025

Profile-XXX: XXX

Product/service name: FACEBOOK HACKING TOOL

Price: $14.99
Unit: lifetime

Payment method: PayPal/Bitcoin/Ethereum



(1) Crimeware/CaaS trading **threads**:

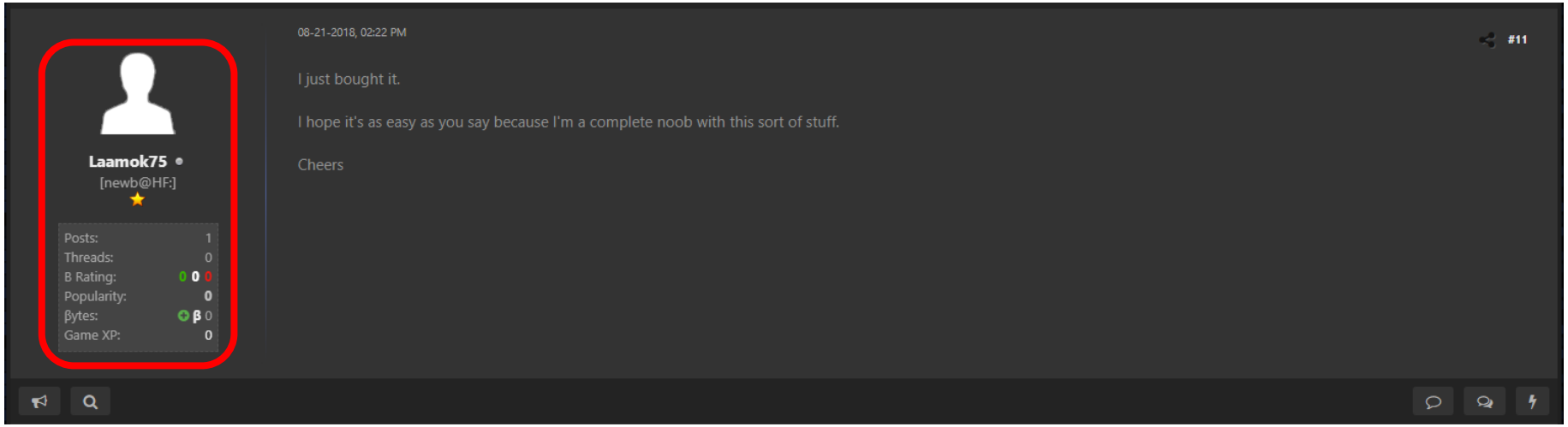
extract use name and profile of vendor, product/service name of each thread, price, payment method, **# of comments (i.e., replies), # of reviews**;

Premium Sellers Section

Thread / Author		Replies	Views	Last Post [asc]
Important Threads				
[$6,000 PER MONTH] => THE PRIVATE CASHFLOW SYSTEM <= | MAKE MONEY ONLINE | EARN TODAY Tokyo [Pages: 1 2 3 4 ... 45]	MAKE $	666	19,420	3 minutes ago Last Post: Stanley
$500/DAY =>Let's Cut The BULLSH*T AND MAKE SOME REAL MONEY<= [100% SUCCESS or REFUND] Vyrez. [Pages: 1 2 3 4 ... 21]	AMAZING	311	4,991	9 minutes ago Last Post: Stanley
$500/DAY =>THE ONLY WORKING and MOST PROFITABLE METHOD ON THE FORUM<= EARN or REFUND Trappy [Pages: 1 2 3 4 ... 153]	AMAZING	2,290	94,635	41 minutes ago Last Post: Stanley
[2019] ==> THE ONLY WORKING WAY TO EARN $350+/DAY ONLINE <== [EARN or 100% REFUND] Sceptic [Pages: 1 2 3 4 ... 20]	TOP DEAL	285	11,255	3 hours ago Last Post: Stanley
PRIVATE MONEY MAKING METHOD | MENTORING | $2000-$3000 MONTHLY | VIDEO EARNINGS PROOF Bezos [Pages: 1 2 3 4 ... 109]	EXCLUSIVE	1,634	97,431	5 hours ago Last Post: Stanley
[#1][$10,000/MONTH] => THE #1 CUTTING-EDGE INCOME PROGRAM ON HF <= [EARN or REFUND!] Zeus [Pages: 1 2 3 4 ... 69]	GREATEST	1,025	49,043	11 hours ago Last Post: Zeus
Normal Threads				
Selling Grubhub glitch get $12 off on every order. Donald Trump Jr. [Pages: 1 2]		18	433	5 minutes ago Last Post: Banana
[AUTOMATED] SPOTIFY PREMIUM LIFETIME UPGRADE! INSTANT! OVER 18K UPGRADES! WORLDWIDE! God-Zeus [Pages: 1 2 3 4 ... 28]		416	15,741	17 minutes ago Last Post: Molec

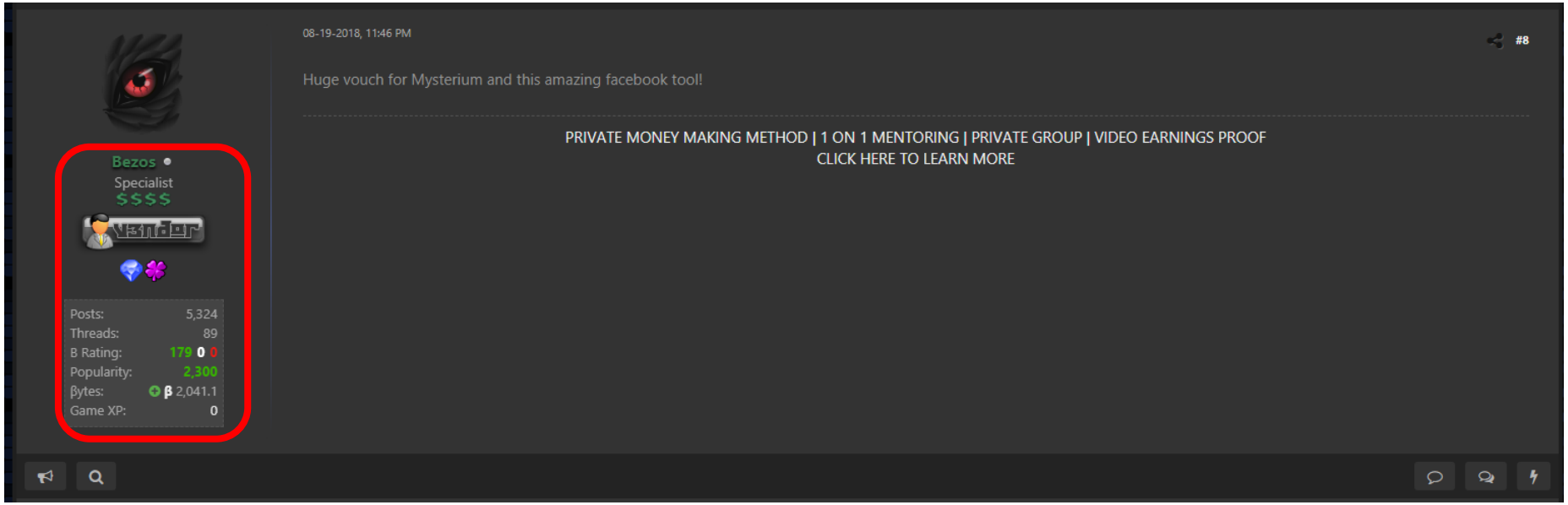
(2) **Comments**: classify each comment

- username and profile of commenter
- trading [**Yes**/No/Uncertain]
- contracted customer [Yes/No]
- review [Positive/Negative/Neutral]
- Q&A [Yes/No]
- other.



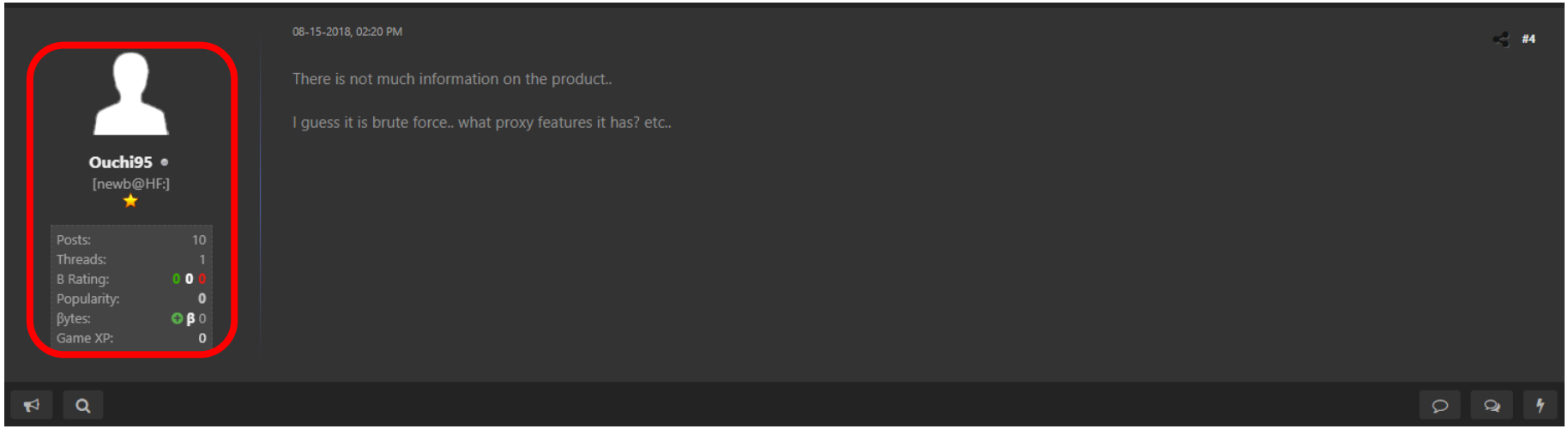
(2) **Comments**: classify each comment

- username and profile of commenter
- trading [Yes/No/Uncertain]
- contracted customer [Yes/No]
- review [**Positive**/Negative/Neutral]
- Q&A [Yes/No]
- other.



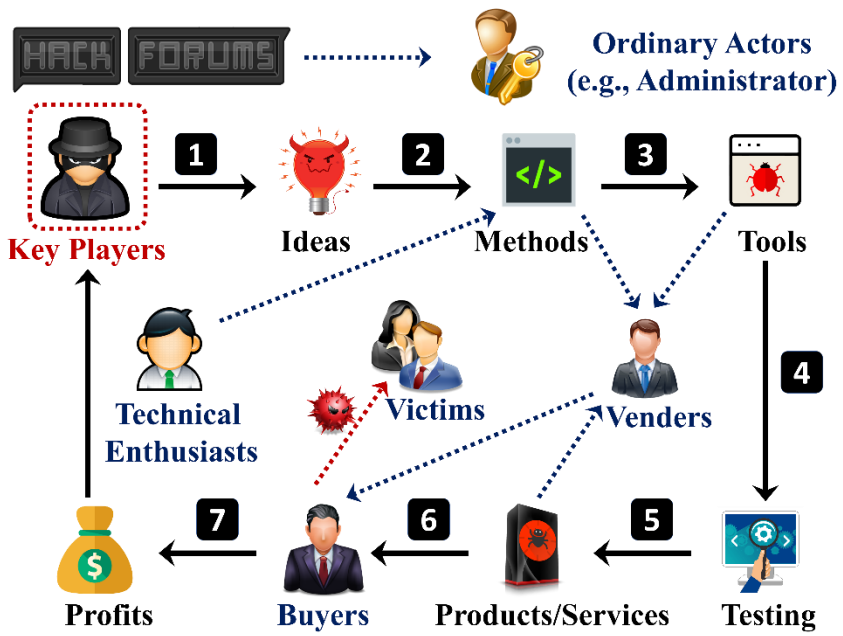
(2) **Comments**: classify each comment

- username and profile of commenter
- trading [Yes/No/Uncertain]
- contracted customer [Yes/No]
- review [Positive/Negative/Neutral]
- Q&A [**Yes**/No]
- other.



T1: Gain Deep Insights into the Online Underground Ecosystem

3. Based on the above steps, please develop your own solutions/tools for in-depth analysis: (1) find out the **key players** (i.e., most active vendors and buyers) for the kind of crimeware (P1) and CaaS (S1) you explore; (2) further analyze the top key players (one vendor and one buyer) for P1 and S1 to find out: i) whether he/she is an individual or organization; ii) what other products he/she sell or buy; iii) how he/she influent others in the market; iv) if he/she is active in other markets and how he/she will have the impacts in the cyberspace, etc. Describe the storyline and provide the case studies to elaborate your findings.



Role of key players in the value chain in Hack Forums

1. Key players generate ideas by analysis of market demands
2. Exploit the vulnerabilities and devise the methods
3. Develop the tools using the devised methods
4. Testing the tools by purchasing some accounts and bots
5. Produce and advertise illicit products or services
6. Sell the illicit products or services
7. Monetize the illicit products or services

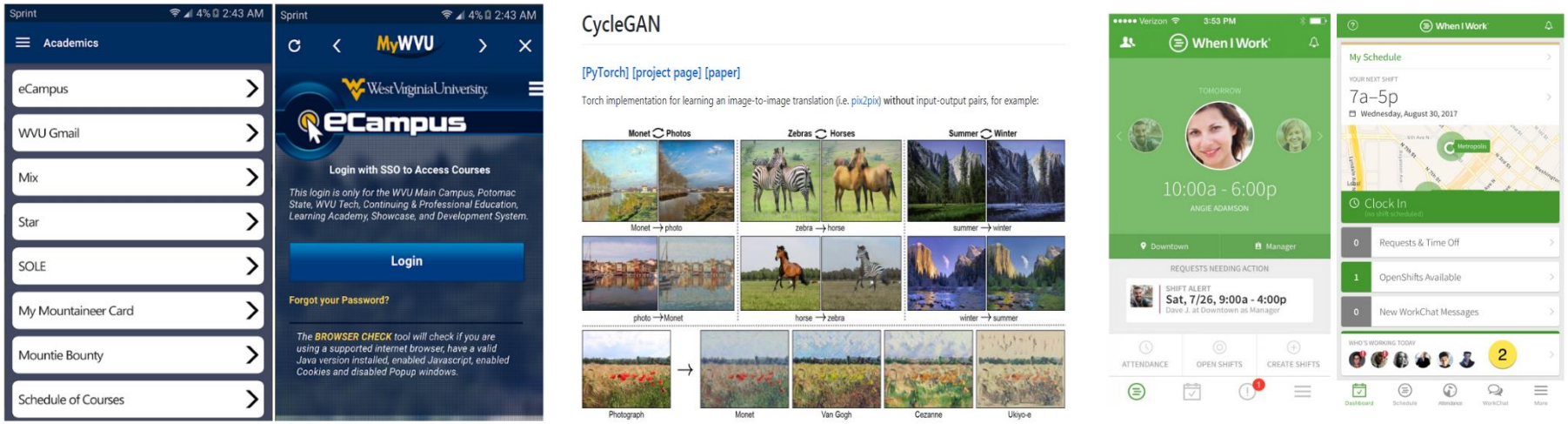
T1: Gain Deep Insights into the Online Underground Ecosystem

4. Based on the above findings and analysis, devise your solutions to help inform effective countermeasures.
5. A fully developed project report with required format should be submitted.
6. Finally, present your project in the class.

T2: The study of code security problem in social coding platforms

Project Background and Motivation

- Software is everywhere.



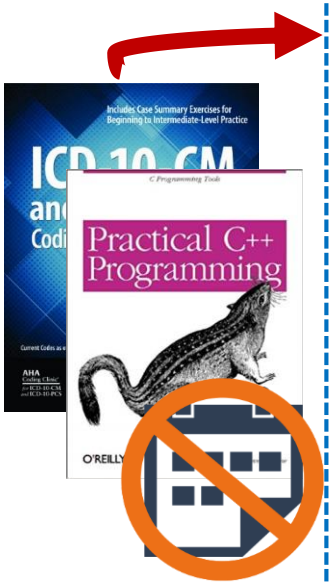
(a) eCampus software for education community. (b) CycleGAN developed by research community. (c) A check-in app for business community.

Figure 1: Example software developed and used for education, research and business communities.

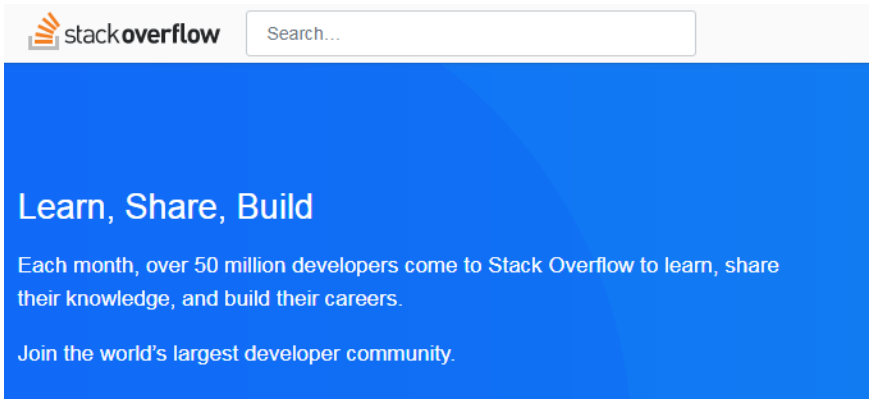
- There have been more than 1 billion software products available worldwide;
- It's estimated that the global software market revenue will reach over $507 billions in 2021.

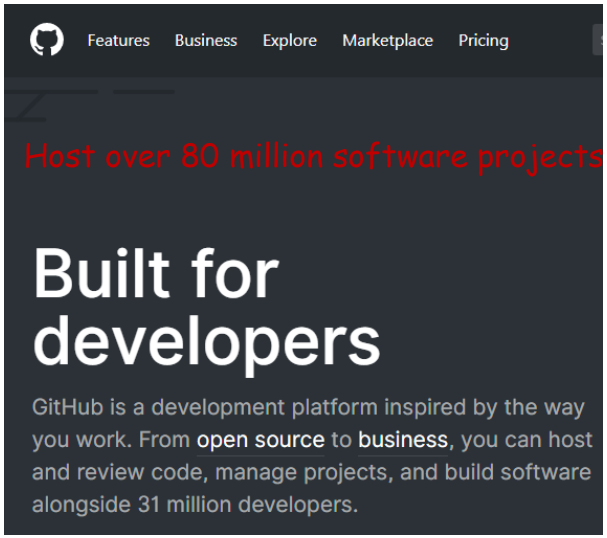
Modern Software Programming Cyberinfrastructure

- Modern software programming cyberinfrastructure (CI), consisting of online discussion platforms like Stack Overflow and social coding repositories such as Github, offers an open-source and collaborative environment for scientific communities to expedite the process of software development.



Modern software programming cyberinfrastructure (CI)

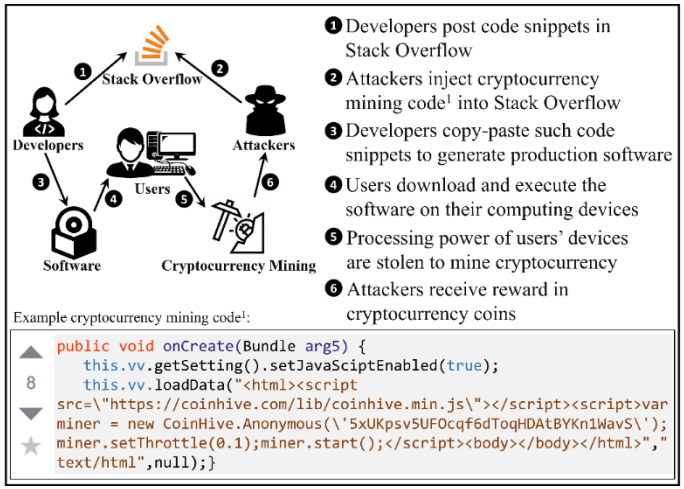




Scientific Credibility of Modern Software Programming CI

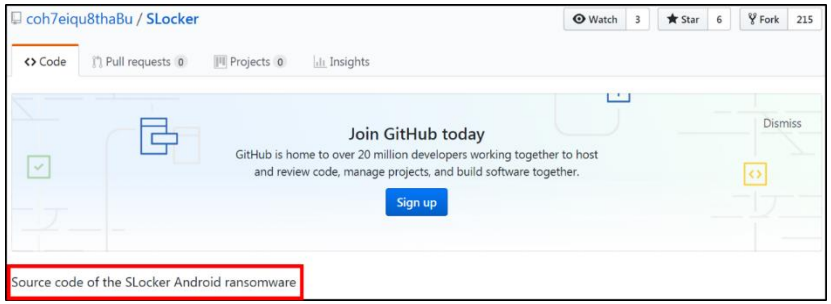
Despite the apparent benefits of this new social coding paradigm, its potential security-related risks have been largely overlooked; insecure or malicious codes could be easily embedded (e.g., through forking or committing) and distributed (e.g., through copy-and-paste), which could severely damage the scientific credibility of CI.

- Can one trust such code snippets or existing software project files?
- In other words, how much do we know about the **scientific credibility** of Stack Overflow and GitHub from the **security** point of view?



Example of code security attacks in Stack Overflow.

- There has been no principled ways of dealing with insecure or malicious codes in modern software programming CI.



Android ransomware published on GitHub.

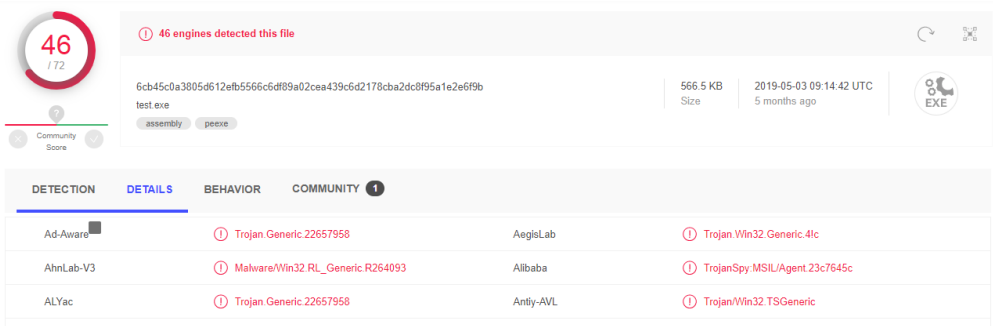
T2: Study of social coding security problem

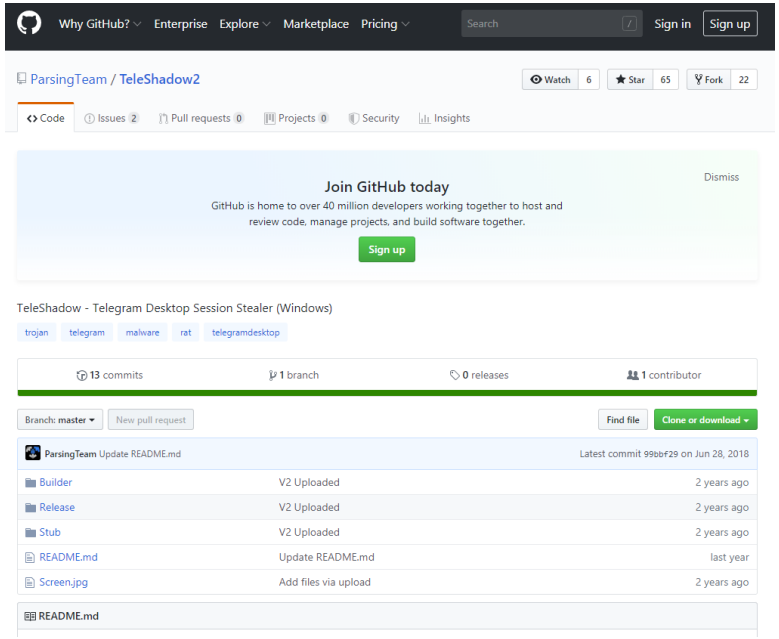
1. You are first asked to explore and devise your own methods to identify the malicious codes/projects hosted in GitHub. Describe and summarize how you identify them. ***Note**: you can regard the code/project is malicious if over 1/4 of AV engines in VT detected it as malicious.*

Example

https://github.com/ParsingTeam/TeleShadow2/

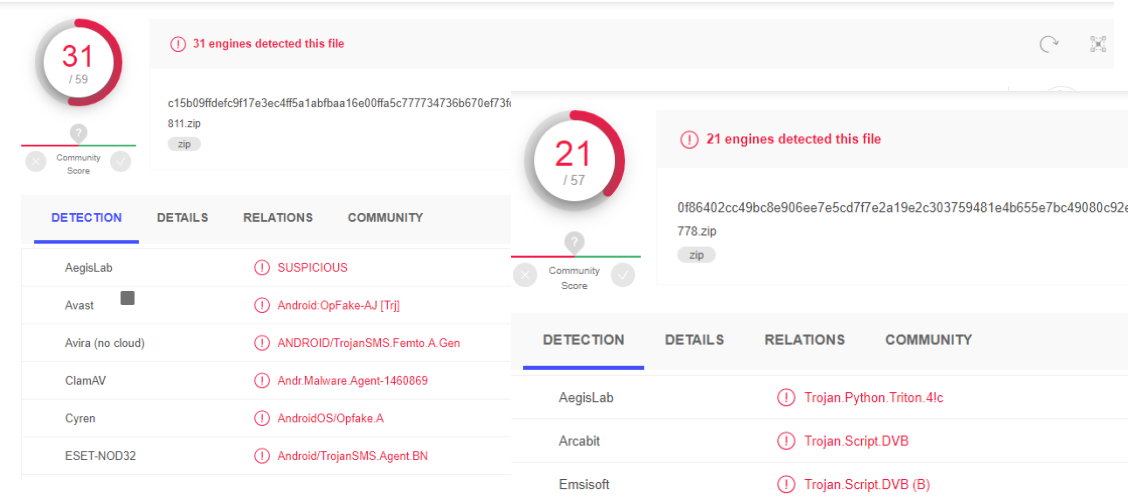
https://www.virustotal.com/gui/file/6cb45c0a3805d612efb5566c6df89a02cea439c6d2178cba2dc8f95a1e2e6f9b/detection

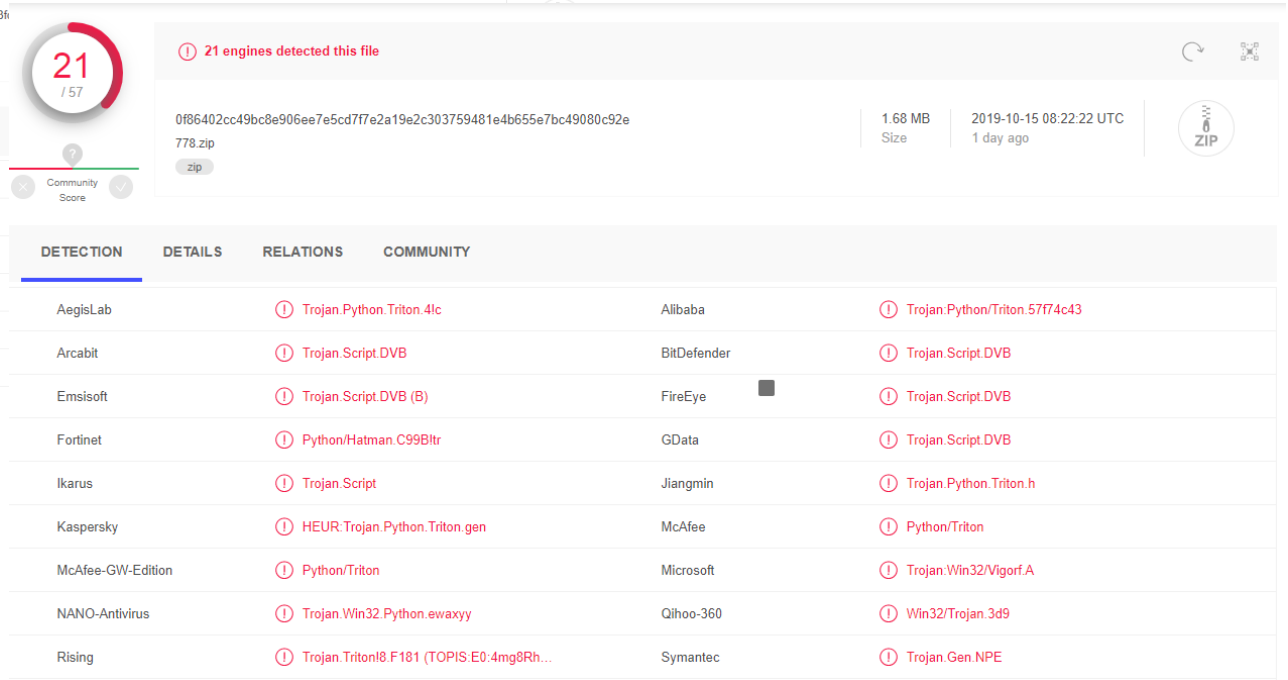




T2: Study of social coding security problem

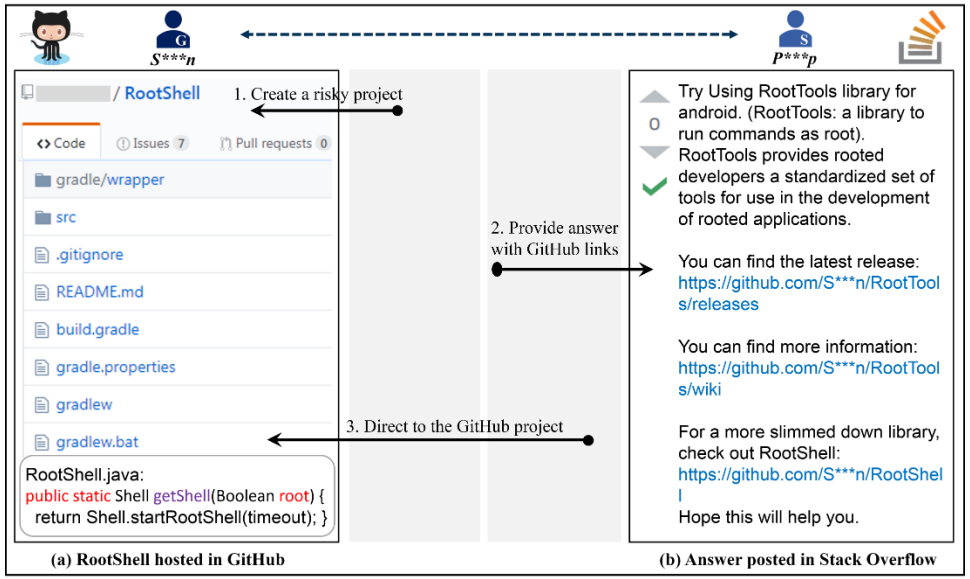
2. In order to leverage the pre-identified samples to automate the detection of malicious codes/projects in GitHub, you should propose your own solutions and assess the effectiveness of your developed tools. You need to submit the ground truth using the required template in Canvas (each group needs to detect > 100 malicious projects hosted in GitHub). Your submission will be validated by cross-validation during grading.





T2: Study of social coding security problem

3. Based on the detected malicious codes/projects hosted in GitHub, you are asked to develop your own approaches and tools to further analyze how these threats are propagated or disseminated in the cyberspace: (1) Who are the key players? What is his/her profile, his/her reputation, geo-location, etc? (2) How they disseminate the malicious codes/projects hosted in GitHub and interact with other users cross-platforms (e.g., between Stack Overflow and GitHub to propagate or disseminate the malicious codes/projects)? Afterwards, please further analyze and figure out their dissemination networks (i.e., the communities hiding behind if any) and how they interact with each other. You need to find out at least 10 solid cases to demonstrate how the malicious codes/projects are disseminated in the cyberspace. Describe the storyline and provide the case studies to elaborate your findings.



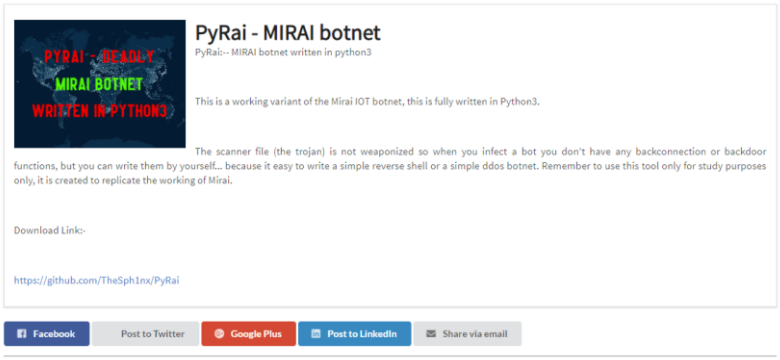
Interplay between GitHub and Stack Overflow.

T2: Study of social coding security problem

3. Based on the detected malicious codes/projects hosted in GitHub, you are asked to develop your own approaches and tools to further analyze how these threats are propagated or disseminated in the cyberspace: (1) Who are the key players? What is his/her profile, his/her reputation, geo-location, etc? (2) How they disseminate the malicious codes/projects hosted in GitHub and interact with other users cross-platforms (e.g., between Stack Overflow and GitHub to propagate or disseminate the malicious codes/projects)? Afterwards, please further analyze and figure out their dissemination networks (i.e., the communities hiding behind if any) and how they interact with each other. You need to find out at least 10 solid cases to demonstrate how the malicious codes/projects are disseminated in the cyberspace. Describe the storyline and provide the case studies to elaborate your findings.

GitHub Link: https://github.com/TheSph1nx/PyRai

https://datasec.az/project/66/

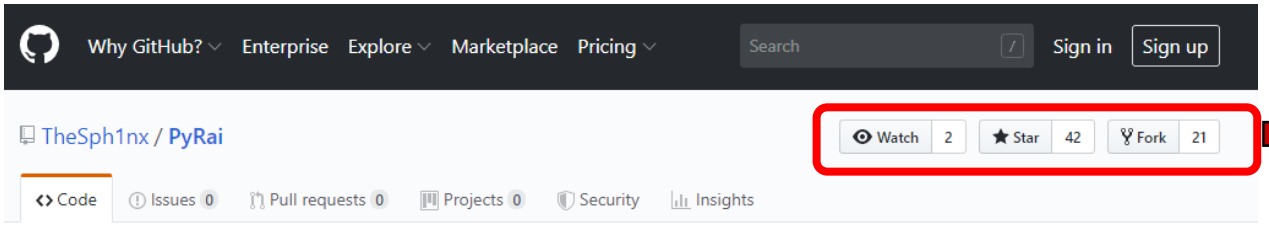


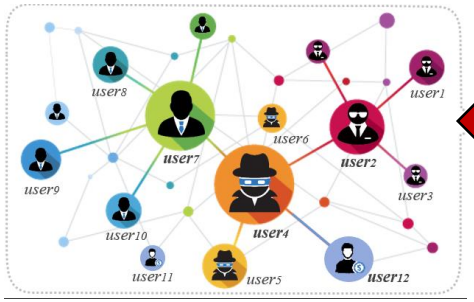
https://hi-in.facebook.com/ncybersec/posts/1247362645434457

T2: Study of social coding security problem

3. Based on the detected malicious codes/projects hosted in GitHub, you are asked to develop your own approaches and tools to further analyze how these threats are propagated or disseminated in the cyberspace: (1) Who are the key players? What is his/her profile, his/her reputation, geo-location, etc? (2) How they disseminate the malicious codes/projects hosted in GitHub and interact with other users cross-platforms (e.g., between Stack Overflow and GitHub to propagate or disseminate the malicious codes/projects)? Afterwards, please further analyze and figure out their dissemination networks (i.e., the communities hiding behind if any) and how they interact with each other. You need to find out at least 10 solid cases to demonstrate how the malicious codes/projects are disseminated in the cyberspace. Describe the storyline and provide the case studies to elaborate your findings.

GitHub Link: https://github.com/TheSph1nx/PyRai





T2: Study of code security problem in Stack Overflow & GitHub

4. Based on the above findings and analysis, devise your solutions to help inform effective countermeasures.
5. A fully developed project report with required format should be submitted.
6. Finally, present your project in the class.


Thank you!