THE UNIVERSITY OF
MELBOURNE

Wrapping Up

COMP90073
Security Analytics

Sarah Erfani, Yi Han
CIS

Semester 2, 2021

Outline

- Exam
- Subject revision
- Assignment feedback

Exam Instructions

- Wednesday 03/Nov, 3:00pm, LMS.
- Worth 60 marks, 30 mark hurdle.
- 15 minutes reading time, 2 hours writing time.
- Answer all questions.
- Note that questions are not of equal value.
- 2–3 sentences sufficient for when brief descriptive answer requested.
- Please use your script book for the long answer question, clearly marking where the response starts. Any pages which are not labelled as forming part of the response to that question number will not be considered during marking.
- A sample exam will be available soon.

Types of Questions

There are a mix of question types on the exam.

- **Conceptual:** A question which tests or requires you to define or explain a concept, term, or algorithm introduced in the subject.
- **Problem solving:** A question which asks you to use a specific algorithm or formula to solve a problem on some data.
- **Application:** A question which asks you to demonstrate that you have gained a high-level understanding of the methods and algorithms covered in this subject, and can apply that understanding.

A Comment on Mathematical Concepts

We expect you to be able to do:

- Remember simple, key formulas
- Read and understand more complex formulas that have been presented for core concepts, provided “bare”.
  - E.g., attacker’s objective functions in adversarial attacks against machine learning models

- Addition, subtraction, multiplication, division
- Reducing and ordering of fractions
- Gradient-descent based method for generating adversarial samples

Week 1: Introduction to Cybersecurity & Security Analytics Use Cases and Data

- Core cyber security principle
  - Explain CIA triad
  - Apply the appropriate controls to protect CIA

- Key access control concepts
  - Describe access control and four key attributes
  - Explain "Defense in Depth"

- Security analytics use cases and data
  - Explain seven common use cases
  - Explain four data sources

Week 2: Cybersecurity Landscape

- Cyber Kill Chain
  - Explain seven steps of cyber kill chain
  - Model cyber attacks using cyber kill chain

Week 3: Network Security & Attacks

- Fundamentals of Networking Protocols
  - Understand DHCP & DNS protocols and TCP three-way handshake

- Network Attacks
  - Compare different types of attacks
  - Understand how network attacks work
  - Describe examples of different types of attacks

- Network Security Systems
  - Explain DMZ and network segmentation
  - Explain NAT & PAT process
  - Compare the difference between IDS and IPS

Week 4: Botnet & DDoS Deep Dive & Business Context for Cybersecurity Management

- Botnet Deep Dive
  - Explain phases of botnet lifecycle
  - Compare the difference between push and pull based propagation methods

- DDoS Deep Dive
  - Compare three types of DDoS attacks

- Information Security Management Governance
  - Determine qualitative risks
  - Calculate quantitative risks

Week 5: Intro to Anomaly Detection, Clustering and Density-based methods

- Describe shortcomings of convectional security systems
- Discus the objective anomaly detection
- Define different types of anomalies
- Discuss operation of iForest, and describe the advantages of this method
- Apply clustering algorithms to identify anomalies
- Discuss differences between distance and density based methods

Week 6: Anomaly Detection in Evolving Data Streams

- Characterise the differences between batch and incremental learning
- Describe the operation and properties of HS-tree algorithm
- Describe an efficient approach to extend LOF to incremental learning

Week 6: Anomaly Detection Using Support Vector Machine

- Describe the operation of SVDD/OCSVM
- Characterise the key parameters of SVDD/OCSVM
- Derive the dual formulation SVDD/OCSVM from the primal formulation

Week 7: Autoencoders and their Applications

- Describe operation and training of autoencoder
- Identify anomalies using an autoencoder
- Characterise properties of different types of autoencoders
- Characterise the key parameters different autoencoders' loss function

Week 8: Graph Anomaly Detection

- Graphs cannot always be treated as points lying in a multi-dimensional space independently.
- Preserve data structure with node embedding
- Characterise the properties of random walk and graph convolutional network
- Apply graph embedding for anomaly detection

Week 8: Contrast Mining

- Explain the advantage of contrast mining in cybersecurity problems
- Compare and contrast alerts from different datasets
- Find frequent patterns using FP-Growth algorithm

Week 9: Adversarial Machine Learning – Vulnerabilities (Part I)

- Evasion attacks
  - Indiscriminate: $\arg\min_{\delta\in[0,1]^d} \|\delta\| - c\cdot f_{true}(x+\delta)$
  - Targeted: $\arg\min_{\delta\in[0,1]^d} \|\delta\| + c\cdot f_{target}(x+\delta)$
  - Gradient-descent based approach to generate adversarial samples
  - Automatic differentiation
- Poisoning attacks
  - Attacker's objective: $O_A(D,\hat{\theta}_D) = \|\hat{\theta}_D - \theta^*\| + \|D - D_0\|_2$
  - $\hat{\theta}_D$ (θ^*): parameter of the poisoned (targeted) model
  - D (D_0): poisoned (original) training dataset
- Transferability
  - Black-box attacks

Week 10: Adversarial Machine Learning – Vulnerabilities (Part II), Explanation, Detection & Defence

- Adversarial attacks in domains other than computer vision (malware detection)
- Potential locations of adversarial samples
  - Off the data manifold of legitimate data
- Why are machine learning models vulnerable?
  - Insufficient training data
  - Unnecessary features
- How to defend against adversarial machine learning?
  - Data-driven defences
    - Filtering adversarial samples
    - Adversarial training
    - Project to lower dimension
  - Learner robustification
    - Distillation
    - Stability training
  - Adaptive attackers

THE UNIVERSITY OF MELBOURNE

Week 11: Adversarial Reinforcement Learning

- Reinforcement learning
  - State, action, reward
  - Value function, policy, model
  - Q-learning → Q-network → DQN → DDQN
- Adversarial reinforcement learning
  - Manipulate the states observed by the agent
  - Cross entropy loss: $J = -\sum_i p_i log \pi_i$
    - π_i: probability of taking action a_i
    - $p_i = \begin{cases} 1, \text{if } a_i = \text{optimal action} \\ 0, \qquad\qquad\qquad \text{otherwise} \end{cases}$
    - Maximise J → minimise the probability of taking the optimal action
  - Test time/training time
  - Timing of the attack
- Defence – adversarial training

$$\begin{pmatrix} 0.07 \\ 0.15 \\ \dots \\ 0.02 \\ \boxed{0.61} \end{pmatrix} \longleftrightarrow \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \\ \boxed{1} \end{pmatrix}$$

Week 12: Guest Lecture

- No examinable material

Wrapping Up

- I hope you enjoyed this introduction to security analytics
- Maybe we'll see you in PhD programs
- Thank you for your patient attention
- Good luck with your exams and future studies