someone please help me with these notes :-(

Morning lecture

6841 ONLY:

- there will be exploitation in the finals
- answer k out of n questions, where one of them is an exploitation question
- use `python -c` in terminal if you wish
- tooling: stub command environment, gdb

Privacy lecture by Adam

Privacy in the modern age: we're stuffed

- best one sentence description
- a lot easier to execute security argument than privacy argument

this lecture is held under Chatham House Rule:
noun
a rule or principle according to which information disclosed during a meeting may be reported by those present, but the source of that information may not be explicitly or implicitly identified.

Story:
Adam and his sister both studied at UNSW, engineering and psychology respectively. One day Adam was listening to a podcast on the Stanford Prison Experiment for an ethics course (not his favourite). In 2009, Adam believes that it was more important that engineers and computer scientists study ethics than doctors and scientists.

"**Data is the new oil**" - coined 5-10 years ago
Data is refined and carefully protected just like oil, however people forget that oil is a huge source of danger. eg. storing too much, oil spills, long term unknowns. This leads to massive reputation damage.

Issues around data collection:

- **data linking**: network effect of linking data sources - 63% of the population can be uniquely identified by the combination of their gender, date of birth, and zip code
- **de-anonymisation**: all anonymised data runs the risk of being reversed. it is often difficult to draw a sharp distinction between personal information and de-identified data. Possible Bayes rule reverse inference.
- **massive scale**: data is being collected at a scale and pace that could not be imagined in other times
- **massive data-breach**: data is increasingly being disclosed in massive data breaches with serious impacts
- **phishing expeditions**: large and persistent datasets provide the temptation for phishing expeditions in the future

“There's just so much bloody data" - Adam
eg. location data - mac addresses, pinging towers
When we walk down the road, we are spewing information.

Roomba maps your house and uploads it to servers, Bose track your music history

bodies that collect Adam’s data:

- university
- government eg. tax department, welfare
- tech companies eg. bose, fitbit, spotify
- large tech eg. google, facebook

personal data:
- political leanings
- friend groups
- news sources
- health data
- communication history
- purchase history
- browsing history
- criminal records
- credit scores

bad if:
- government found out political leanings
- tax depot got personal docs
- insurance companies got health records

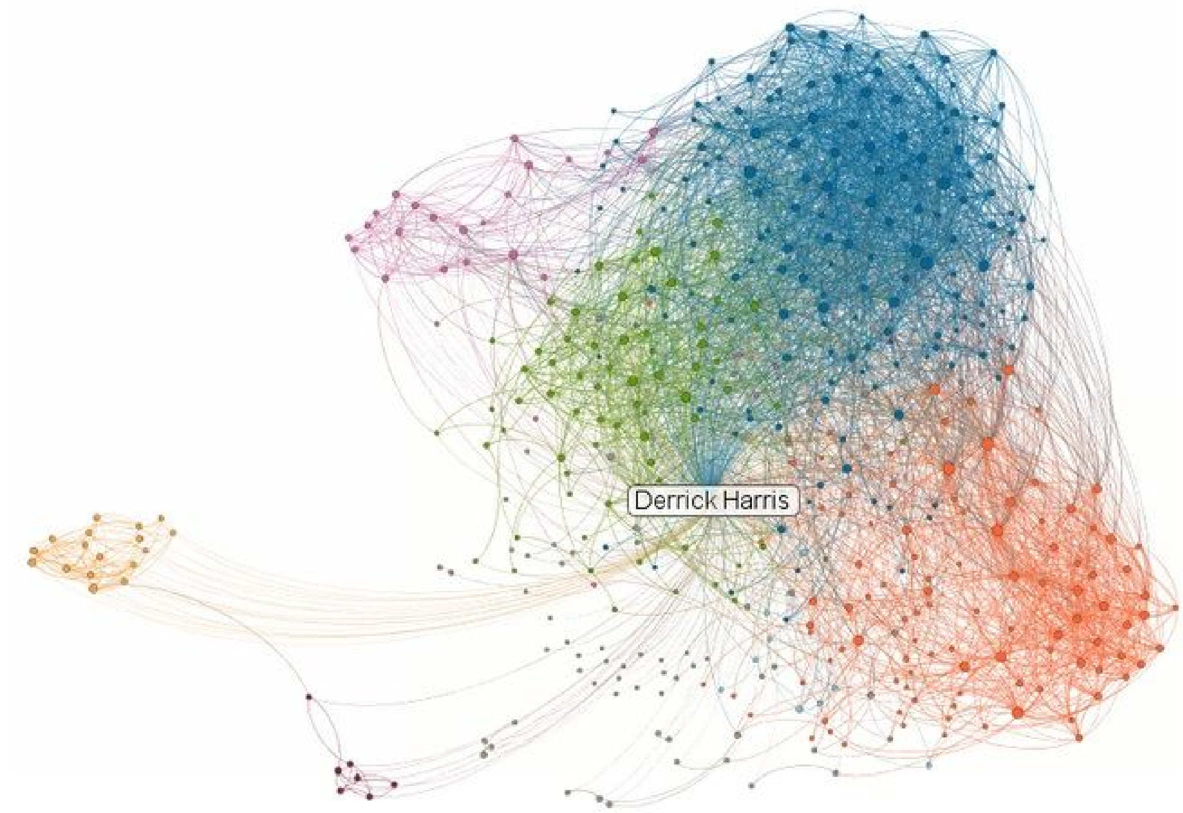
big tech companies have a lot of our data

Google Takeout
Warning: think before you download. Do you expose yourself to more risk downloading it? These tools weren't built for users, they were actually built for the government for law enforcement so that they could self select the data that they needed.

Adam went to a tech con at Bent's Basin, in the middle of nowhere. There was no reception (data or wifi) but Google was still able to track his location as shown in Google Takeout. He also found 5 years of search history.

iPhone data
Apple has a backup software which dumps all of your phone’s content onto your computer. It contains contacts, photos, notes, call history and text messages. Adam found over 87,000 texts since 2008.

Meta-data is incredibly powerful. The government probably creates a giant graph to link people together like below:



App data
eg. MS word - Adam found an old case study from COMP3441 (now COMP6441)
Tinder had an exposed SQLite database in backup which contained caches of conversations and unique identifiers.

he's single ladies and gents ;)

Facebook
some of the things you can download:
- messages
- timeline info
- email addresses added to your account (including those removed)
- addresses
- check-in history
- ads you clicked on
- ad topics
- ip addresses
- active sessions

Facebook has a lot more data that doesn't appear in the list above. It is their intellectual property, but it is also information about us.

How can we use this data?
A day in the life of Adam - 13th July 2017
Using information from his iPhone, Facebook, Google, Opal card, Dropbox, and Tinder

web history - how to make a strong password, trello board
opal - eastwood to town hall, town hall to eastwood
photos + GPS data - 363 george street, qvb t2 store

what we know:
visited city after 1pm
attended a 'cyber' event
evidence of presentation
visited qvb stores
home at 7pm

The meta-data revealed a lot of information, but not all the information

---

Q&A
Incognito mode doesn't store browser history and cookies

Companies are using browser fingerprinting instead of cookies. When you visit a website your screen resolution, plugins installed and other information is sent to the site. Incognito mode protects you from shoulder surfing but not traffic inference, ISP, security (eg. phishing), tracking and shadow fingerprinting.

Under the Australian legislation, companies like Facebook have to hand over their information about you. They don't have to disclose the inferences made about you
eg. race, job, marital status, sexual preference, political leanings

Data can be used against you
Bank data contains ID data eg. drivers licence or passport; and spending data

Privacy concern: risks around bank data being made publicly available
This data is retained for 7 years for tax purposes.
If you don't want your spending to be tracked, use cash

Adam suspects cash will be illegal soon. There was a legislation that stated any cash transactions over $10k will be made illegal. Unsure if legislation has been passed...

Companies collaborate closely with the government (banks especially)
Financial data is a huge source of information. Intelligence agencies collaborate closely with a bunch of companies to detect spies.

Having a VPN essentially means not trusting your ISP. It creates an encrypted tunnel to another computer and protects from snooping companies, man in the middle attacks. You can still get phished. It is not secure after the VPN terminates.

It's a BIG problem. Categorised into three sections:

- Private industry
- Government
- Intelligence community

Adam could readily find GPS coordinates of photos on Flickr!

Private companies:
data is the new oil, all going crazy about it
"data lakes" - streams coming together into the lake
run analytic reporting, make inferences

Aim: hoover as much data as possible and put it in one place

collect all the data, get it into one spot, oh no too much access restrictions, oh no it's encrypted

wrong attitude to take: "of course it's encrypted, of course there's access control"
walk back from this thinking
examine the possibility that nothing is perfect and done on time
normally people want results so non-functional requirements often scrapped (eg. privacy)

hundreds of companies struggling to keep their computers up to date
less than 50% of businesses have patching within a month
false to assume business will be up to date

what inferences can you make from a data set?
mental illness? divorce? pregnancy?

Quantium
data analytics, owned by Woolworths
massive data mining based on shopping habits
Every time you scan your Everyday Rewards card, it is pulled into a data set and the data ends up on your Facebook to target you
Facebook states that the company received "de-identified" purchase data

Coles
Flyby vouchers etc.
Adam’s theory is that they only care about the first time you sign up because it links your credit card to a name, email, phone number

Qantas
Frequent Flyer Program - one of the biggest businesses inside Qantas
Qantas loyalty ($369m) > Qantas international ($327m)
The loyalty program has an estimated worth of $4bil and Qantas itself worth ~$9bil

Facebook:
not the best reputation in privacy space
eg. tampering of federal election in the states

politicians x tech companies

Workplace surveillance:
everything done on work laptop can be monitored
apps on mobile phones being used to track people
skype, hoover, slack
many instances where company found something to fire undesirable employee

unsw was tracking location throughout campus
access point - triangulate signal strengths
DNS requests are in the plain, unsw has access to this
whether they are storing it is another question

Issues around data collection:
often hear don't worry you can't de-anonymise it, it's not linked etc.
it's secure, it only does x under y condition
it's end to end encrypted
it's algorithms not people
it's locked down, only x can access it
we have thorough auditing in place

all this means nothing ^

eg. imessages - end to end encrypted, but icloud is not

There isn't enough public data sets to train AI on stuff eg. detection of nude photographs

Anonymising strategies:
- redaction
- encryption/hashing
- pseudonyms (unique identifier)
- statistical noise/'binning'
- aggregation

every one of these strategies is breakable
depends how determined the attacker is

4 data points in mobile cell towers enough to identify 95% of the population

Takeaways:
- anonymisations is not a certainty
- most data breaches had some sort of identifier between the records
- the pattern becomes the fingerprint, not the data itself
- to adequately anonymise data, might render the data as useless

The government:

- centrelink
- my health record - no one was asking for this; incredibly valuable data, probably selling the data back to drug companies and researchers
- opal: location history
- mygov: starting to link everything together
- drivers licence

Truism about government data collection:

1. Collection laws will emerge and re-emerge in different forms until passed
2. Collection is always going to increase in scope; it never diminishes
3. Terms of usage always start narrow, but quickly broaden
4. The least bad thing is no change in the status quo; most bad is robo cop scenario
5. Access is automated

Snowden discoveries:

- PRISM
- XKeyscore
- Tempora
- ECHELON

Mandatory metadata retention
- incoming/outcoming telephone caller id
- date, time, duration of call
- location of the device which call was made
- unique identifier for each phone
- email address

"It's only metadata"

Access and Assistance bill (2018)
https://www.homeaffairs.gov.au/about-us/our-portfolios/national-security/lawful-access-telecommunications/myths-assistance-access-act
---

Is it justified?
How do you fight it?

What about meta information?

Adam's opinion:

- **Start with engineers**: Understand the ethics around it. Take a stand if necessary. Broaden our perspectives
- **Advocacy**: Educate other people. Complain to governments and organisations, it’s our duty as technical people in the room
- **Design**: It's not all or nothing. Engineer a genuine choice. Engineer safer ways eg. on device. Minimise data along the way
- **Story Telling**: Allow people to fully understand the impact eg. Blackmirror, 1984

In memory of John Gibson
2 months earlier, the Ashley Maddison breach occurred. He was so affected about the data breach and what the companies had about him that took his own life.

Evening lecture

Rootkits seminar

root: root, or admin; highest possible level of access privilege
kit: software that grants root-level access to the system

a rootkit can:
- conceal itself
- execute any process
- make changes to the system
- track usage of the system

not malicious by itself
can enable malware
often bundled with malware
Zeus aka ZeuS aka Zbot - trojan
uses rootkit to hide keylogger

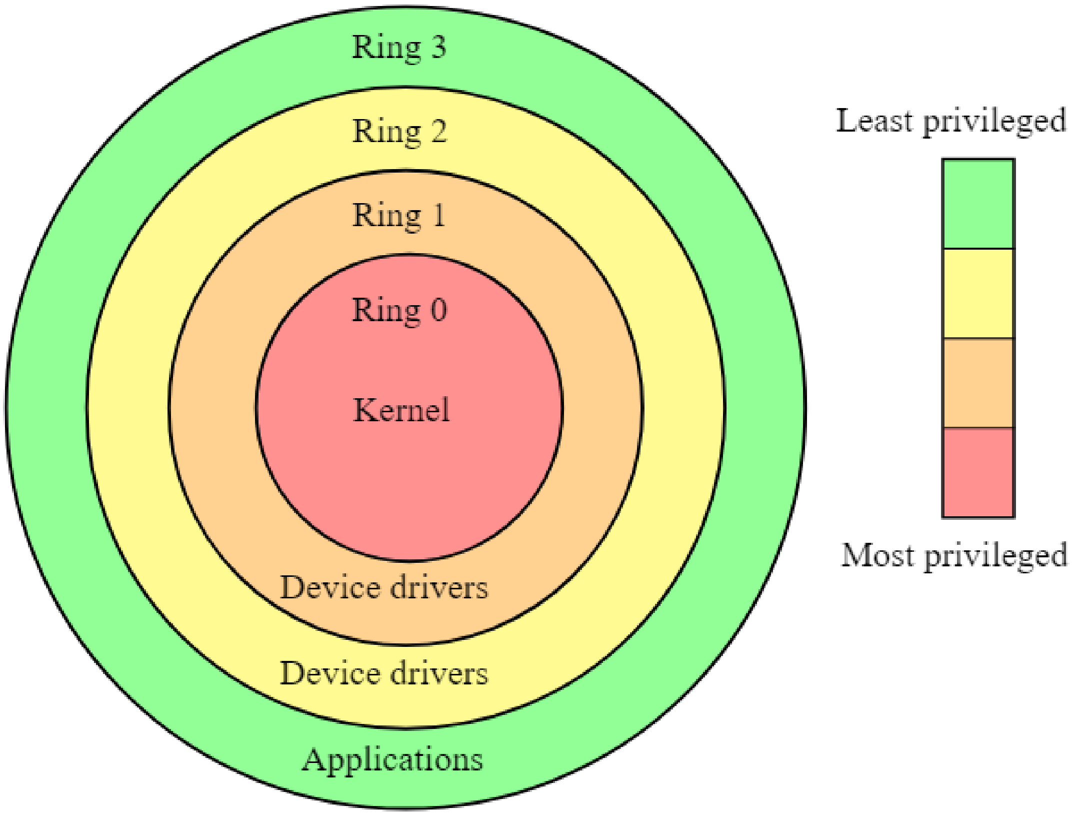
installation:
- phishing attacks
- social engineering
- inserting usb into system
- someone willingly granting access to system

history:
- earliest rootkit - admin tools that replaced legitimate tools on UNIX
pretty easy to detect
- 1980 - Ken Thompson
- 1986 - 'brain virus', not really a rootkit uses cloaking (stealth virus)

- 1990 - first real rootkit
- 1999 - first malicious rootkit for windows OS - NTRootkit
- 2005 - sony BMG modifies operating system to tamper with disc copying
- 2009 - first rootkit for Mac OS X
- 2009 - Zeus infects 3.6 million devices in the US

concept of rootkits is actually quite old
gotten a lot more sophisticated as time went on

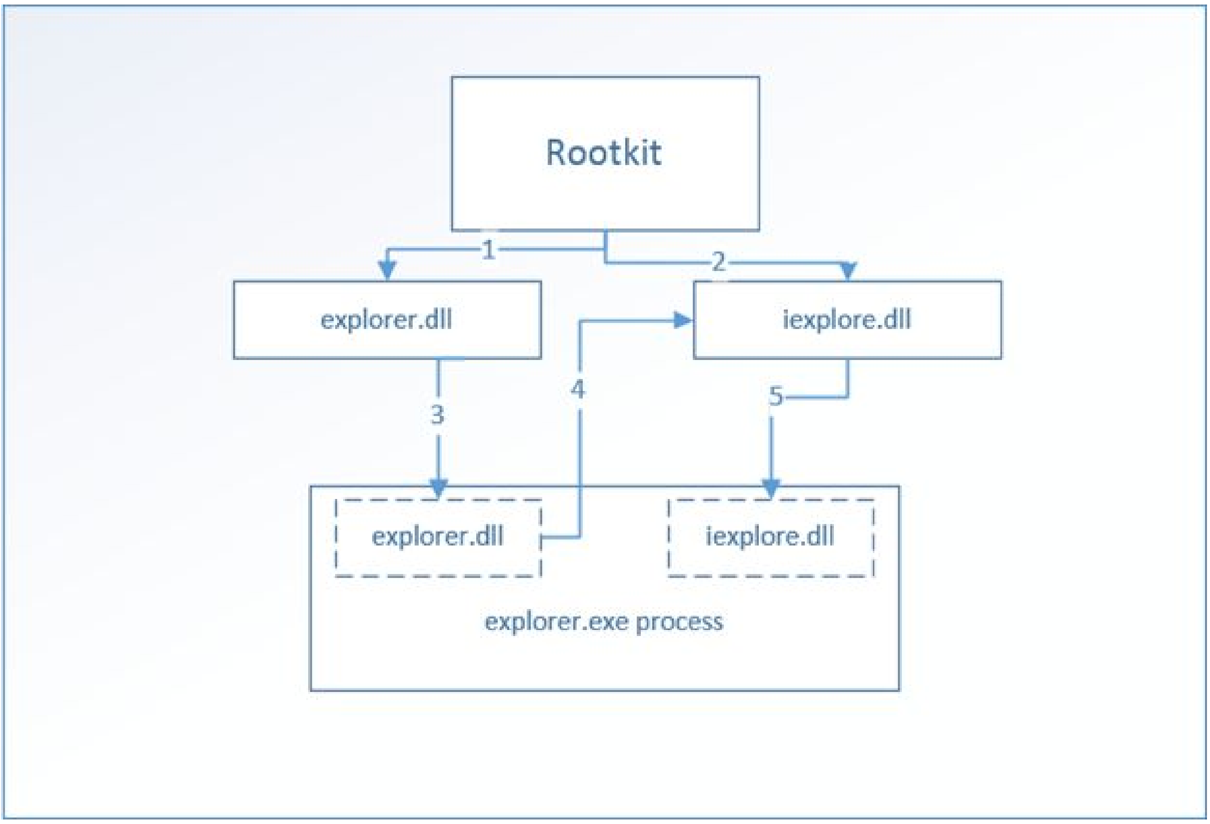
Privilege ring (not completely accurate, more inside kernel)



Types of rootkits:
- usermode

modifies user-level apps or shared libraries
remote access: backdoor sshd
privilege escalation: backdoor su
hiding:
  - process - replace ps, pidof, top
  - network - replace netstat, ifconfig
  - file hiding - replace ls, find

DLL injection
injecting into explorer.exe process



- kernel
violates trust of users and processes
alters objects stored in kernel memory
alter results of system calls eg. list files in dir
pro - hard to detect by traditional anti-malware software
con - harder to write

- system call hooking (BSD):
one of the most popular techniques
sysent - system call table of syscalls
have index point to your hook
syscall j now points to your hook
so it'll run your program

- memory based:
never write to disk
    - no physical presence
    - hard to detect
no persistence, sometimes that's ok
good for one-off, specific, targeted attacks
SucKit (2001)
    - pattern searches /dev/kmem and located syscall table
    - patches /dev/kmem by overwriting entries in the syscall table
    - undetectable by LKM detection methods

- bootkits:
replaces the legitimate boot loader with a boot loader under the attacker's control
usually performs the transition to 'protected mode' thus intercepting encryption and passwords
harder to detect
removing a Bootkit may 'brick' your system
can attack full disk encryption systems

- hypervisor level rootkits
exploit hardware virtualisation features such as Intel VT and AMD-V
runs at ring -1 (before the kernel) and hosts the OS as a virtual machine
intercepts hardware calls made by the oS
does not have to modify kernel, thus harder to detect than regular Bootkits
only detectable through extremely low level monitoring such as measuring hardware latencies

detection:
- boot into a different OS and check the contents of the drive
- RKHunter - hashes files and compares them with known good hash
- catch the OS lying
eg. checking netstat, nmap
- cat and mouse game, defence is always playing catch up

a lot of these tools (RKHunter, Tripwire, antivirus) can be evaded though

further reading material:
Designing BSD Rootkits
Rootkits & Bootkits
Rootkits Arsenal
http://www.phrack.org/
Reflections on trusting trust

modify your own kernel (in a virtual machine)!

Lecture by Lachlan

6447 assignment is to write a rootkit
exam has not been made yet

Lightning Talks

Hayden (tutor)
CA - certificate authentication
attack - typosquatting
registering a domain similar to existing one
eg. google -> gooogle

dnstwist - open source python tool
https://github.com/elceef/dnstwist

takes list of domains
generates a list of permutations based on given rules
has a feature to check if domain is registered

Jarrod
buffer overflow 101
writing your code inside a buffer instead of redirecting to your code
very cool rick roll
all the code is on github

Anon (unknown)
a guy's mum's work company got hacked
corporate emails the company had got compromised
someone had access to the password to the emails
suppliers sent invoice to company
hacker sent follow up email using subdomain of supplier email
footer identical
only thing different "sorry disregard previous email, here is the updated bank account"
lady who makes payments actually already made the payment to the legit bank account

received suspicious email requesting for payment
normal procedure is printing and stamping then handed to the lady to process payment
that one physical interaction saved the attack from happening

Final exam

has Richard sent me (caff) the draft exam? no.
Richard is going to dump a bunch of content on the openlearning finals page such as exam skeleton

6841 buffer overflow WILL be in final exam
a section where you have to answer k out of n questions
where one will be a buffer overflow question
two stage overflow
first stage pretty easy
second stage need to know how it works

last practice buffer overflow is not realistic to do in exam environment
rest is up to Richard to decide

8:45am exam

Questions:

Q: Do extended students get more time?
A: I don't think so. Richard likes to write a 2 hour exam and give you 3 hours to do it

Q: Is the only difference between 6841 and 6441 the seminars?
A: 6841 we expect more technical understanding of things such as buffer overflows. Suspicion is that 6841 component will only be the k out of n section.

Q: What if 6441 students do the 6841 questions?
A: Then I'll be really impressed. Probably no extra marks (unsure though)

"That's some hickity hackity stuff, damn" - Lachlan

"I'm really proud of you and your spirit and the work you have put into the course and how you have all worked together. India is lovely but I miss you all and wish I was there to share the final week with you. Please send me a huge class selfie if you can take one Lachlan! I'll be running a revision session in the week before the exam - will sort out a room and time when I get back and post it on open learning." - Richard

Where do we go from here?
https://sec.edu.au/summit
Brendan Hopper - COMP6447 lecturer
Great opportunity if you are looking for a job!
20th September at UNSW in Scientia Building

http://secedusummit.eventbrite.com.au
Discount code: STUDENT6441

Courses

COMP6[84]43 - Web Application Security and Penetration Testing

- The website has a search bar for posts
  Bet you I can get it to give me a list of all email addresses
- We've signed in with Google Auth, we're safe
  But, like, how do you prove that you're the same user who signed in?
- We do all our work on a remote server, the user can't change anything
  I. Don't. Believe. You.

COMP6[84]45 - Digital Forensics and Incident Response

- I deleted my file, it's gone. Confidential data is deleted.
  Just because you can't see it, doesn't mean it's gone
- I'm throwing out my computer, I don't need it anymore
  Can I have it? Please?
- I got a virus!
  Where is it?
  What has it done?
  How can we remove it?

- Phones use flash memory, you can't recover data from that
  Come on, really???

COMP6447 - System and Software Security Assessment

- Here's a program
  Hack it
- Here's an Operating System
  Hack it
- Here's a…
  Hack it

binary, heap exploitation

COMP6448 - Security Engineering Masterclass

- Advanced class, aimed at post 6447
- Typically partners with a company to do advanced research
  Read as: hack whatever they say is 'unhackable'
- Runs over summer, when there is an available project

Security Project [AB] - COMP930[12]

- 9301 - 6 UoC
- 9302 - 12 UoC
- If you have anything security related in mind, come and get 6 UoC for it!
- Email Anatoli, CC Lachlan

COMP4337 - Securing Wireless Networks

- nothing to do with secEDU

We love having our students come back to teach, you guys are awesome

- HS1917 - 1511 taught to high school students
  HS1953 - 6441 taught to high school students
  Outreach - expanding to wider NSW
- 6[84]41 Tutoring