A Practical Guide to Cybercrime

Dr Richard Clayton
Director, Cambridge Cybercrime Centre

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

GOTO Berlin
17th November 2017

Outline

- Cambridge Cybercrime Centre
- Ransomware
- Business Email Compromise
- DDoS extortion
- Passwords

My background

- I’ve been looking at online abuse (spam, phishing, malware, DDoS etc) for two decades
- My general approach is data driven (I count things)
- I have obtained many datasets from industry under NDAs and that has underpinned the work I have done (in collaboration with some very smart people)
- BUT this is a long and tedious process, and we’re beginning to realise that no papers in this field can be reproduced (data cannot be shared, results cannot be compared, conclusions cannot be validated)
- This does not really look like science…

Cambridge Cybercrime Centre

- I have 5 years funding from EPSRC (+ some other money)
- Currently 6.5 of us
- We are interdisciplinary

  Computer Science, Criminology & Psychology
- Our approach is data driven. We aim to leverage our neutral academic status to obtain data and build one of the largest and most diverse datasets that any organisation holds
- We will mine and correlate this data to extract information about criminal activity. We will learn more about crime 'in the cloud', detect it better & faster and determine what forensics looks like in this space (and where appropriate work with LEAs)

Others can play too

- We have started the process of renegotiating existing NDAs
- We will collate the data, add value to it, put it into collections; and then make it available to others under one simple NDA agreement which is between the researcher and us
- We cannot make the data entirely public (or open) but we *will* be making it available to legitimate academics
- We will have a 'catalogue' of data that can be used in specialist research without the need to learn all about the web scraping, whois limits, duplicated data and all the other complexity
  - it will be easy to set MSc work in this area since it will not take 2 years to get the data together
  - we aim to see more *science* by letting people run different techniques on the same data and compare results

This is not a competition

- We will use this data in Cambridge – we will have world class researchers doing world class work
- BUT at the end of the first five years I want to be judged not on how many papers we wrote in Cambridge, but how many papers were written across the whole of academia because we helped to make it possible
- We (and the people using our datasets) will find new ways to prevent crime, to detect and deter criminals – and in the end, that's why society funds our work

https://www.cambridgecybercrime.uk/process.html

Cambridge Cybercrime Centre: Process for working with our data

This page sets out the steps in the process for obtaining data from the Cybercrime Centre.

Assess whether you will be allowed to use our data

Our datasets are intended for research and analysis into methods to find, understand, investigate and counter cybercrime so your project must clearly fall into this space. Although we do not require researchers to be academics, there are significant restrictions on using our data for commercial purposes.

Although some of our data was generated internally and so we can make it available for other types of project and for commercial purposes, much of our data has come from third parties and they have only provided us with the data because of the framework under which it will be shared.

Identify the data you wish to use

We describe our various datasets on this page [LINK]. The descriptions are public and necessarily fairly high level. We do however try to indicate the size of the datasets, the period over which they was collected, along with any known biases.

We strongly encourage the use of prepacked datasets rather than "live feeds". Although a live feed may be superficially attractive it makes it harder to arrange that other researchers can receive the same data that you did -- a key aim of the Cybercrime Centre is to enable reproducible research. If the issue is that you need to collect a further "field" over and above what we supply then talk with us and we may well be able to do this for you.

Read about our legal framework

It is important that you understand the basis on which we share data and the paperwork that will need to be signed.

There's several pages of explanations and FAQs about our agreements, starting here at https://www.cambridgecybercrime.uk/data.html, which you should read before contacting us.

Make an application

You will need to make a formal application to use our data. In the first instance you should send an email to the Director of the Cybercrime Centre,

Ransomware

- Executing a program on your machine encrypts your data, and you must pay for the decryption key to get your data back
- AIDS Trojan 1989 (arrived on a 5.25” disk!)
- Academic work on using public key dates from 1996
- Various attacks in 2005/2006
- First big success in 2013 with cryptolocker
  - requested payment in bitcoin (trying to avoid “follow the money”)
  - criminals put a lot of effort into customer support
  - lots of people paid up and almost all got their data back
- Now lots of variants, and lots of spam containing it
  - and some doesn’t work properly so can never decrypt
  - and some is badly designed (check out: nomoreransom.org)

Ransomware avoidance

- Use sophisticated anti-spam systems
  - your copy of SpamAssassin isn't good enough any more
- Run anti-virus
  - the problem is that criminals don't ship the malware until it evades AV, so this only works if you answer your email very slowly
- Don't click on attachments
  - training can help by reducing the number of people who open malicious attachments (not everyone understands the risk)
- Don't give everyone write access to every "share"
  - ransomware cannot encrypt what it cannot write
- Don't expose your database online
  - problems are occurring with MongoDB, MySQL etc
  - automated scanning means that unlikely just one encryption!

Practical thinking about ransomware

- Ransomware is just a Business Continuity Issue
  - same threat as a cup of coffee
  - same threat as a rogue employee
  - same threat as a burning building
  - same threat as a flooded datacentre
- Asking people not to click isn’t going to work
  - but do they need Word on the system ?
  - But do they need scripting enabled in PDFs ?
- The real fix is backups
  - which must not be writeable except when backing up!
  - preferably offsite
- Backups need to be tested and restore needs to be practised
  - even with backups it may take days to recover

Business Email Compromise (BEC)

- Several different types of attack
- Fake Invoice
  - email system compromised
  - replacement invoice issued with criminal's details
  - usually involves a "look alike" domain
    - arnazon.com, qotoberlin.de, netvvorksolutions.com
- CEO fraud
  - Please pay this supplier, I forgot before I left for the conference
    - don't ring, I'm in sessions all morning
    - lookalike domains here too (or verysimilar@webmailsystem.com)
- Can involve significant losses
  - $3m for a boatload of coal
  - $1m for a shipment of palm oil
    - FBI says $3000m since Jan 2015

BEC avoidance

- Check incoming email for look-alike domains
- Apply DMARC tests
  - i.e. check for SPF or DKIM passes
- Flag email where reply-to and from are different
- Set your email client to display <the@address.string>
  - major email clients rethinking this issue
  - still a problem on mobile
- Label email coming from outside the company
- Use S/MIME or PGP to validate email

Practical thinking about BEC

- Agree on bank account details up front
  - change contract to specify bank details not “as nominated”
  - agree at the start of your house purchase what accounts to use
- Don’t accept changes to payment destinations by email
  - insist all changes are made by snail mail, personal visit
- If a change is being made by email check it “out of band”
  - NB: use the phone number from the filing cabinet not the email !
- Share stories with your peers (& at dinner parties)
  - this type of fraud is not as well-known as it should be
- Empower your accounts department to say “NO”
  - No Purchase Order, No Payment!
  - pay a bonus for standing up to the fake CEO (and the real one)

DDoS extortion

We are Armada Collective

Your network will be DDoS-ed starting [date] if you don't pay protection fee – 10 Bitcoins @ [Bitcoin Address].

If you don't pay by [date], attack will start, yours service going down permanently price to stop will increase to 20 BTC and will go up 10 BTC for every day of attack.

This is not a joke.

Our attacks are extremely powerful - sometimes over 1 Tbps per second. And we pass CloudFlare and others remote protections! So, no cheap protection will help.

Do not reply, we will not read. Pay and we will know its you. AND YOU WILL NEVER AGAIN HEAR FROM US!

Bitcoin is anonymous, nobody will ever know you cooperated.

DDoS Extortion: what happens next

Practical thinking about DDoS Extortion

- Occasionally there will be a one hour DDoS against your website
- After that, nothing happens
- So DO NOT PAY

- If you care about a one hour DDoS then invest in an anti-DDoS solution of some kind !
- Share the email with the police, your trade body etc
  - if all the bitcoin wallets are the same then it’s fake
  - they may be able to link this with other attacks
  - … and bitcoin can be sometimes be traced

Passwords – the brute force analysis

- A safe is protected by a 4 digit code
  - so there are 10000 possible combinations
  - 10 to the power of 4
- How long until you open it ?
  - you might crack it at your first try
  - you might crack it after 9999 failures
  - on average it will take 5000 attempts
  - at 10 seconds per attempt it will take over 13 hours (on average)
- If 6 digits then 100 times longer (10^6 = over 8 weeks)
- If 4 alpha characters (A..Z) cracking time is over 26 days
- … and for 6 alpha characters it's 49 years
- If eight characters from A..Z a..z 0..9 then 62^8 = 34m years

Computer passwords

- Computer will compare password from user with stored secret
  - obvious risk if hold passwords in plain text
    - but you can send really good "reminders"
- Needham & Guy (1963) proposed 1-way hash
  - store Hash(secret)
  - hash user value and check for exact match
- Unix hash (from 1991) was 25 rounds of DES
  - more recently we use MD5, SHA-1, SHA-256 &c
- Reversing the one-way hashes is "impossible"
- But computers (& GPUs) are fast so "brute force" attack…
  - NVIDIA GeForce 8800 Ultra: 200m MD5's per second
  - so length 8 of A..Z falls in 8 minutes
  - so length 8 of A..Z a..z 0..9 falls in 6.3 days

Parallel cracking

- Easy to parallelise cracking tasks
  - split search n ways amongst n machines
  - dish out task blocks to these machines
  - good when machines different speeds
  - can all do a random search (this avoids communication costs and is tolerant of cheating)
- Hence usual metric is to consider number of hashes per dollar
- Alternatively one can try small number of passwords against many accounts
  - it may not matter which password is cracked, just how many
  - perhaps just one is sufficient [cf DirtyCow])

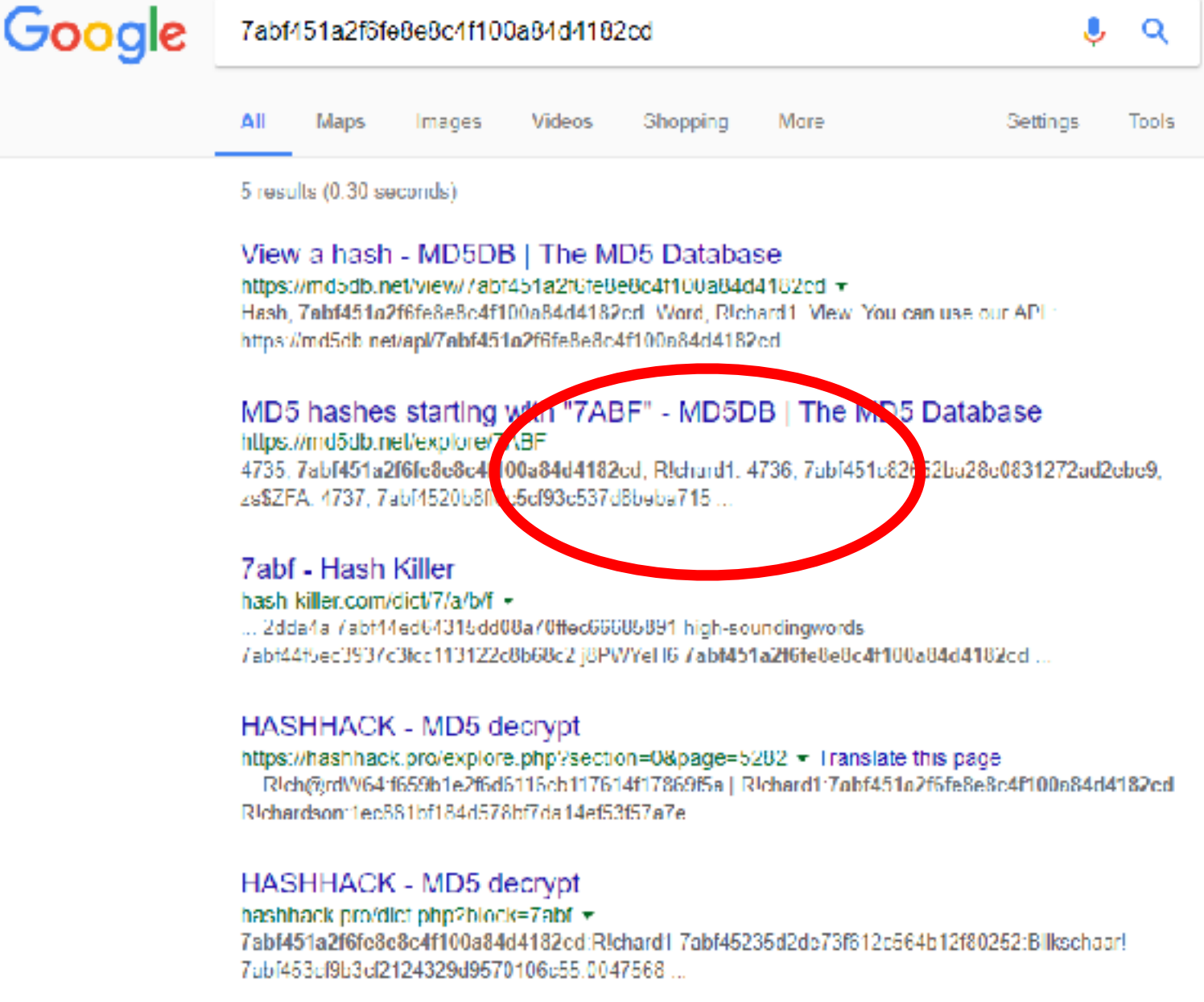
Real world hash cracking

- R!chard1

  MD5(R!chard1)

  7abf451a2f6fe8e8c4f100a84d4182cd

- Internet has many sites with hashes
  - the sets of tables are sometimes (incorrectly) called rainbow tables
  - see (e.g.) Wikipedia for details of how rainbow tables actually work


Google
7abf451a2f6fe8e8c4f100a84d4182cd
All
Maps
Images
Videos
Shopping
More
Settings
Tools
5 results (0.30 seconds)
View a hash - MD5DB | The MD5 Database
https://md5db.net/view/7abf451a2f6fe8e8c4f100a84d4182cd
Hash, 7abf451a2f6fe8e8c4f100a84d4182cd Word, Richard1 View You can use our API
https://md5db.net/api/7abf451a2f6fe8e8c4f100a84d4182cd
MD5 hashes starting with "7ABF" - MD5DB | The MD5 Database
https://md5db.net/explore/7ABF
4735, 7abf451a2f6fe8e8c4f100a84d4182cd, Richard1, 4736, 7abf451a82f52bd28c0831272ad2ebc9,
zs$ZFA, 4737, 7abf1520b8f5cf93c537d8beba715 ...
7abf - Hash Killer
hash killer.com/dict/7/a/bf
... 2dda4a 7abf41ed04315dd08a70ffed66006891 high-soundingwords
7abf44fbec0937c3fcc1131222d8b68c2 jUPWYeII6 7abf451a2f6fe8e8c4f100a84d4182cd ...
HASHHACK - MD5 decrypt
https://hashhack.pro/explore.php?section=0&page=5202 Translate this page
Rich@rdW64:f659b1e2f6d5115ch117614f17869f5a | Richard1:7abf451a2f6fe8e8c4f100a84d4182cd
Richardson:1ec881bf184d578bf7da14ef53f57a7e
HASHHACK - MD5 decrypt
hashhack.pro/dict.php?block=7abf
7abf451a2f6fe8e8c4f100a84d4182cd:Richard1 7abf45235d2db73f612c564b12f80252:Bilkschaar1
7abf453cf9b3cd2124329d9570106c55:3047568 ...

Modern protection: salts & slow hashes

- To protect against hash tables (and rainbow tables) passwords are usually "salted" with a random value chosen by the system
  - calculate hash(salt_value | password_text)
  - now attacker needs 2^s tables for each password (for s bits in salt)
  - salt is stored *en claire* next to the hashed password
  - rarely, for small salt lengths, the system brute forces it!
- Modern trend is towards hash functions designed to be slow to run on GPUs and ASICs
  - generally achieved by writing & reading lots of memory.
  - these hashes also tend to be tuneable (you select the number of interactions to provide the performance you need)
  - e.g. bcrypt, Argon2i

Practical thinking about passwords

- General Principle: you cannot assess a security solution without first determining what your "threat model" might be
- Are attackers Online or Offline ?
  - can attacker steal file and then do their cracking on their own kit ?
  - or do attackers have to present all their guesses to your systems ?
- Are attacks Targeted or Untargeted ?
  - does the attacker win if they crack one specific password, a percentage of passwords, or any password at all ?
- How much can you impose on your users
  - password length, character sets, system-chosen passwords?
- Are attackers local or remote
  - is writing down a password a disaster or somewhat desirable ?
- Why can't I deploy a two-factor solution ?

Latest advice on passwords

- New advice out of NIST (2017: 800-63B) on passwords
  - pay attention to length, not to character sets
- Change passwords only when compromise suspected
  - change rules usually imposed by auditors and not evidence based
  - in practice humans “cheat”
    - verysecret1, verysecret2, verysecret3…
    - Secret!Jan, Secret!Feb, Secret!Mar
- For offline attacks limit is size of attacker’s wallet
- Online, the system can set the limits, so make sure it does
  - <n> tries and then a timeout
    - better is that tries get exponentially slower (e.g. iPhone)
  - limits need to be
    - per account : one account being attacked
    - per IP: stop attacks on n accounts in parallel

What criminals do

- Given a password file, criminals brute force the passwords
  - if not encrypted then of course trivial
  - if not salted then may be just a lookup
  - otherwise use mangled dictionary approach
    - Passw0rd, pa$$word, Password, Password1
- Then they try the username/password combination everywhere
  - so after a merchant compromise they can attack email accounts, Skype, banking, Facebook etc etc
- Good guys are also brute force the passwords
  - force password change on users as needed
  - perhaps prevent future use of this password
- Around 90% of the 2012 Linkedin password leak were broken !
  - so file theft => all users must be assumed to be compromised

Some inconvenient truths

- Note that an incorrect password (or two) followed by the correct one is often a good indication of it being authorised human !
- Passwords that are always correct come from mobile phones!
- Passwords that are always incorrect come from mobile phones!
- Need to balance barring the bad guy with the risk of them perpetrating a denial of service attack on the account's legitimate owner
- Malware keyloggers…. game over!
- Compromise of clear text password file…. Game over!
  - and remember that 90% figure from LinkedIn
- Password managers are a good way forward, but they have a chequered history regarding their own security

Summary

- Cambridge Cybercrime Centre
  - driving a step change in cybercrime research in many disciplines
- Ransomware is merely a Business Continuity Issue
  - back up your data & practice restoring it
- Business Email Compromise is addressed by following procedures and empowering your accounts department
- DDoS extortion attempts should be ignored
- Passwords should be hashed and salted
  - and changed only when needed
- User should use long passwords, and not reuse their email or banking password elsewhere
  - complexity is a red herring, pay attention to attack techniques

blog: https://www.lightbluetouchpaper.org

Cambridge Cybercrime Centre

data: https://cambridgecybercrime.uk/process.html

UNIVERSITY OF CAMBRIDGE
Computer Laboratory