CSCI 4140
Quantum Computing and Cryptography

Mark Green

Faculty of Science

Ontario Tech

Introduction

- Start with an introduction to security
- All information has a value, what it’s worth to the person who wants to steal it
- There is also a cost associated with stealing the information
- We are mainly interested in the computational cost, but there are others
- The basic idea behind information security is to make the cost greater than the value of the information

Introduction

- Hacker don’t do this for the fun of it, they have a profit motive
- This profit may be monetary, but could also be social causes
- In addition, if hacking us is too expensive, they will look elsewhere
- The value of information also changes with time
- In some cases it’s only valuable for a few days or months
- In this case you need to be able to crack the codes quickly, or the value is gone
- This is typical of a lot of financial information

Introduction

- Most businesses must keep data for 7 years, usually for income tax purposes
- After 7 years this data can be destroyed, and even if its stolen it will not be of much use
- Some government data must be held secure for at least 50 years
- In some cases it can be released to the public after that time, in other cases it must still be kept secret

Introduction

- In the past this data was physical, for example, it was on paper
- All you needed to do was lock it away in a safe place, you just needed to worry about physical security
- With digital data the problems are quite different
- There can be many copies of the data and you may not be able to control its distribution
- Example: it can be intercepted when transferred over a network, you may not know that this has happened

Introduction

- The solution to this problem is to encrypt the data
- Now the question is how safe is the encrypted data?
- If it’s only short term, it is probably reasonably safe, but what about data that needs to be stored for 50 years
- Technology can change a lot in that time, new techniques could appear that make the encryption useless
- Suggested solution: unencrypt that data and then encrypt it with a better technique, think about this for a second

Introduction

- No, this doesn’t work, if someone has already copied the data, they have it with the old encryption and it can be hacked
- You need to think like a hacker, they don’t follow the rules
- Consider passwords, if we are lucky all the services that we use encrypt them
- If we have b different characters that can be used in a password, and they must be at least n characters long, there is the potential for $N=b^n$ different passwords
- How secure is this?

Introduction

- First thought: guessing passwords requires sequential search, from first year we know this will takes N/2 time on average
- If b and n are large enough this will take a long time
- But, no hacker does this, there are files of common passwords that are easily available
- The top few passwords are remarkably common
- If we just need to crack a few passwords out of a long list, this essentially becomes a constant time algorithm

Vectors – Risk Assessment

- How can we be attacked and where will quantum computers be a problem?
- Any encryption technique requires at least one key, these keys need to be communicated between the parties involved in the communications
- This is the weak point
- This occurs when you set up a https connection, uses the TLS protocol
- The current techniques are based on factoring large numbers and similar mathematical problems

Vectors – Risk Assessment

- We know that Shor's algorithm can solve these problems relatively quickly
- It's estimated that breaking the RSA 1024 algorithm requires approximately 2300 logical qubits and a about a day of computing time
- This is beyond current systems, but 5 or 10 years from now it could be quite possible
- What about QKD? It requires specialized hardware at both ends making it out of reach for most applications

Vectors – Risk Assessment

- Once we have a secure key we can use better encryption algorithms
- It is somewhat believed that these algorithms are relatively safe
- The best classical algorithms use brute force sequential search over the key space, which is currently 128 bits – 2^{128} keys
- This can be expanded to 192 and 256 bits
- Since it is sequential search it has been suggested that Grover's algorithm can be used
- Naively this will require 2^{64} iterations, in the worst case

Vectors – Risk Assessment

- The conclusion is that this won't happen any time soon
- Remember, this is a worst case number, assumes that all hackers have extremely bad luck
- Our own investigations indicate that Grover's algorithm doesn't always behave the way we expect it to
- Could there be more here that allows us to crack the code easier
- We have a piece of encrypted text and we are searching or a key that converts this into plain text

Vectors – Risk Assessment

- We need to have an oracle, but how do we build one?
- If we knew what the plain text was we could construct an oracle that produces 1 when we've found the correct key
- But, do we know what the plain text is? After all it's encrypted
- In some cases we do, with http all messages start with a header that is basically the same for all messages, we know what this is
- The same thing occurs in other Internet protocols, we can make a good guess of what should appear at certain points in the message

Vectors – Risk Assessment

- This gives us an oracle, but it still takes a long time, or does it?
- Remember what Grover's algorithm does, amplitude amplification
- At the beginning all N values are equally likely
- The location that we are interested in gets amplified on each iteration
- In our examples we've seen that even after one iteration there can be a significant change in amplitudes
- If we run for say 10 iterations, we won't have the solution, but we could have a small number of potential solutions with relatively high probability

Vectors – Risk Assessment

- The basic approach is as follows:
  - Run Grover's algorithm for a small number of iterations (20?)
  - Measure the result
  - On a classical computer, check whether the result decrypts the message, this is a fast operation
  - If not go back to the first step
- Since we are dealing with probabilities, we are likely to get a different result on each iteration, the correct result will be one of them
- This isn't guaranteed, but it could work well enough

Vectors – Risk Assessment

- If we have already cracked the key when it was distributed, we don’t even need to do this much work
- Another threat is certificates and digital signatures
- These are basically public key systems, there is a private key that is used to encrypt the certificate and then a public key is used to decrypt to check if it is valid
- These certificates are used by websites, how you can be sure that you are connecting to a trusted website

Vectors – Risk Assessment

- There are a small number of Certificate Authorities (CA) that issue these certifications and they are valid for a certain length of time
- If the secret key is determined the process is no longer reliable
- A hacker can create his own certifications that will appear to be legitimate, they could impersonate a bank
- RSA is used for this, so again Shor's algorithm is a way to attack this
- It will be about 10 years before this is a real problem

Vectors – Risk Assessment

- The last attack vector we will examine is passwords
- Hopefully passwords are kept in an encrypted form, some websites don’t bother to do this, we can offer no help here
- On most Unix-like systems encrypted password are stored in the /etc/passwd file, which can be read by any user
- When a user logins in their password is encrypted and checked against the copy in /etc/passwd
- If they match everything is okay

Vectors – Risk Assessment

- The idea is that going from the encrypted password to the plain text one is computationally difficult
- Does this hold up for quantum computers?
- Assume we have b=70 different characters and password of length n=10, this give us 70^{10} possible passwords
- Sequential search through this large space is impractical, but could we use Grover's algorithm?
- Remember we have one encrypted password, and many possible plain text ones

Vectors – Risk Assessment

- Constructing the oracle is relatively easy, it returns 1 when a plain text password encrypts to the encrypted password, we know the encryption algorithm – remember Linux is open source
- With b=70, we can easily encode each character in 7 qubits, with n=10, this requires 70 qubits total
- If Grover runs well enough on physical qubits, we will have a large enough quantum computer in 2021
- If we need logical qubits we may need to wait until 2025 or 2026
- This is getting pretty close

Vectors – Risk Assessment

- There are several things we can do in the near term
- One is to force passwords to be long, for example 20 characters
- There is a limit to what people will put up with
- Another approach is to ask users to submit 5 or 10 candidate passwords, then choose the one that produces the least Grover amplification
- Note, we may need a quantum computer to do this evaluation, since we currently can’t classically simulate a 70 qubit computer

Vectors – Risk Assessment

- The current wisdom is that the hashing codes that are used in software distribution are safe
- These hashes are used to detect changes that are made to the software after it has been packaged
- Used to prevent unwanted modification to program code
- I’ve not has the opportunity to look into this in any detail, but the arguments seem to be reasonable

Solutions

- The National Institutes of Standards (NIST) in the US is working on a new set of cryptography standards that will be safe from quantum attacks
- Need to have one set of standards that everyone agrees on that can be implemented on all of our software
- They are in the third round of evaluations, down to about a half dozen candidate standards
- You can find these candidates on their website https://csrc.nist.gov/Projects/post-quantum-cryptography/post-quantum-cryptography-standardization

Solutions

- These candidates are open for public comment
- I believe all of them have been implemented and you can download the code and try it out
- They hope to settle on a single standard by 2024
- These will need to be open standards that can be used by all software companies
- Need to be widely used as soon as possible

Deployment

- At first it looks like 2024 will be plenty of time
- The soonest we are expecting any problems is 2025 or 2026
- The problem is that we need to update a large number of computers and their software
- Clearly all the servers need to be updated, but all the clients will need to be updated to talk to them
- We can’t continue to support the old standard, since that will be a back door into the severs

Deployment

- So how long will this take?
- The current best estimate is 10 years
- This is based on the replacement of other encryption techniques in the past
- May need to split the Internet, one part using the new secure encryption, and the other part using the old scheme
- Shrink the old part over time until it is eliminated, or only has services that don’t need security

What Have We Missed?

- There are clearly cases that have been missed
- What about WIFI and cell phones?
- Cell phone standards evolve relatively quickly, so this may not be a major problem
- People tend to update their phones every few years, so older phones do tend to disappear relatively quickly, there are relatively few 10 year old phones
- So maybe we don’t need to worry very much about cell phones

What Have We Missed?

- WIFI is a different story
- This is a consumer technology, we expect WIFI enabled devices to be usable for a significant length of time
- Don’t want to buy a new high end fridge because its WIFI is out dated
- It will take much longer to update this infrastructure
- This is a major weak point, since WIFI is broadcast and can be intercepted outside of the home

Summary

- Examined some of the issues with computer security
- Examined some of the areas where our current systems are vulnerable to quantum computers
- Most of the potential problems are 6 to 10 years away
- We must move quickly to be prepared for this