

CS302 Information Security and Cryptography – Lecture 4

B Tech III Div. A and B
(Jan-Apr 2023) – 2 Feb 2023

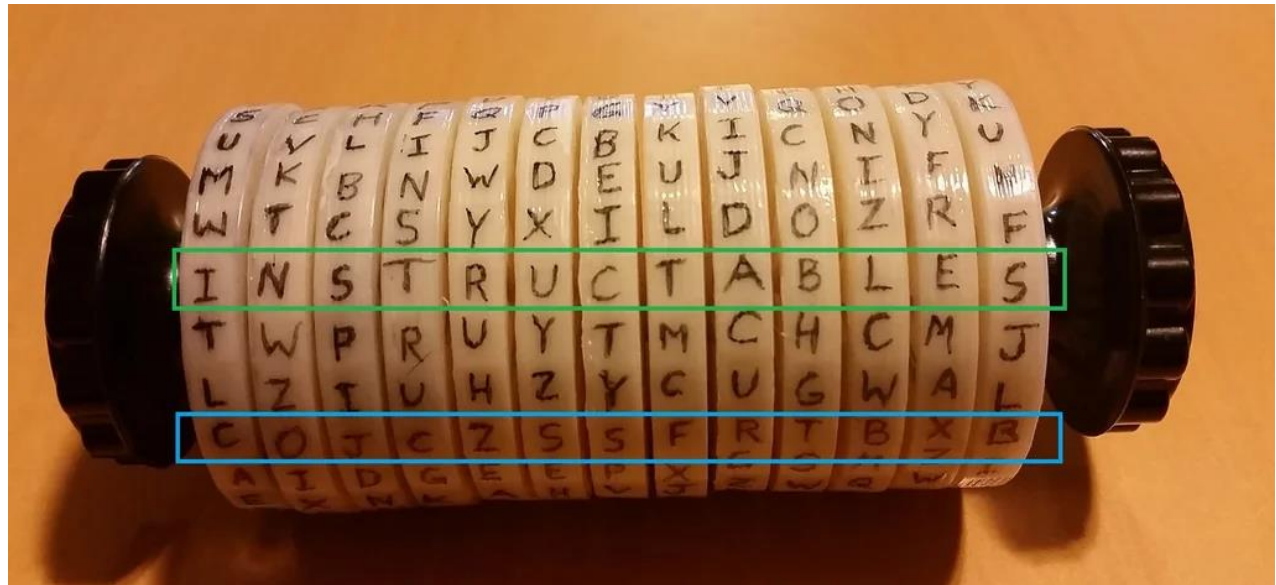
Dhiren Patel

Agenda

- Machine ciphers
- Machine ciphers (Mechanical)
- Machine cipher (ElectroMechanical)
- Steganography (Hiding? Camouflaged?)
- Changing Attack Landscape (Social Engineering Attacks (SEA))
- ChatGPT: Dark side

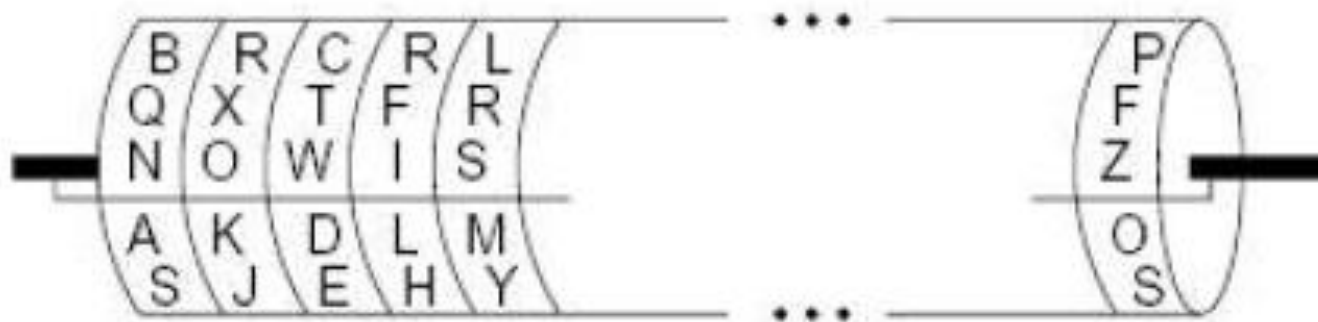
Wheel Cypher (Cipher)

- The Jefferson disk, also called the Bazeris Cylinder or wheel cypher as named by Thomas Jefferson



Machine cipher (Mechanical)

- The *Jefferson cylinder* implements a poly-alphabetic substitution cipher while avoiding complex machinery, extensive user computations, and Vigenère table.
- A solid cylinder 6 inches long is sliced into 36 disks. A rod inserted through the cylinder axis allows the disks to rotate. The periphery of each disk is divided into 26 parts.



Jefferson Cylinder

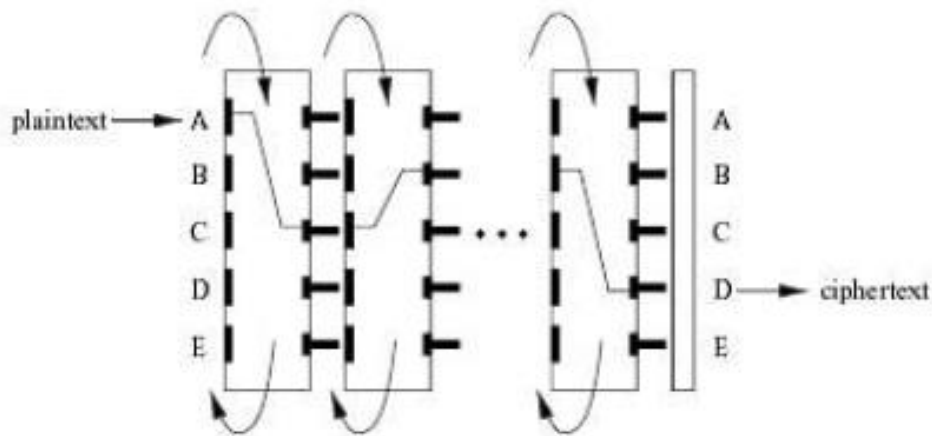
- On each disk, the letters A–Z are inscribed in a (different) random ordering.
- Plaintext messages are encrypted in 36-character blocks.
- A reference bar is placed along the cylinder's length.
- Each of the 36 wheels is individually rotated to bring the appropriate character (matching the plaintext block) into position along the reference line.
- The 25 other parallel reference positions then each define a ciphertext, from which, one is selected as the ciphertext to transmit.

Jefferson cylinder

- The ciphertext is decrypted by rotating each of the 36 disks to obtain characters along a fixed reference line matching the ciphertext.
- The other 25 reference positions are examined for a recognizable plaintext.
- Reordering disks (1 through 36) alters the polyalphabetic substitution key.
- The number of possible orderings is $36! \approx 3.72 \times 10^{41}$
- Both parties may agree beforehand on an index 1 through 25 specifying the offset between plaintext and ciphertext lines (or use additional shift cipher) to defeat an enemy with identical cylinder.

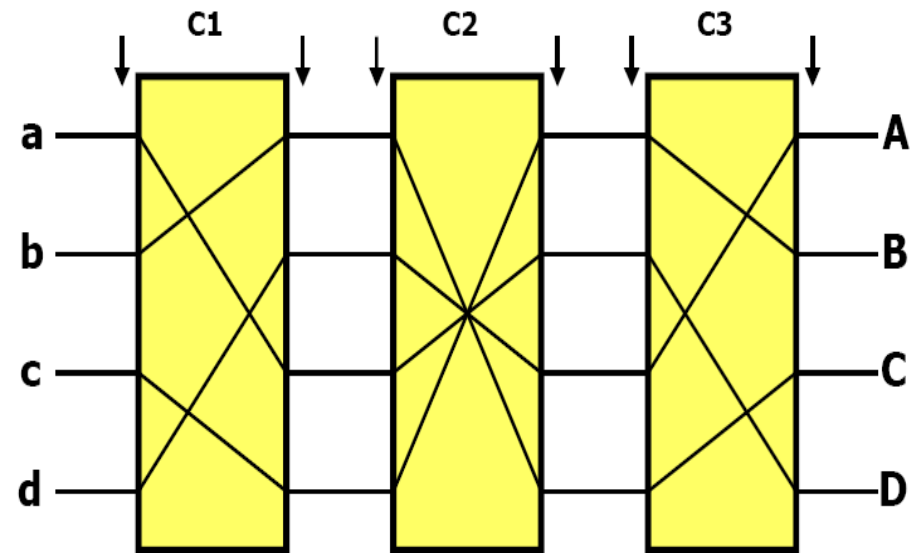
Rotor machines

- A plaintext character input to the first rotor generates an output which is input to the second rotor, and so on, until the final ciphertext character emerges from the last.
- A generic rotor machine consists of a number of *rotors* (*wired codewheels* through which electric pulses can flow) each implementing a different fixed mono-alphabetic substitution

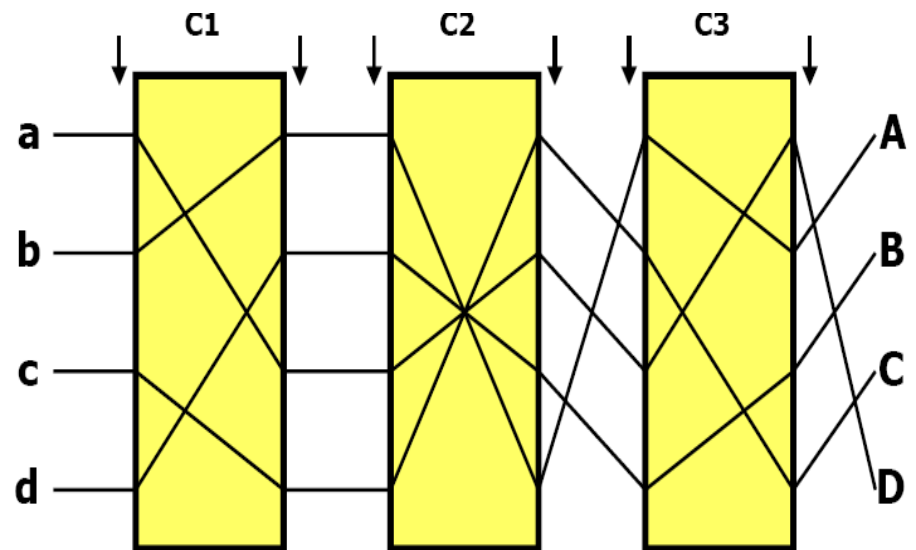


Rotor machines

- provide polyalphabetic substitution
- Internal cross-connections, providing a substitution using a continuously changing alphabet
- The cipher key is defined by the fixed wheel wirings and initial rotor positions
- Two properties desirable for security are: (1) long periods; and (2) state changes (concerns the motion of rotors relative to each other).



This setting maps: $a \rightarrow D$, $b \rightarrow C$, $c \rightarrow B$, $d \rightarrow A$

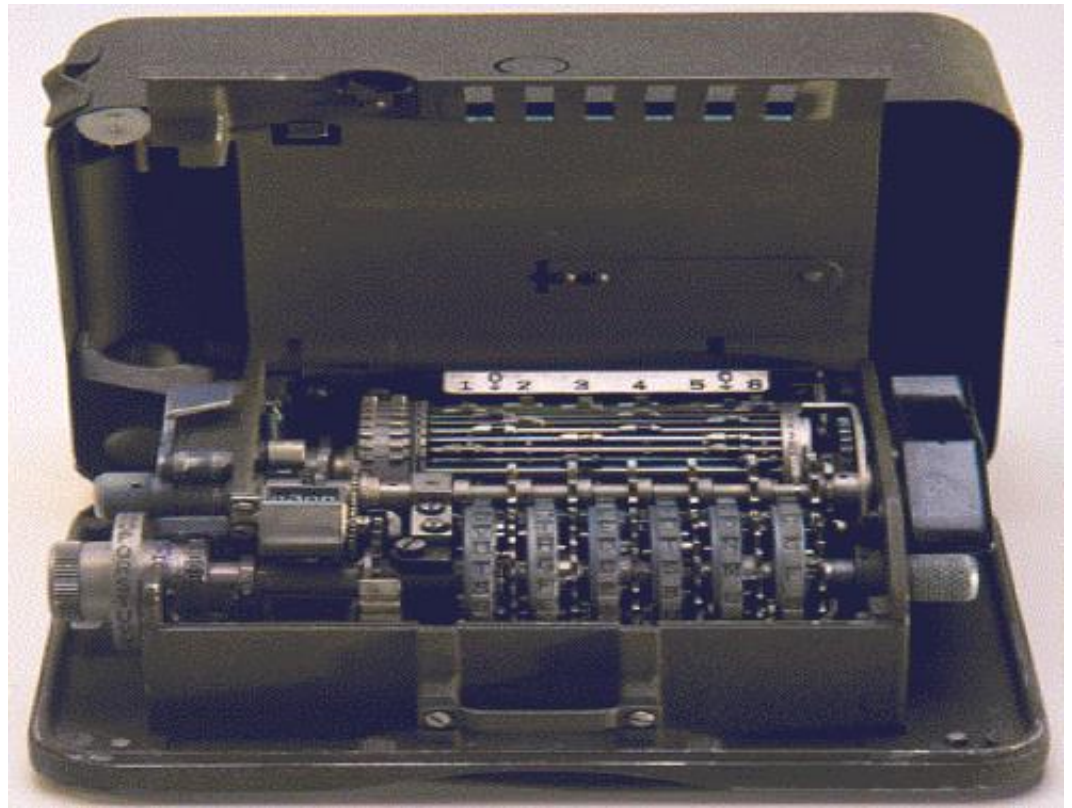


This setting maps: $a \rightarrow D$, $b \rightarrow A$, $c \rightarrow C$, $d \rightarrow B$

Rotor Machines

1. In indicating machines, ciphertext output characters are indicated by means such as lighted lamps or displayed characters in output apertures.
2. In printing machines, ciphertext is printed or typewritten onto an output medium such as paper.
3. With on-line machines, output characters are produced in electronic form suitable for direct transmission over telecommunications media.

Enigma and M209 Cipher



Enigma

- The Enigma machine is a cipher device developed and used in the early-to mid-20th century to protect commercial, diplomatic, and military communication.
- invented by German engineer **Arthur Scherbius** in 1918.
- It was employed extensively by Nazi Germany during World War II, in all branches of the German military.

Imitation Game (movie 2014)

- 1939, Alan Turing travels to Bletchley Park UK. Under the direction of Commander Alastair Denniston, he joins the cryptography team. The team analyze the Enigma machine, which the Nazis use to send coded messages.
- They could successfully decrypted 2 msgs per minute.
- Breaking Enigma played a huge role in ending the World war II.

Pocket code breaker (with Morse code)



Steganography: Information Hiding

- Steganography hides the existence of a message by transmitting information through various carriers. Its goal is to prevent the detection of a secret message.
- **Steganography** is the art and science of writing hidden messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message

The word *steganography*

- is of Greek origin and means "concealed writing"
- Cryptography
 - Conceals the meaning of message
- Steganography
 - Conceals the existence of messages

Steganography: Classical example

- Prisoners used a “tap code” to communicate
- Tap instruments...
- Availability in cell
- Tumbler
- Whistling
- Spoon (on anything)
- Fingers (on wall, on tumbler)
- Feet (on floor)
- Head (on wall)
- Sample code , . , . , , ,

Tap code

	1	2	3	4	5
1	A . .	B . ..	C, K	D	E
2	F .. .	G	H	I	J
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

- , . . , , => WAIT

Example – Spy message

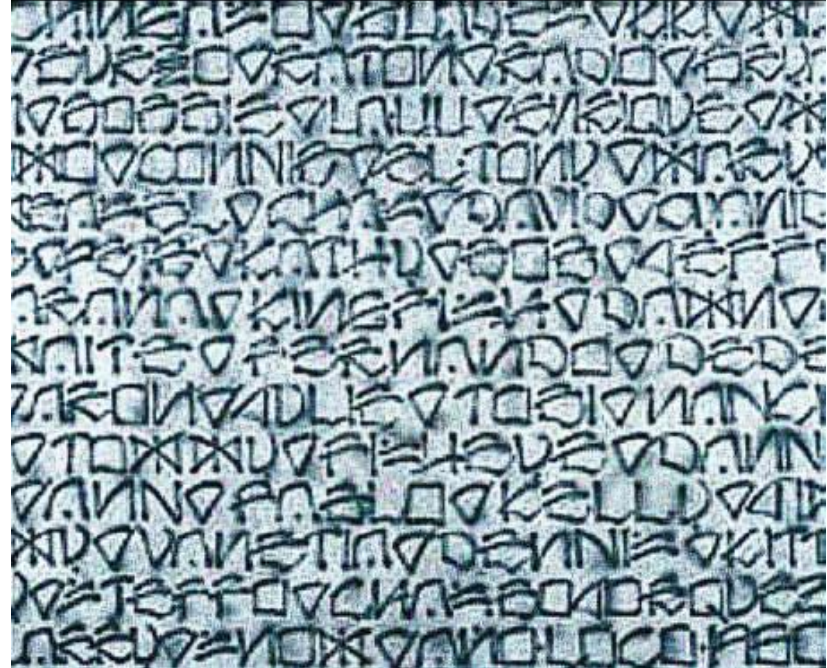
- Apparently neutral's protest is thoroughly discounted and ignored. Isman hard hit. Blockade issue affects pretext for embargo on byproducts, ejecting suets and vegetable oils.
- Trick = Collect second letter from each word

Hidden message within a message using Null Cipher

- Apparently neutral's protest is thoroughly discounted and ignored. Isman hard hit. Blockade issue affects pretext for embargo on byproducts, ejecting suets and vegetable oils.
- pershingsailsfromnyjune1
- pershing sails from ny june 1
- Pershing sails from NY June 1

More classical techniques

- Pin puncture,
- over written characters,
- writing using invisible ink,
- Modified alphabet



Modern Steganography

- In a digital message, the secret information is inserted or "hidden" into the "container data" <any type of digital data file>
- It doesn't appear to be anything other than what it is eg. A picture or music file.
- An encrypted file on the other hand cries out '*I contain sensitive information!!!*'
- Computer files (images, sounds recordings, even disks) contain unused or insignificant areas of data.
- Steganography takes advantage of these areas, replacing them with information.

More techniques

- Holography
- IR controls
- Pagers
- Colored glasses to filter wavelengths
- Inks – magnetic, thermo-chromic, photo-cromic
- Jargon speak
- Blank areas on Memory
- HTML code

Modern Techniques

- Least Significant Bit
- Low-Bit Encoding (image)
- Spread Spectrum
- Echo Data Hiding (audio)
- Perceptual Masking (audio–differential sounds)
- Discrete Cosine Transform (video)
- Textual steganography (spacing, coding)
- unused space in the packet headers

Original



New



Digital Watermarking

- Steganography can also be used to place a hidden "trademark" in images, music, and software
- Using a variety of techniques - images, music, movies can be imprinted with digital watermarks.
- Digital watermarks - used as a means of preventing modification

Digital Watermarking

- Method of embedding information into a digital signal in a way that is difficult to remove.
- Visible Watermarking: Information is Visible
e.g. Company logo embedded in an image.

Invisible Watermarking

- Information is added as a digital data in an audio, picture or video that cannot be perceived as such
- Applications....
- Protection of Data alteration.
- Access Control System for digital content Distribution
- E.g. apple music store

Changed attack landscape (Privacy? Social Engineering Attacks)

Social Engineering

- Social engineering is a manipulation technique that exploits human error to gain private information, access, or valuables.
- Scams based on social engineering are built around how people think and act.
- As such, social engineering attacks are especially useful for manipulating a user's behavior. Once an attacker understands what motivates a user's actions, they can deceive and manipulate the user effectively.

The Attack Cycle

- **Prepare** by gathering background information on you or a larger group you are a part of.
- **Infiltrate** by establishing a relationship or initiating an interaction, started by building trust.
- **Exploit the victim** once trust and a weakness are established to advance the attack.
- **Disengage** once the user has taken the desired action.

SEA

- Many employees and consumers don't realize that just a few pieces of information can give hackers access to multiple networks and accounts.

Traits of SEA

- Among most attacks, you'll find yourself being misled into the following behaviors:
- **Heightened emotions** : Emotional manipulation gives attackers the upper hand in any interaction.
- You are far more likely to take irrational or risky actions when in an enhanced emotional state.
- **Urgency**: Time-sensitive opportunities or requests are another reliable tool in an attacker's arsenal.
- You may be motivated to compromise yourself under the guise of a serious problem that needs immediate attention.
- Alternatively, you may be exposed to a prize or reward that may disappear if you do not act quickly. Either approach overrides your critical thinking ability.

SEA

- **Trust:** Believability is invaluable and essential to a social engineering attack.
- Since the attacker is ultimately lying to you, confidence plays an important role here.
- They've done enough research on you to craft a narrative that's easy to believe and unlikely to rouse suspicion.

Phishing Attacks



Baiting Attacks

- **Baiting** abuses your natural curiosity to coax you into exposing yourself to an attacker. Typically, potential for something free or exclusive is the manipulation used to exploit you. The attack usually involves infecting you with malware.

Phishing Attacks

- **Phishing** attackers pretend to be a trusted institution or individual in an attempt to persuade you to expose personal data and other valuables.
- **Spam phishing**, or mass phishing, is a widespread attack aimed at many users. These attacks are non-personalized and try to catch any unsuspecting person.
- **Spear phishing** and by extension, **whaling**, use personalized info to target particular users. Whaling attacks specifically aim at high-value targets like celebrities, upper management, and high government officials.

SEA - phishing

- **Voice phishing ()** phone calls may be automated message systems recording all your inputs.
- **SMS phishing ()** texts or mobile app messages might include a web link or a prompt to follow-up via a fraudulent email or phone number.
- **Email phishing** is the most traditional means of phishing, using an email urging you to reply or follow-up by other means. Web links, phone numbers, or malware attachments can be used.

SEA - phishing

- **Angler phishing** - an attacker imitates a trusted company's customer service team. He can intercept your communications with a brand to hijack and divert your conversation into private messages, where they then advance the attack.
- **Search engine phishing** attempt to place links to fake websites at the top of search results. These may be paid ads or use legitimate optimization methods to manipulate search rankings.

SEA

- **URL phishing** links tempt you to travel to phishing websites. These links are commonly delivered in emails, texts, social media messages, and online ads. Attacks hide links in hyperlinked text or buttons, using link-shortening tools, or deceptively spelled URLs.
- **In-session phishing** appears as an interruption to your normal web browsing. For example, you may see such as fake login pop-ups for pages you're currently visiting.

Physical Breach Attacks

- **Physical breaches** involve attackers appearing in-person, posing as someone legitimate to gain access to otherwise unauthorized areas or information.
- Attacks of this nature are most common in enterprise environments, such as governments, businesses, or other organizations.
- Attackers may pretend to be a representative of a known, trusted vendor for the company.

SEA

- **Pretexting Attacks**
- **Pretexting** uses a deceptive identity as the “pretext” for establishing trust, such as directly impersonating a vendor or a facility employee.
- **Access Tailgating Attacks**
- **Tailgating** , or piggybacking, is the act of trailing an authorized staff member into a restricted-access area. Attackers may play on social courtesy to get you to hold the door for them or convince you that they are also authorized to be in the area.

SEA

- **Quid Pro Quo Attacks**
- **Quid pro quo** is a term roughly meaning “a favor for a favor,” which in the context of phishing means an exchange of your personal info for some reward or other compensation. Giveaways or offers to take part in research studies might expose you to this type of attack.
- The exploit comes from getting you excited for something valuable that comes with a low investment on your end.

SEA

- **DNS Spoofing and Cache Poisoning Attacks**
- **DNS spoofing** manipulates your browser and web servers to travel to malicious websites when you enter a legitimate URL. Once infected with this exploit, the redirect will continue unless the inaccurate routing data is cleared from the systems involved.
- **DNS cache poisoning attacks** specifically infect your device with routing instructions for the legitimate URL or multiple URLs to connect to fraudulent websites.

SEA

- **Scareware Attacks**
- **Scareware** is a form of malware used to frighten you into taking an action. This deceptive malware uses alarming warnings that report fake malware infections or claim one of your accounts has been compromised.
- **Watering Hole Attacks**
- **Watering hole attacks** infect popular webpages with malware to impact many users at a time. It requires careful planning on the attacker's part to find weaknesses in specific sites.

SEA

- When malware creators use social engineering techniques, they can lure an unwary user into launching an infected file or opening a link to an infected website.
- **Worm Attacks**
- The cybercriminal will aim to attract the user's attention to the link or infected file – and then get the user to click on it.

SEA

- **Malware Link Delivery Channels**
- the message will usually contain eye-catching or intriguing words that encourage the unsuspecting user to click on the link. This method of penetrating a system can allow the malware to bypass the mail server's antivirus filters.
- **Peer-to-Peer (P2P) Network Attacks**
- P2P networks are also used to distribute malware. A worm or a Trojan virus will appear on the P2P network but will be named in a way that's likely to attract attention and get users to download and launch the file.

??

- **Shaming Infected Users out of Reporting an Attack**
- the victim will be keen to avoid disclosing their own illegal intentions

How to Spot Social Engineering Attacks?

- self-awareness – about privacy and security; slow down and think before doing anything or responding
- **Are my emotions heightened?** When you're especially curious, fearful, or excited, you're less likely to evaluate the consequences of your actions.
- **Did this message come from a legitimate sender?** Inspect email addresses and social media profiles carefully when getting a suspect message.
- **Did my friend actually send this message to me?** It's always good to ask the sender if they were the true sender of the message in question.

Spot and Defend

- **Does the website I'm on have odd details?**
- Irregularities in the URL, poor image quality, old or incorrect company logos, and webpage typos can all be red flags of a fraudulent website.
- **Does this offer sound too good to be true?**
- You should consider why someone is offering you something of value for little gain on their end.
- **Attachments or links suspicious?**
- If a link or file name appears vague or odd in a message, reconsider the authenticity of the whole communication.
- **Can this person prove his/her identity?**

Defend, Protect

- Safe Communication and Account Management Habits
- Never click on links in any emails or messages
- Use multi-factor authentication
- Use strong passwords (and a password manager)
- Avoid sharing names of your schools, pets, place of birth, or other personal details.
- Be very cautious of building online-only friendships

Safe Network Use Habits

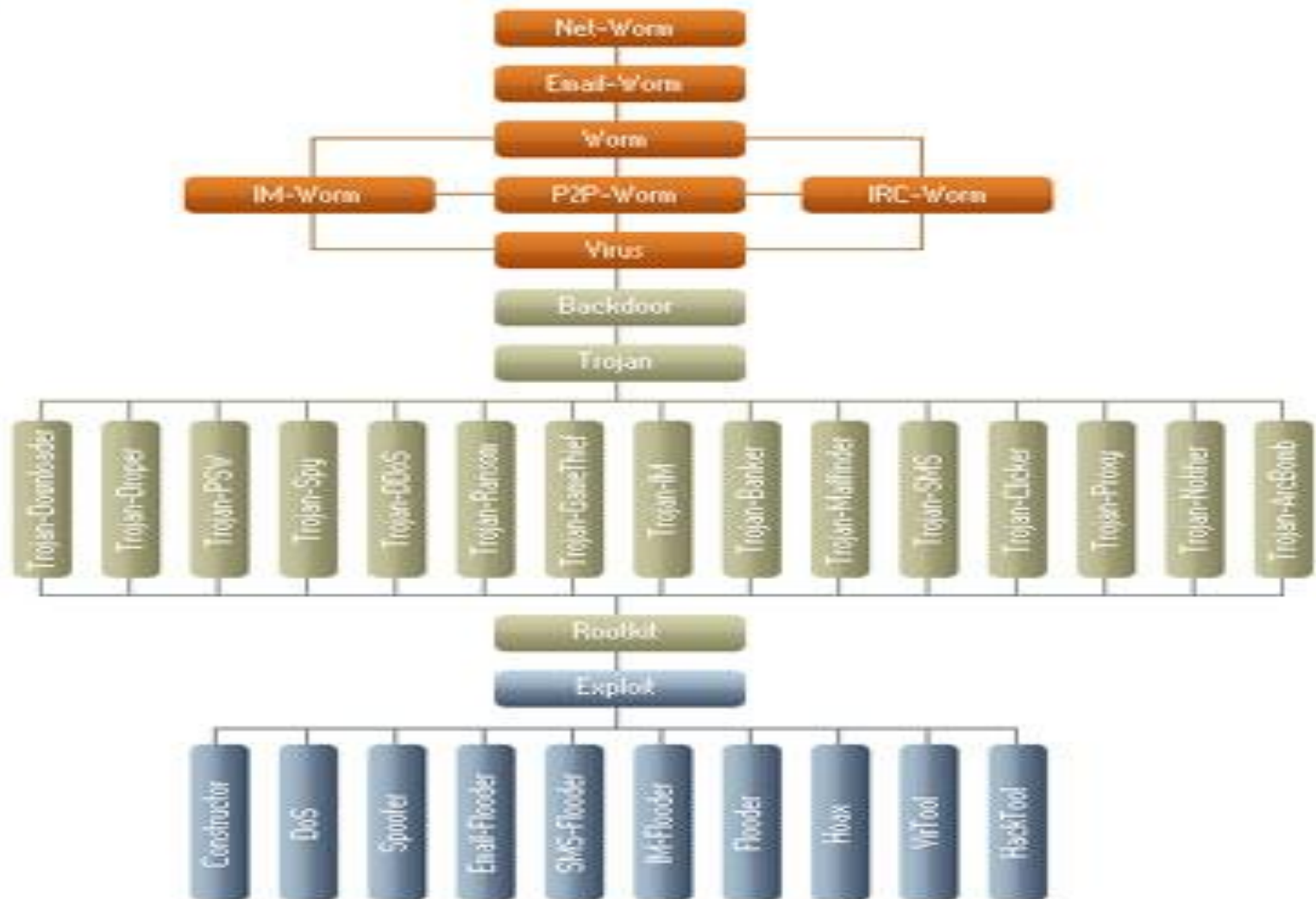
- Never let strangers connect to your primary Wi-Fi network
- Use a VPN
- Keep all network-connected devices and services secure

Safe Device Use Habits

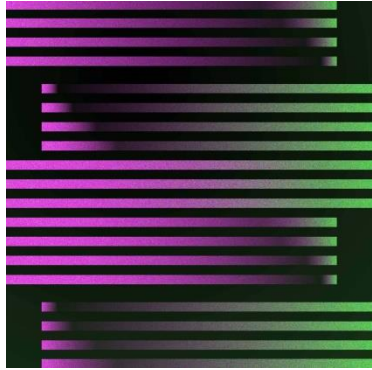
- Use comprehensive internet security software
- Don't ever leave your devices unsecured in public
- Keep all your software updated as soon as available
- Check for known data breaches of your online accounts

Malware classification

- Distinguishing and classifying different types of malware from each other is important to better understanding how they can infect computers and devices, the threat level they pose and how to protect against them.
- **"classification tree"**
- **Malware types with multiple functions**
- **Multiple functions with equal threat levels**



ChatGPT



- ChatGPT ChatBot OpenAI (iterative deployment of increasingly safe and useful AI systems)
- ChatGPT: Optimizing Language Models for Dialogue

What is the temp of Castle Rock CO?

- I am sorry, I am a language model and have no way of accessing current information such as weather data. My knowledge cut-off is from 2021 and I do not have access to current information such as weather or temperature of any location.

which wallet has highest bitcoin?

- The "Satoshi Nakamoto" wallet: This wallet is believed to be controlled by the creator of Bitcoin, Satoshi Nakamoto. It is estimated to contain about 1 million Bitcoin, which would make it one of the largest wallets in existence. However, it's worth noting that the true identity of Satoshi Nakamoto is unknown and there is no way to confirm that this wallet truly belongs to him.

???

- Welcome to (ChatBot open AI) - **the fastest-growing consumer internet app ever**
- The chatbot ChatGPT gives answers which are grammatically correct and read well
- ChatGPT is a chatbot launched by OpenAI in November 2022. It is built on top of OpenAI's GPT-3 family of large language models, and is fine-tuned with both supervised and reinforcement learning techniques.
- ChatGPT reaches 100 million users two months after launch (It had about 590m visits in January 2023 from 100 million unique visitors)

- By comparison it took TikTok about nine months after its global launch to reach 100 million users and Instagram more than two years
- The growing usage, while imposing substantial computing cost on OpenAI, has also provided valuable feedback to help train the chatbot's responses

ChatGPT

- ChatGPT interacts in a conversational way.
- The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.
- this model is trained using Reinforcement Learning from Human Feedback (RLHF)
- an initial model is trained using supervised fine-tuning: human AI trainers provided conversations in which they played both sides—the user and an AI assistant

ChatGPT's Dark Side: An Endless Supply of Polymorphic Malware

- OpenAI's popular new AI tool ChatGPT can be used to create polymorphic malware.
- “[ChatGPT]’s impressive features offer fast and intuitive code examples, which are incredibly beneficial for anyone in the software business,”
- its ability to write sophisticated malware that holds no malicious code is also quite advanced

ChatGPT's Dark Side

- the ChatGPT API could be leveraged within the malware itself, delivering modules to perform different actions as needed. “This results in polymorphic malware that does not exhibit malicious behavior while stored on disk and often does not contain suspicious logic while in memory,”
- “This high level of modularity and adaptability makes it highly evasive to security products that rely on signature-based detection and will be able to bypass measures such as Anti-Malware Scanning Interface (AMSI),”

ChatGPT's Dark Side

- ChatGPT can take the code produced and repeatedly mutate it, creating multiple versions of the same threat.
- “By continuously querying the chatbot and receiving a unique piece of code each time, it is possible to create a polymorphic program that is highly evasive and difficult to detect,”

Summary

- Security Technology, Secure System, Trust??
- Lectures provide you broad insights and awareness of security landscape, examples, use cases, scope, challenges and opportunities!!