Data and Network Security

(Master Degree in Computer Science and Cybersecurity)

Lecture 14

SAPIENZA
UNIVERSITÀ DI ROMA

Outline for today

- Recap last lecture
- Decoy files/Fake document generation

Outline for today

- Recap last lecture
- Decoy files/Fake document generation

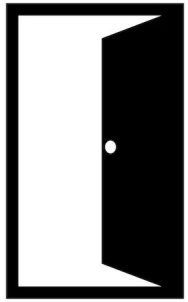
Authentication methods

Authentication is the process of verifying the identity of a user.

Objectives:

- Establish trust
- Prevent unauthorized access
- Protect sensitive information

Authentication methods - Actors

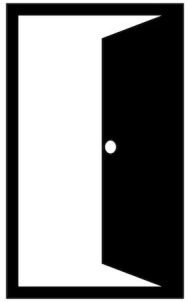


Claimant

Monitor

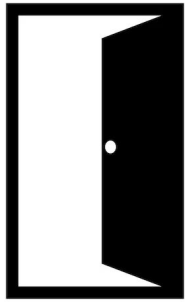
Information system

Authentication methods - Actors - Claimant



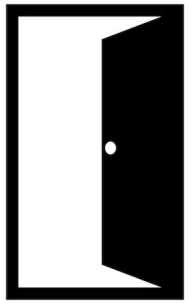
Claimant: entity that **authenticates** to the system, in order to use the services.

Authentication methods - Actors - Monitor



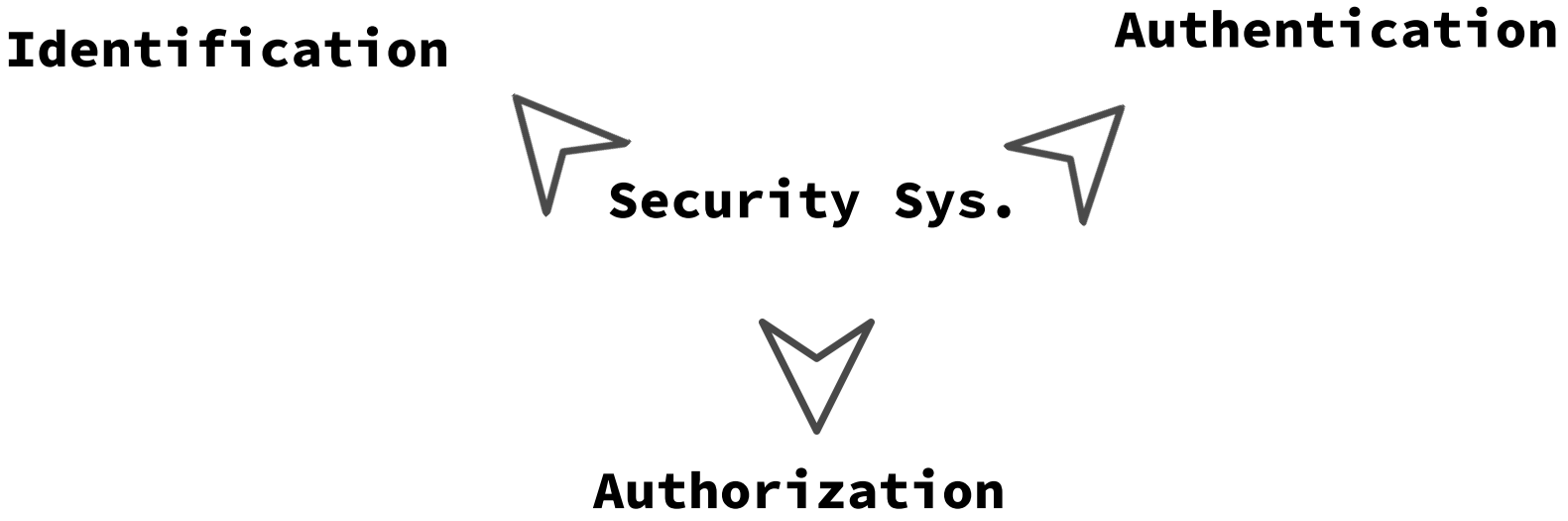
Monitor: the entity that provides an authentication service. It asserts the identity of a claimant and checks if it can grant him/her the use of the required service.

Authentication methods - Actors - Information System



Information System: provides services, such as an access to a computer account, an application, a door unlocking or a network printer, and will let the claimant use its services if the monitor correctly authenticated it

The core of a security system



The core of a security system

- Identification: Communicate your identity to the IS
- Authentication: Proof given by a claimant to assert by a monitor that claimant really corresponds to the identity he provided.
- Authorisation: The granted privileges to the claimant

The core of a security system - another crucial thing

We need a way to connect the claimant and the monitor:

Channel: is a support of communication between the claimant and the monitor.

It can either be considered as:

- confidential,
- authentic,
- secure
- insecure.

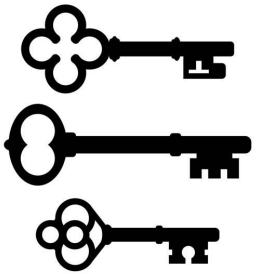
The core of a security system - another crucial thing

Channel is:

- **Confidential** - if it is interception resistant
- **Authentic** - if it is resistant to tampering
- **Secure** - resistant to both interception and tampering
- **Insecure** - obviously none

Authentication types - categories

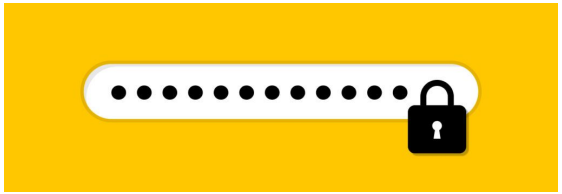
1. Ownership model (you own something)
    1. Physical keys
    2. Smart card
    3. NFC
    4. RFID
    5. Hardware-token

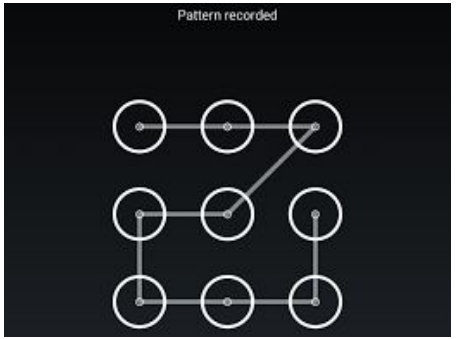


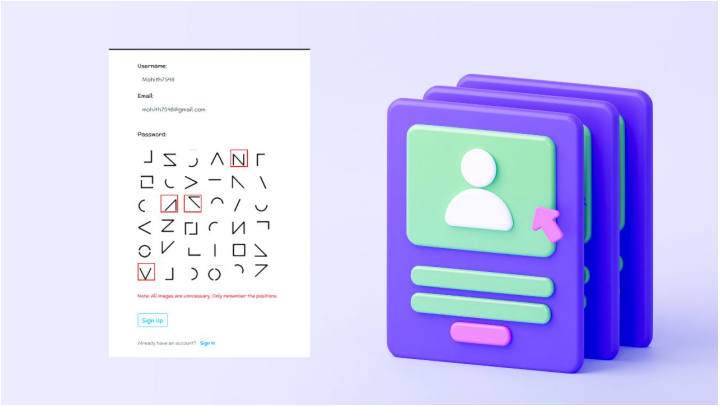
Authentication types - categories

2. Knowledge-based model:

    1. Passwords
    2. PIN code
    3. Lock pattern
    4. Graphical password
    5. Rhyme based
    6. Challenge response

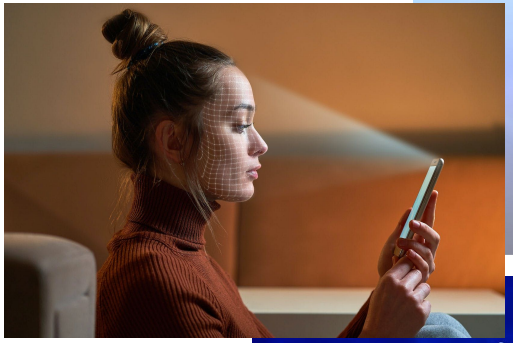


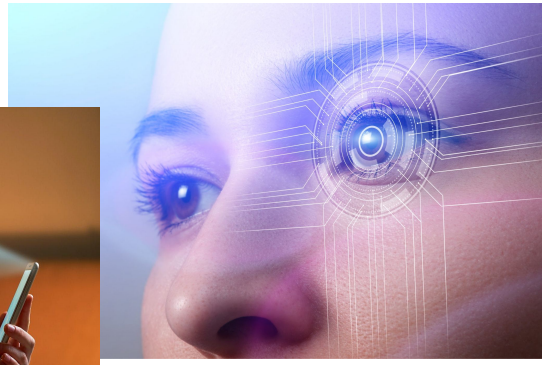


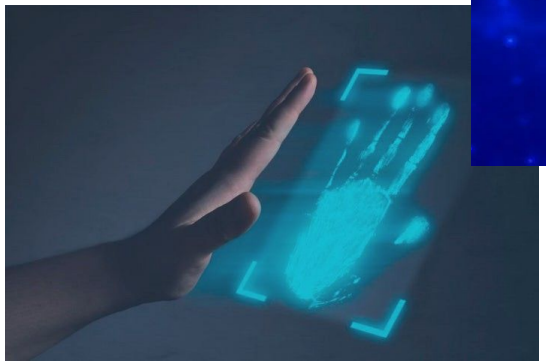


Authentication types - categories

3. Inherent-based model
    1. Fingerprints
    2. Palm
    3. Iris
    4. Voices
    5. Gestures
    6. Face





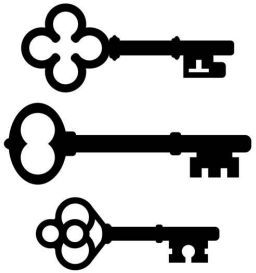


Authentication types - Issues/Drawbacks

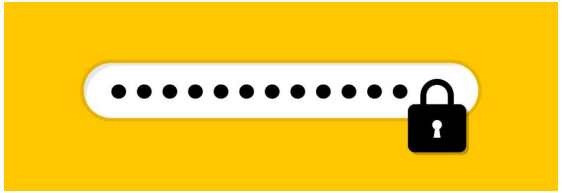
1. Ownership model (you own something)

    1. Losing it
    2. Token stealing
    3. High cost
    4. MITM attack
    5. Usability

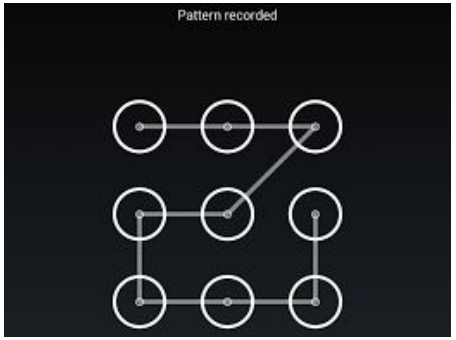
RFID



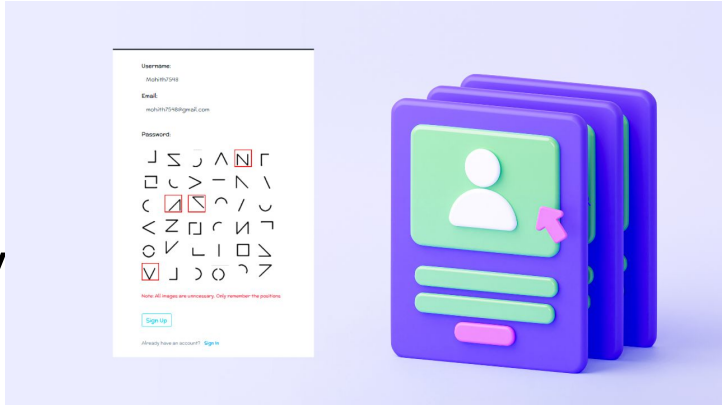
Authentication types - Issues/Drawbacks



2. Knowledge-based model:

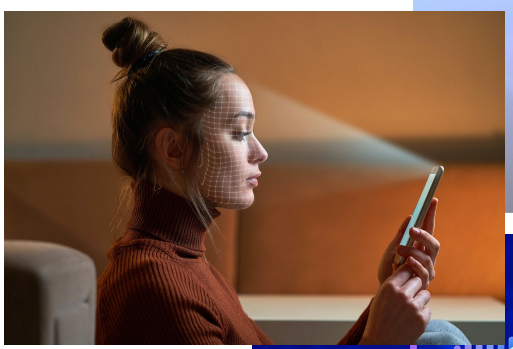


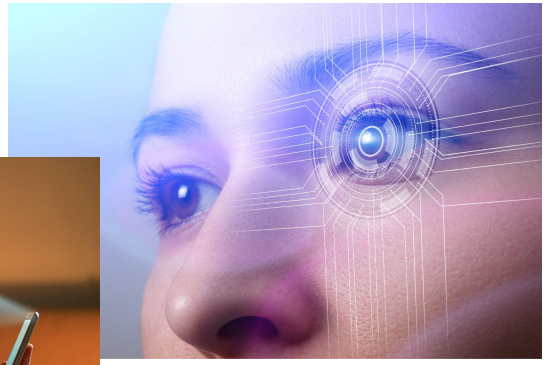
1. Keylogging
2. Shoulder surfing
3. Guessing
4. Brute force
5. Dictionary attacks
6. Screen Capturing
7. MITM attack
8. Memorization capability

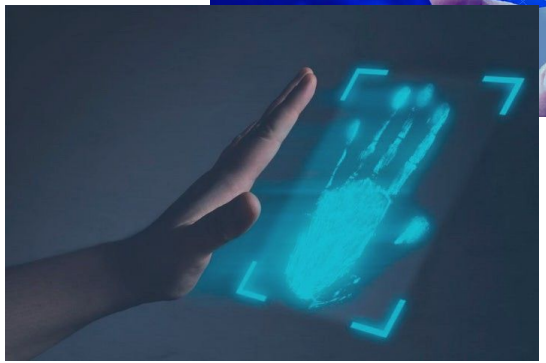


Authentication types - Issues/Drawbacks

3. Inherent-based model
    1. Forgery
    2. Replay
    3. MITM
    4. High costs
    5. Accuracy
    6. Medical (scars)
    7. Lighting conditions
    8. Clothes/Jewelry







Authentication types - Nowadays

Increase in cyber security threats

Authentication systems are being more empowered than before

(Non) Conventional attempts at multi-factor authentication

- Graphical centered systems
- Touch based
- EEG (electroencephalogram) based
- Web-based

Graphical centered systems

Various types of images or shapes are used as password.

- Images can be easily remembered by human than text.
- Human brains can process images easily.
- Utilizing images renders it resistant to dictionary attack, keylogger, social engineering etc.

Two types of graphical password techniques:

- Recognition based
- Recall based

Graphical centered systems

Two types of graphical password techniques:

- **Recognition based**
- **Recall based**

Recognition based - various images are presented to the user and from that user has to recognize the right images in a correct sequence.

Recall based - user has to reproduce something that he/she has created or selected during registration.

Graphical centered systems - some examples

Enrollment phase:

- Users choose some images from a set of 25 pictures
- And after that users are presented with 3 questions and they should select 3 points region-of-answer

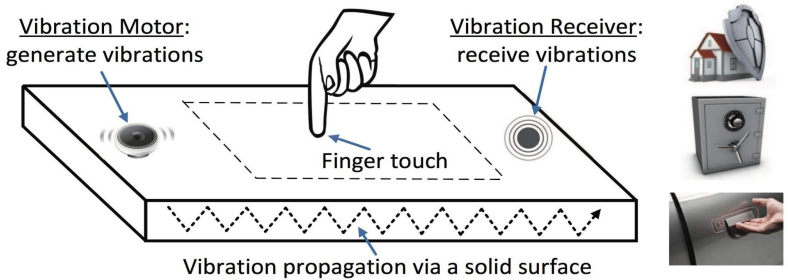
Authentication phase:

- Users should select correct images respectively in the first phase of authentication then they should select three regions of the preselected pictures as for the next step.

Graphical centered systems - some examples

Vibration and pattern:

To pass the authentication step, the user must select the same number of cells and feel the same number of vibration code as he had done in the registration phase.



Touch based systems

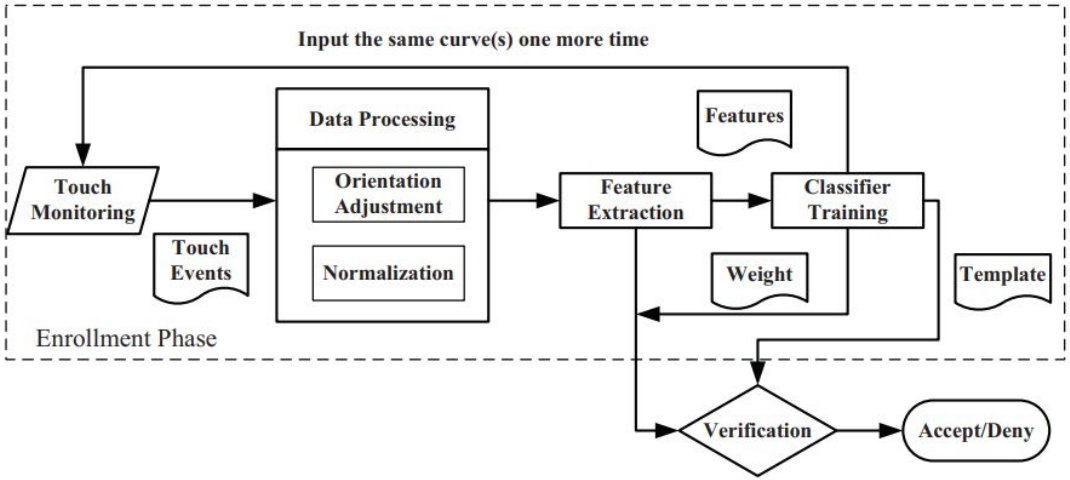
The user is authenticated based on the geometric properties of his drawn curves as well as his behavioral and physiological characteristics.

- No need to look at the screen

Touch based systems - one approach

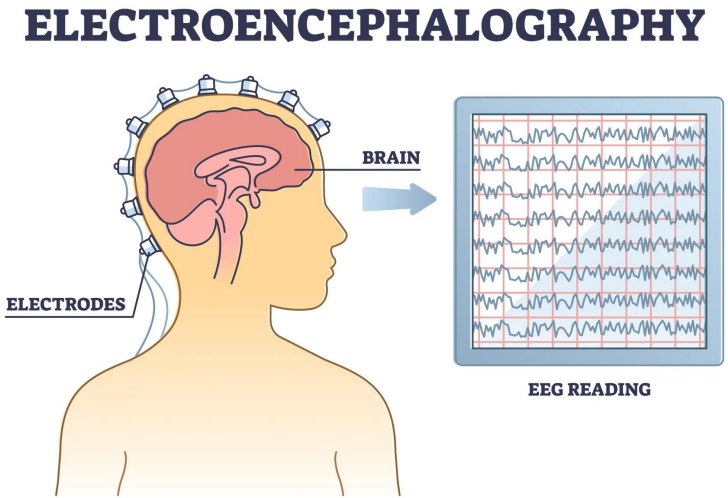
Authentication phase:

- anyone attempting to unlock the mobile device needs to draw on the multi-touch screen, from which a candidate template is extracted.
- If the candidate and authentication templates match, the user is allowed in.



EEG (electroencephalogram) based systems

An **electroencephalogram** (EEG) is a test that measures electrical activity in the brain using small, metal discs (electrodes) attached to the scalp. Brain cells communicate via electrical impulses and are active all the time, even during sleep. This activity shows up as **wavy lines** on an EEG recording.



EEG (electroencephalogram) based systems

Purpose: Strengthen password based authentication to prevent possible reply attacks by demanding also a certain mental state of the user that is entering the password.

Possible issues:

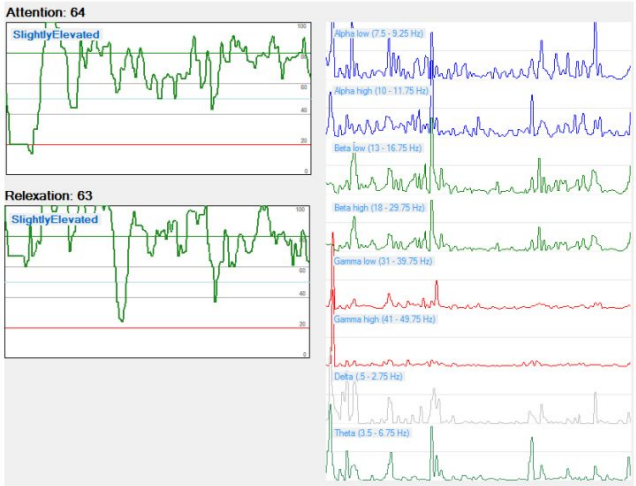
- The main challenge of this research is recreating the brainwave. EEG as a biometric characteristic lacks consistency which depends on **stress**, **fatigue**, **medication**, **environment** (electrical equipment), etc. To cope with this, researchers often use some kind of **stimuli** to help in recreating the valid authentication EEG pattern (imagining letters, text, somebody you care etc).
- EEG devices are not cheap.

EEG (electroencephalogram) based systems

Authenticate with cheap EEG:

Cheap EEG devices are proved effective is recognising the mental state of the user

- using alpha and beta waves to determine users' relaxation and concentration
- Can not measure effects caused by thought (pass-thought)
- **EEG signals that it provides can be used as additional parameters for a password**



Outline for today

- Recap last lecture
- Decoy files/Fake document generation

Cyber Deception

Cyber deception is a cybersecurity strategy designed to mislead and manipulate attackers who are attempting to breach an organization's network or systems.

It involves the deliberate creation of false information, decoys, traps, and other deceptive elements to divert, confuse, or delay attackers, ultimately enhancing the security posture of the organization.

Cyber Deception - some components

1. Decoys
2. Honeypots
3. Deceptive data
4. ….

Cyber Deception - decoys

Decoys are simulated assets, such as servers, workstations, databases, or entire networks, designed to mimic real systems and services within an organization's infrastructure.

- Honeypots are specialized decoy systems that are intentionally left vulnerable to attract and monitor attackers.

Cyber Deception - deceptive data

False or misleading information strategically placed within the network to mislead attackers. This may include fake user credentials, fabricated network paths, or bogus file directories.

Cyber Deception -Rationale

Corporations and government organizations have to take into account that one day their systems might be hacked:

- Protect sensitive information leakage
- Slow down the attacker
- Confuse the attacker

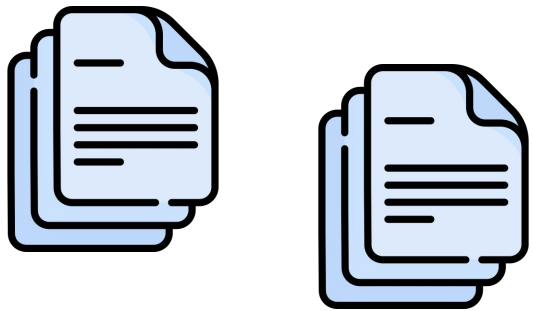
Cyber Deception - fake document generation

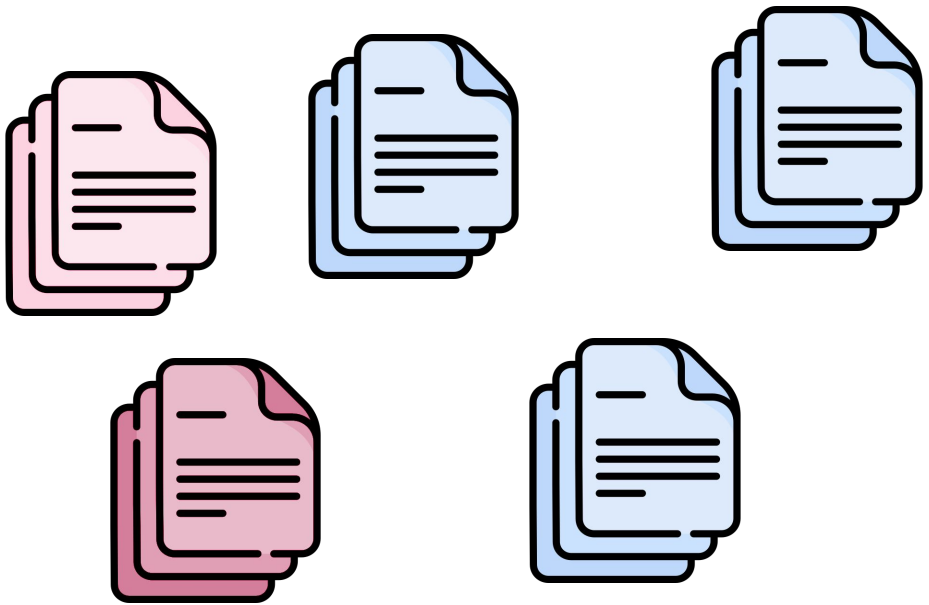
For every real document d, develop methods to automatically generate a set Fake(d) of fake documents that are very similar to d.

Goal:
The attacker who steals documents must wade through a large number of documents in detail in order to separate the real one from the fakes.

Fake documents

We want to increase costs on attackers who wish to steal (exfiltrate) documents from an organization.





Fake documents - adversary's problems

Case 1: unaware that the system has fake files:

External connections and exfiltration of data is monitored closely by most organizations with sensitive information.

High probability that the attacker will steal an incorrect file. If such files include, complex designs for an aircraft/submarine/missile, the attacker would incur actual costs (dollars) and time delays as he tries to execute the design, only to find months later that it doesn't work.

Fake documents - adversary's problems

Case 2: aware that the system has fake files:

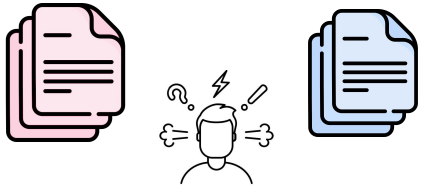
Option 1: Spend more time within the system picking and trying to choose the right document -> increases risk of discovery

Fake documents - adversary's problems

Case 2: aware that the system has fake files:

Option 2: exfiltrate many documents and go through them at your own pace within his facility/network -> increases risk of discovery

Fake documents - defenders problems



Generate “believable” fake version ***d’*** of an original document ***d*** to be successful in deceiving the attacker.

Fake document generation

Generate “believable” fake version ***d’*** of an original document ***d*** to be successful in deceiving the attacker.

Focus today will be on the generation of believable fake documents for technical topics such as engineering documents

Fake document generation - steps

1. **Extraction of concepts**
2. **Documents to Multi-layer Graphs**
3. **Employ Meta-Centrality**
4. **Generating Fake Documents**

Fake document generation - Extract concepts

Extracting candidate concepts:

- Given a document d:
  - Extract equations and formulas (using the html tags present in the document)
  - Use a parser to identify and extract all noun phrases of (size 5 or less).

Fake document generation - Extract concepts

Discarding redundant concepts:

- Once candidate concepts are identified:
  - Discard redundant concepts. If one concept is a proper subset of another concept, discard the smaller one and only consider the larger one.

Fake document generation - Extract concepts

Discarding useless concepts:

- Need to discard concepts that are not related to the domain of the repository. For instance, if we seek to protect intellectual property of an aerospace company, we would remove documents related to the company's employee health plans.
- Define set of rules to perform this discarding

Fake document generation - from document to multi-layer graphs

After concept extraction a graph is built consisting of two layers – a **document layer** and a **domain layer.**

The multi-layer graph associated with a document captures both the relationships between concepts inside the document as well as across the domain related to the document.

A vertex in the multi-layer graph of a document is a concept and an edge connecting two vertices represents the similarity between the two concepts.

Fake document generation - Document layer

Given a document d:

Document layer (d) is a graph whose:

- Vertices are concepts extracted from d
- Edges are existent between concepts if they appear within a sliding window of k or less.
- Each edge has a weight that is the number of times the two concepts appear within the sliding window in the whole document.

Fake document generation - Document layer

Given a document d:

Document layer (d) is a graph whose:

- Vertices are concepts extracted from d
- Edges are existent between concepts if they appear within a sliding window of k or less.
- Each edge has a weight that is the number of times the two concepts appear within the sliding window in the whole document.

The document layer captures the idea that if two concepts appear together multiple times inside a document, they are likely to be linked.

Fake document generation - terminology

Context of a Concept:
Given a repository of documents and a concept ***c***, the domain of ***c*** consists of the set of all words appearing **k** positions before or after ***c*** in any document in the repository.

Fake document generation - similarity definition

Similarity between two concepts (A, B):
It is represented by the Jaccard similarity of two concepts and their contexts C(A), C(B).

$$\text{Jaccard coefficient } (A,B) = |C(A) \cap C(B)| / |C(A) \cup C(B)|$$

Fake document generation - Domain Layer

For a document d and a repository R:

- The domain layer is a graph where:
  - Vertices are the set of concepts
  - Edges exist between two concepts if Jaccard sim. > 0
  - Edge weight equal to the Jaccard similarity value.

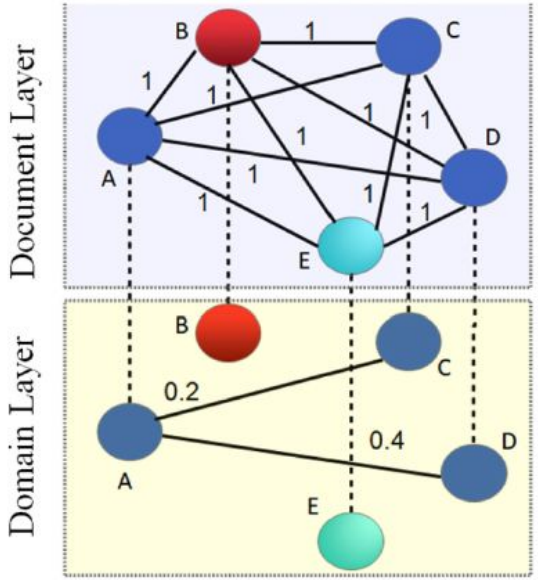
Fake document generation - Domain Layer

For a document d and a repository R:

- The domain layer is a graph where:
  - Vertices are the set of concepts
  - Edges exist between two concepts if Jaccard sim. > 0
  - Edge weight equal to the Jaccard similarity value.

The idea behind the domain layer stems from a popular concept called “distributional semantics” in linguistics which suggests that the words that are used and occur in the same contexts tend to have similar meanings.

Fake document generation - Multi Layer Graph



For a document d and a repository R, the two layers:
<document layer; domain layer> are associated together to form the multi layer graph MLG(d) where:

- Vertices is the union of vertices of the layers.
- Edges are union of the edges of the layers
- There exist a inter-layer edge between same concept in both layers

Fake document generation - Meta centrality

Centrality - how central concepts are in the MLG structure.
It tells us how important concepts are.

We can use different centrality measures:
Degree centrality - CD(c) = degree(c)/number of edges

Betweenness centrality - measures the probability a node appears in the shortest path between two other nodes.
Total number of shortest paths passing through c divided by the number of shortest paths there exist between any two nodes for all pairs except from c.

Fake document generation - Meta centrality

Centrality - how central concepts are in the MLG structure.
It tells us how important concepts are.

We can use different centrality measures:
Closeness centrality: - reciprocal of the sum of the shortest path from node c to every other node in the graph.

Page Rank - uses random walks to identify individuals who are commonly encountered along such walks. Those individuals are viewed as central.

Fake document generation - Meta centrality

- Calculate the meta-centrality value of each concept.
- Rank them in decreasing order.

Fake document generation

After all this process we get the meta-centrality rank of each concept in document ***d***.

Now we need to generate fake versions of ***d*** by replacing some concepts with “similar concepts” in such a way that the fake documents appear “believable” to the attackers.

Fake document generation

Meta-centrality provides a rank of concepts in a document.

- Do not know a-priori if replacing top-ranked concepts would generate a believable document.

- Replacement of a concept by another concept incurs a specific cost in terms of believability, depending both on what is replaced and what the replacement is.

Set of Fake documents

Given an original document ***d***, we define a set Fake(***d***) of fake documents as an ordered list of documents {d1,d2,d3…} where each fake document d_i is generated by replacing selected concepts in **d** by other alternative concepts.

Which candidate concepts can replace a given concept?

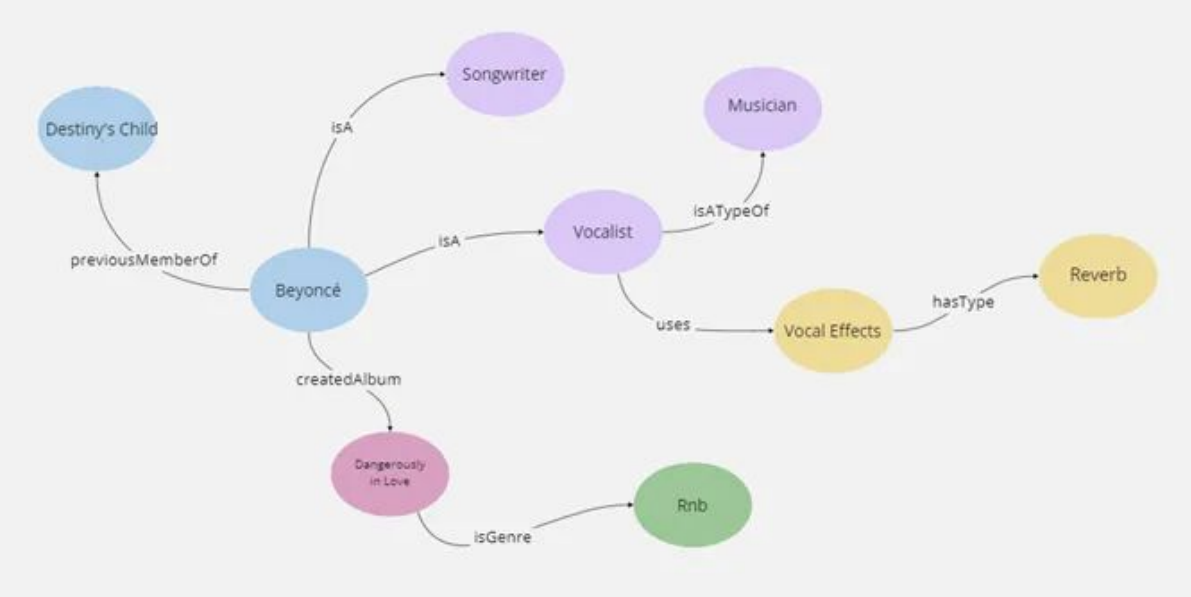
To identify the candidate concepts to replace c, use a domain-specific ontology.

Ontology - model the knowledge about a domain by identifying specific classes of concepts, defining each instance of concepts through properties/attributes, including also different types of relationships between concepts.

- Each substitution should be done by a concept in the ontology whose cost is below a certain threshold.
- By requiring that the total substitution cost is below a given threshold -> concepts in the real document are replaced by concepts “nearby” in the ontology, thus enhancing **believability**.

Which candidate concepts can replace a given concept?

Ontology - model the knowledge about a domain by identifying specific classes of concepts, defining each instance of concepts through properties/attributes, including also different types of relationships between concepts.



Which candidate concepts can replace a given concept?

To identify the candidate concepts to replace c, use a domain-specific ontology.

Ontology - model the knowledge about a domain by identifying specific classes of concepts, defining each instance of concepts through properties/attributes, including also different types of relationships between concepts.

- Each substitution should be done by a concept in the ontology whose cost is below a certain threshold.
- By requiring that the total substitution cost is below a given threshold -> concepts in the real document are replaced by concepts “nearby” in the ontology, thus enhancing **believability**.

Which concepts should be replaced and how many?

Map the fake document generation problem to an optimization problem constrained on the **overall budget** to generate fake documents and the number of candidate concepts to be replaced.

- select these concepts from different positions of the rank-list obtained from the meta-centrality values.

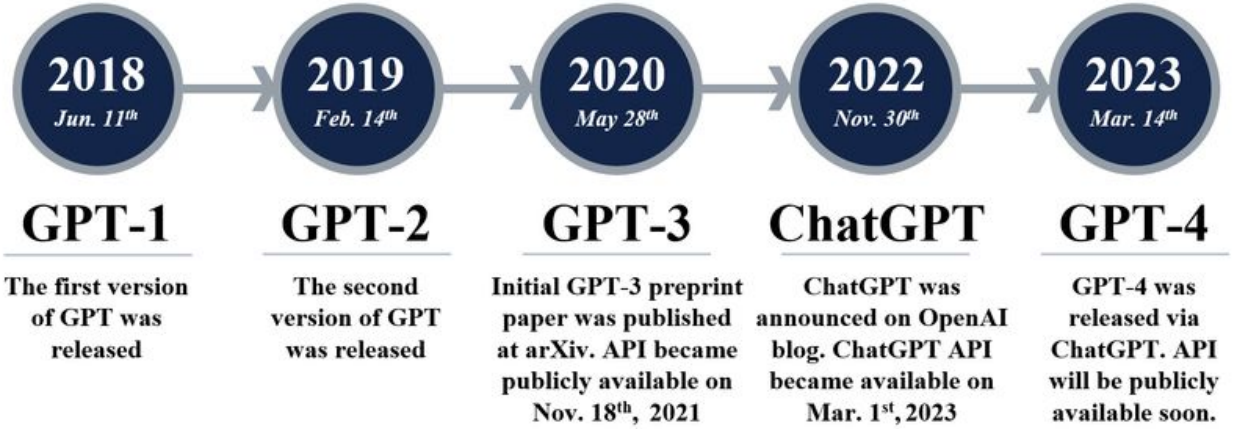
Fake document generation problem - final definition

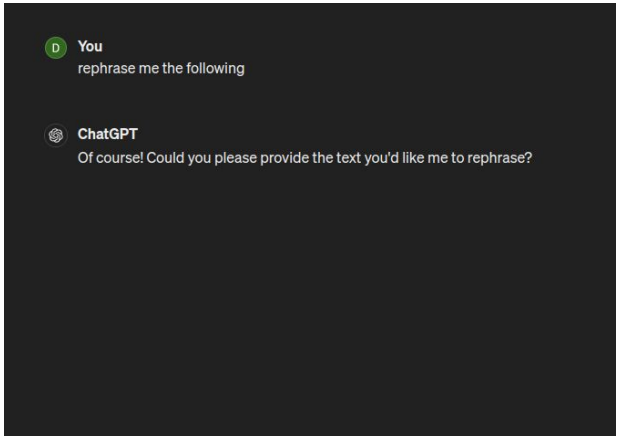
- Have:
  - Document d
  - Budget B

- Extract:
  - Set of concepts C = {c1, c2 ….}
  - Each concept has two attributes:
    - Meta centrality
    - Replacement cost with an alternative concept

The overall problem is to generate fake documents from d by replacing a set of C in such a way that the meta-centrality ranking of the selected concept will be minimum and the substitution cost will remain bounded by a certain budget.

Nowadays


2018
Jun. 11th
GPT-1
The first version of GPT was released
2019
Feb. 14th
GPT-2
The second version of GPT was released
2020
May 28th
GPT-3
Initial GPT-3 preprint paper was published at arXiv. API became publicly available on Nov. 18th, 2021
2022
Nov. 30th
ChatGPT
ChatGPT was announced on OpenAI blog. ChatGPT API became available on Mar. 1st, 2023
2023
Mar. 14th
GPT-4
GPT-4 was released via ChatGPT. API will be publicly available soon.


You
rephrase me the following
ChatGPT
Of course! Could you please provide the text you'd like me to rephrase?

Reading Material

1. Fake document generation: Link-1, Link-2