ACCESS CONTROL

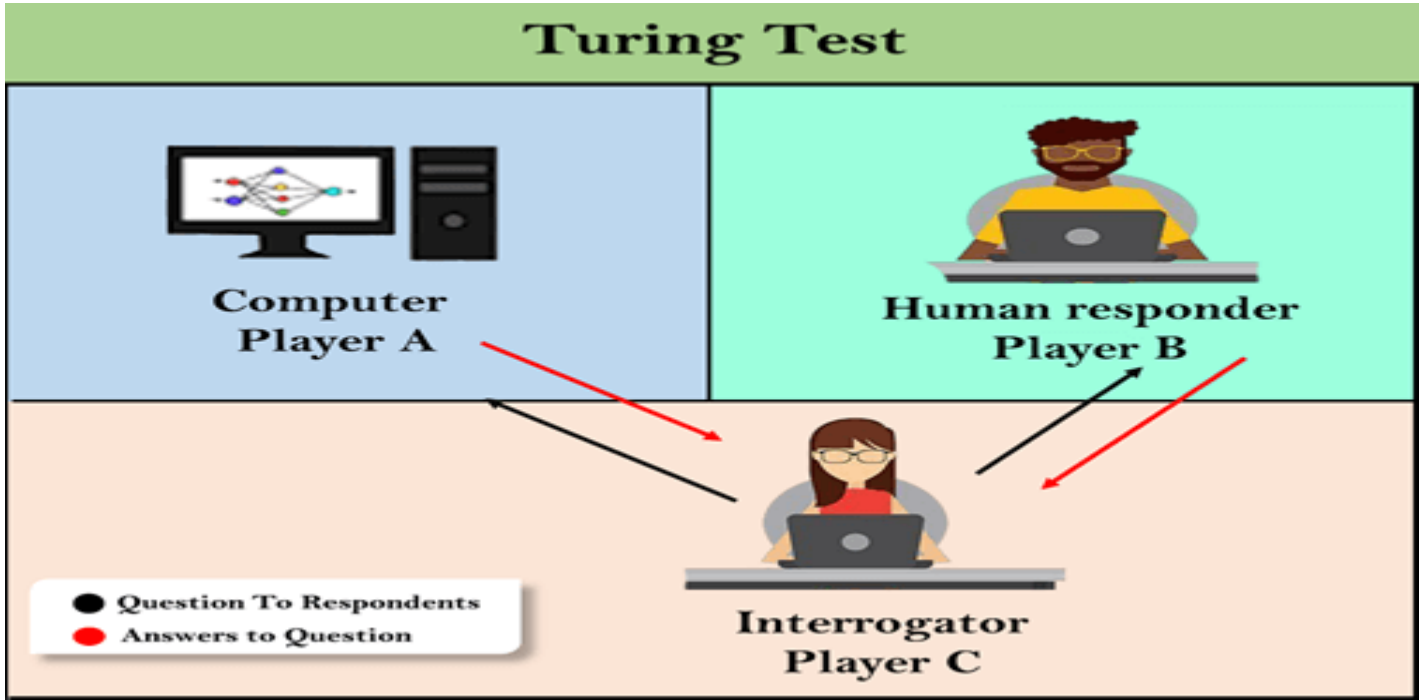
Prepared By

-Anooja Joy

CAPTCHA

ACTING HUMANLY: TURING TEST

- Turing Test is the gold standard in AI
- No computer can pass this today
  - But some claim they are close to passing



Alan Turing

- If interrogator cannot reliably distinguish human from computer, then computer possess intelligence.

CAPTCHA

- CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart). Inverse Turing Test
- CAPTCHA is a Completely Automated test which is generated and scored by a computer that a human can pass, but a machine cannot pass(even if it has access to source code to generate test) with probability better than guessing.
- CAPTCHA **is a program that can generate a test and grade it, but itself cannot pass.**(similar to some professors)
- CAPTCHA **is an access control mechanism to restrict access to resources to humansie,** Only humans get access (not bots/computers)

CAPTCHA

- Computing problems that must be solved to break CAPTCHA(eg: Automatic Recognition, Distorted Text, Distorted Audio) are considered difficult problems in AI.
- The requirements for a CAPTCHA include
  - It must be **easy** for most **humans** to pass.
  - It must be **difficult** or **impossible** for a machines to pass, (even if the machine has access to the CAPTCHA software)
  - **Randomness** in generating CAPTCHA..
  - Have **different types of CAPTCHA** depending on individual requirement

FEATURES TO BE POSSESSED BY CAPTCHA

1. **Accessibility.** CAPTCHAs must be accessible. CAPTCHAs based solely on reading text – or other visual-perception tasks – prevent visually impaired users from accessing the protected resource. Such CAPTCHAs may make a site incompatible with Section 508 in the United States. Any implementation of a CAPTCHA should allow blind users to get around the barrier, for example, by permitting users to opt for an audio or sound CAPTCHA.
2. **Image Security.** CAPTCHA images of text should be distorted randomly before being presented to the user. Many implementations of CAPTCHAs use undistorted text, or text with only minor distortions. These implementations are vulnerable to simple automated attacks.
3. **Script Security.** Building a secure CAPTCHA code is not easy. In addition to making the images unreadable by computers, the system should ensure that there are no easy ways around it at the script level. Common examples of insecurities in this respect include: (1) Systems that pass the answer to the CAPTCHA in plain text as part of the web form. (2) Systems where a solution to the same CAPTCHA can be used multiple times (this makes the CAPTCHA vulnerable to so-called "replay attacks"). Most CAPTCHA scripts found freely on the Web are vulnerable to these types of attacks.
4. **Security Even After Wide-Spread Adoption.** There are various "CAPTCHAs" that would be insecure if a significant number of sites started using them. An example of such a puzzle is asking text-based questions, such as a mathematical question ("what is 1+1"). Since a parser could easily be written that would allow bots to bypass this test, such "CAPTCHAs" rely on the fact that few sites use them, and thus that a bot author has no incentive to program their bot to solve that challenge. True CAPTCHAs should be secure even after a significant number of websites adopt them.

APPLICATIONS OF CAPTCHA

1. **Preventing Comment Spam in Blogs.** Most bloggers are familiar with programs that submit bogus comments, usually for the purpose of raising search engine ranks of some website. This is called comment spam. By using a CAPTCHA, only humans can enter comments on a blog. There is no need to make users sign up before they enter a comment, and no legitimate comments are ever lost!
2. **Protecting Website Registration.** Several companies (Yahoo!, Microsoft, etc.) offer free email services. Up until a few years ago, most of these services suffered from a specific type of attack: "bots" that would sign up for thousands of email accounts every minute. The solution to this problem was to use CAPTCHAs to ensure that only humans obtain free accounts. In general, free services should be protected with a CAPTCHA in order to prevent abuse by automated scripts.
3. **Protecting Email Addresses From Scrapers.** Spammers crawl the Web in search of email addresses posted in clear text. CAPTCHAs provide an effective mechanism to hide your email address from Web scrapers. The idea is to require users to solve a CAPTCHA before showing your email address. A free and secure implementation that uses CAPTCHAs to obfuscate an email address can be found at reCAPTCHA MailHide.
4. **Search Engine Bots.** It is sometimes desirable to keep webpages unindexed to prevent others from finding them easily. There is an html tag to prevent search engine bots from reading web pages. The tag, however, doesn't guarantee that bots won't read a web page; it only serves to say "no bots, please." Search engine bots, since they usually belong to large companies, respect web pages that don't want to allow them in. However, in order to truly guarantee that bots won't enter a web site, CAPTCHAs are needed.

APPLICATIONS OF CAPTCHA

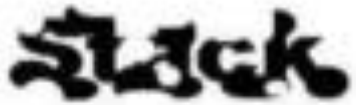
1. **Online Polls.** In November 1999, *http://www.slashdot.org* released an online poll asking which was the best graduate school in computer science (a dangerous question to ask over the web!). As is the case with most online polls, IP addresses of voters were recorded in order to prevent single users from voting more than once. However, students at Carnegie Mellon found a way to stuff the ballots using programs that voted for CMU thousands of times. CMU's score started growing rapidly. The next day, students at MIT wrote their own program and the poll became a contest between voting "bots." MIT finished with 21,156 votes, Carnegie Mellon with 21,032 and every other school with less than 1,000. Can the result of any online poll be trusted? Not unless the poll ensures that only humans can vote.
2. **Preventing Dictionary Attacks.** CAPTCHAs can also be used to prevent dictionary attacks in password systems. The idea is simple: prevent a computer from being able to iterate through the entire space of passwords by requiring it to solve a CAPTCHA after a certain number of unsuccessful logins. This is better than the classic approach of locking an account after a sequence of unsuccessful logins, since doing so allows an attacker to lock accounts at will.
3. **Worms and Spam.** CAPTCHAs also offer a plausible solution against email worms and spam: "I will only accept an email if I know there is a human behind the other computer." A few companies are already marketing this idea.

CAPTCHA TYPES

- TEXT BASED
  - Gimpy, ez-Gimpy
  - Gimpy-R
  - Simard's HIP
- GRAPHIC BASED
  - Bongo
  - Pix
  - 3D
- AUDIO BASED
- Video based

TEXT BASED CAPTCHA's

(a) reCAPTCHA

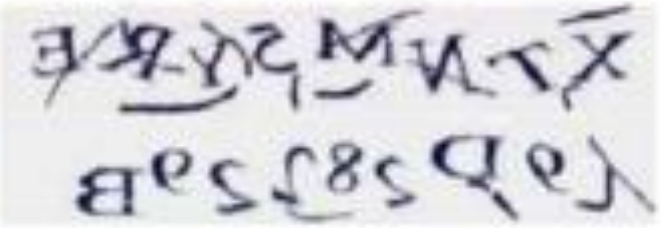


(b) buffle Text

(c) Gimpy

(a) E-Z gimpy

(b) Drag&Drop



(c) MSN

Gimpy

- Designed by Yahoo and CMU (Carnegie Mellon University)
- **Gimpy** works by choosing a certain number of words from a dictionary, and then displaying them corrupted, distorted and overlapped manner. **Gimpy** then asks the user enter a subset of the words in the image.
- Gimpy is based on the human has the ability to read extremely distorted text otherwhile the computer programs don‘t have such type of ability to do the same.

ez-gimpy

- Modified version of Gimpy. Randomly picks a single word from a dictionary and applies distortion to the text. The user is then asked to identify the text correctly.
- Used by Yahoo in Messenger
- Not a good implementation, already broken by OCRs

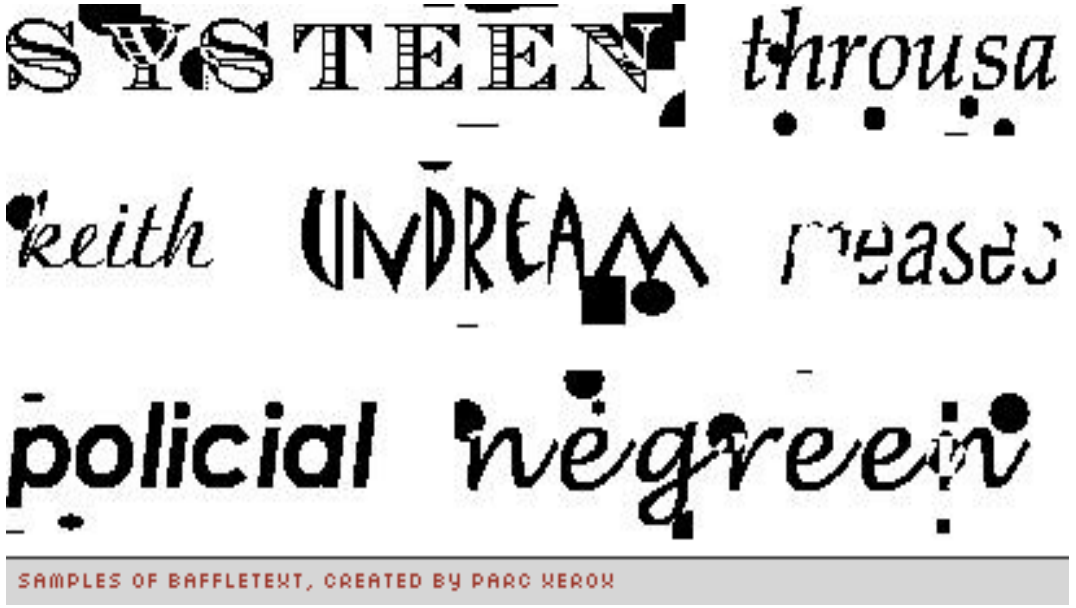
Gimpy-r

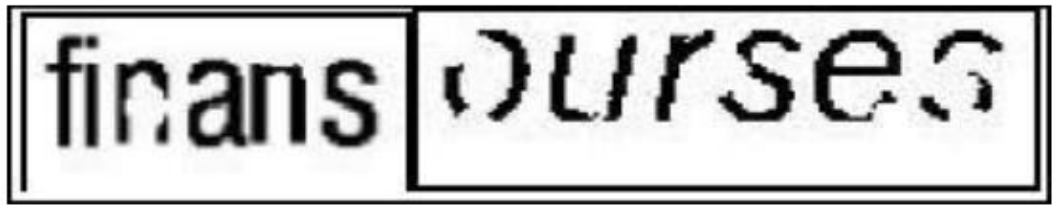
- Google CAPTCHA- Instead of complete word individual letters are noised.
- Pick random letters , Distort them , add noise and background

- **Simcard's HIP (Human Interaction Proofs) (MSN)-**
- Distort using arcs
  - Pick random letters and numbers
  - Distort them and add arcs

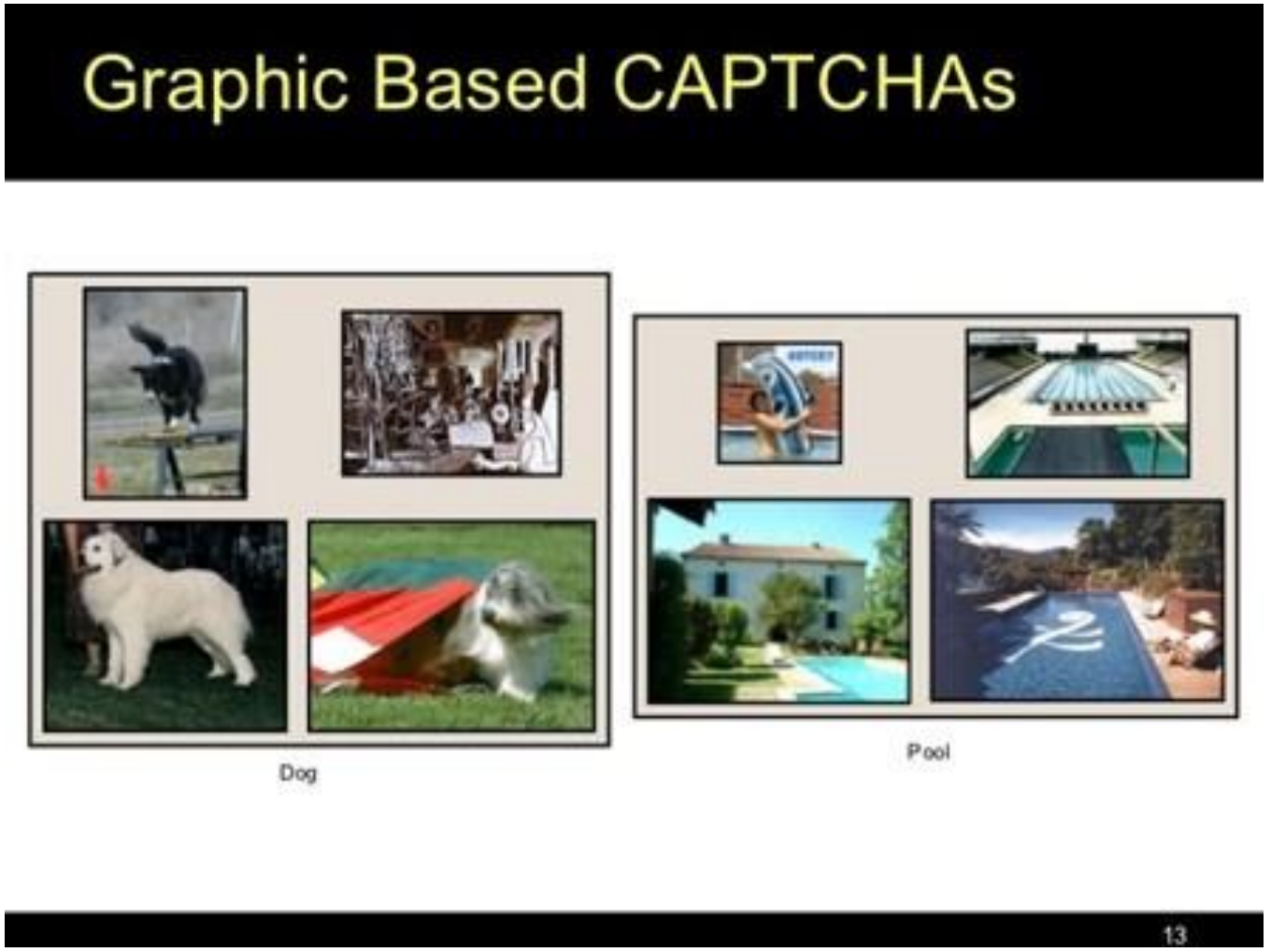
BAFFLE TEXT

- This doesn't contain dictionary words, but it picks up random alphabets to create a nonsense but pronounceable text. Distortions are then added to this text and the user is challenged to guess the right word.
- This technique overcomes the drawback of Gimpy CAPTCHA because, Gimpy uses dictionary words and hence, clever bots could be designed to check the dictionary for the matching word by brute-force.





GRAPHIC BASED CAPTCHAS



Pix

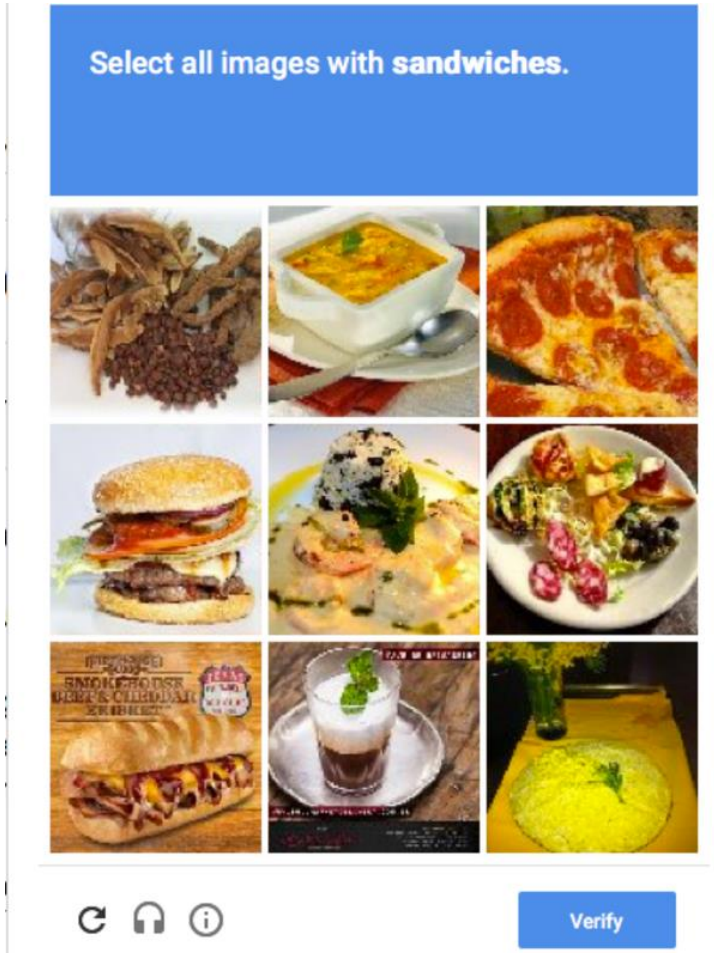
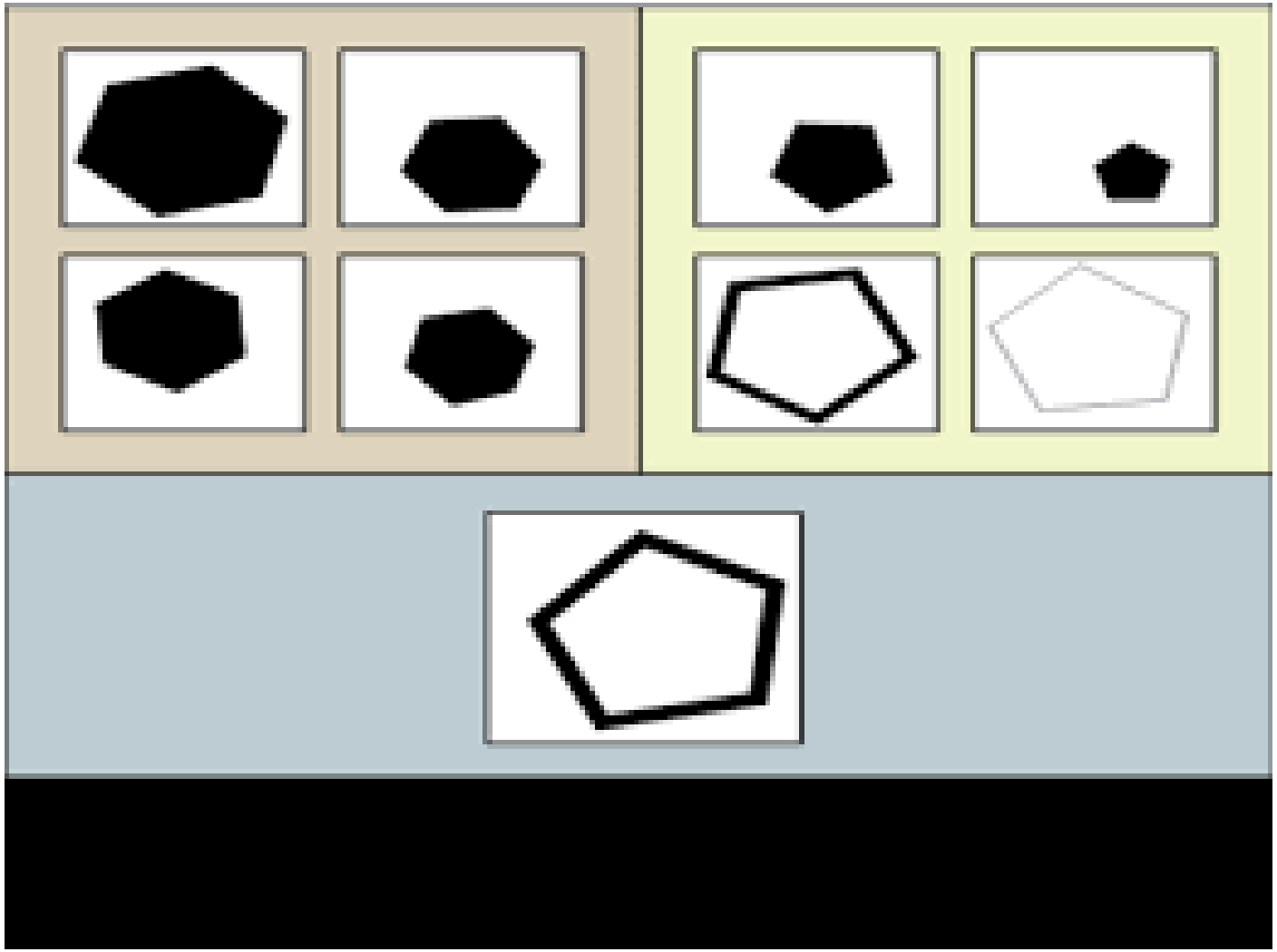


Fig 2 Images Based Captcl

- Uses a large database of labeled images
- Shows a set of images, user has to recognize the common feature among those

Bongo

- Display 2 series of blocks
- User must find characteristics that sets 2 apart
- User is asked to which series given block belongs to



Audio Based

- Consists of downloadable audio clip. picks a word or a sequence of numbers at random, renders the word or the numbers into a sound clip and distorts the sound clip; it then presents the distorted sound clip to the user and asks users to enter its contents
- User listens and enters the spoken word
- Helps visually disabled users

AUTHENTICATION METHODS

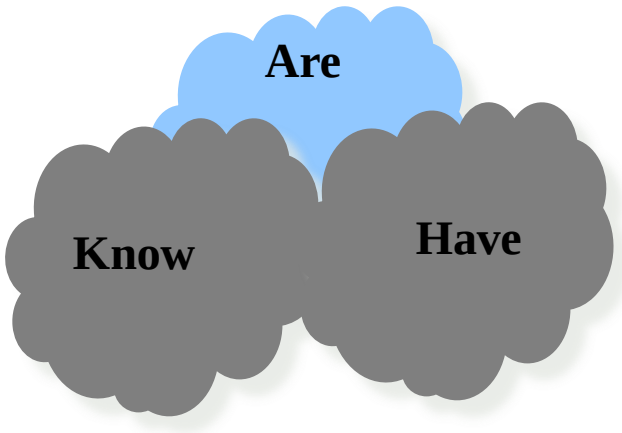
User Authentication Methods

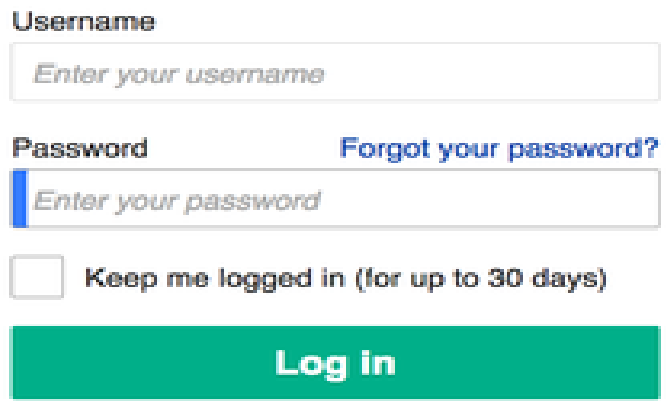
- Using a method to validate users who attempt to access a computer system or resources, to ensure they are authorized.
- Authenticating a Human to a machine to convince someone or something claiming to be.
- Types of user authentication
  - Something you know

    E.g: user account names and passwords
  - Something you have

    Eg: Smart cards or other security tokens
  - Something you are

    Eg: Biometrics





PASSWORD AUTHENTICATION

- ***An ideal password is something user knows so that a machine can verify that user know and attacker doesn't know or can't guess even with the access to unlimited computing resources.***

Eg: PIN for an ATM, password of web applications, Social security number, Mother's maiden name, Date of birth, Name of your pet

Why Passwords are more preferred or popular?

- **Cost:** passwords are free compared to smart cards and biometric devices
- **Convenience:** easier for sysadmin to reset a compromised password than to issue and configure a new smart card.

Password Variants

- **Cognitive Passwords:** created through several experience-based questions. Fact or opinion-based questions

Eg: favorite color, college name, birthplace

- **One-time passwords:** used in sensitive cases

Eg: Token devices, prepaid cards

- **Passphrases:** sequence of characters ie no longer than password. User enters phrase into application which transforms value into password.

Eg: FSa7Yago(four score and seven years ago)

- **Personal identification number (PIN):** an arbitrary string of characters where the permissible characters are constrained to be numeric

Eg:347865

Cryptographic keys

- If cryptographic key is 64 bits, then 2^{64} possible keys where the attacker must try about 2^{63} keys to find correct one.

Passwords

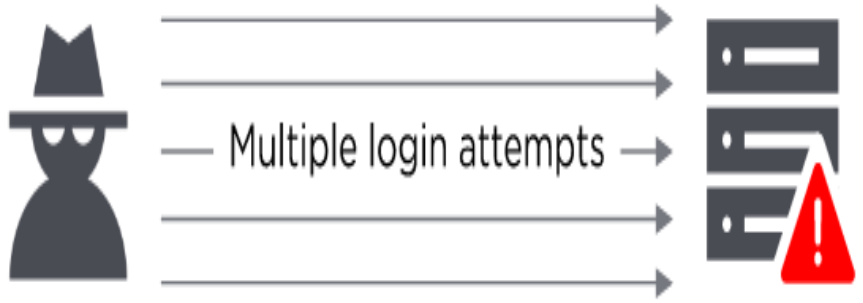
- If password is 8 characters long, with 256 possible choices for each character, then there are $256^8 = 2^{64}$ possible passwords.
- But users do not select passwords at random since they must memorize and a clever attacker can make fewer than 2^{63} guesses to crack it.(Dictionary attack)

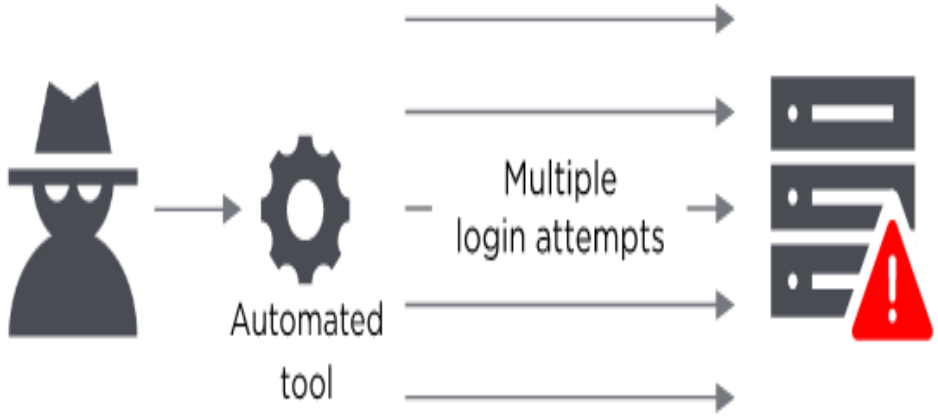
DICTIONARY ATTACK FAILURE FACTORS

- The average number of guesses the attacker must make to guess the correct password is determined by how unpredictable the password is or randomness of password, including **how long the password is, what set of symbols it is drawn from, and how it is created.**
- The ease with which an attacker can check the validity of a guessed password is determined by **how the password is stored, how the checking is done, and any limitation on trying passwords**

COMMON ATTACKS TO PASSWORDS

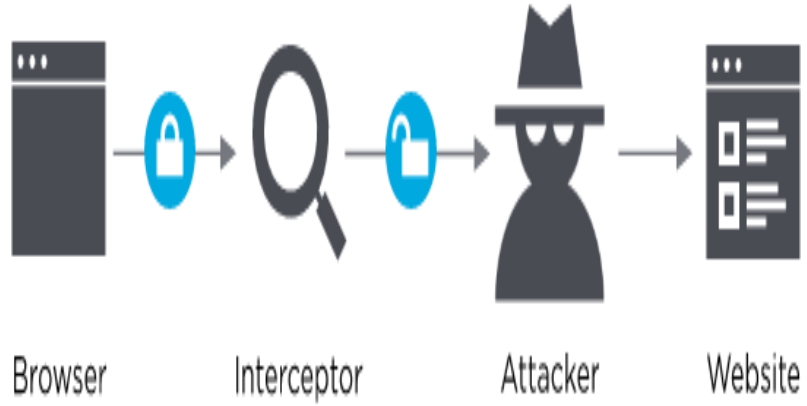
- **Dictionary attack:** An attack that takes advantage of the fact people tend to use common words and short passwords. The hacker uses a list of common words, the dictionary, and tries them, often with numbers before and/or after the words, against accounts in a company for each username. Trudy pre-computes h(x) for all x in a dictionary of common passwords. Suppose Trudy gets access to password file containing hashed passwords. She only needs to compare hashes to her pre-computed dictionary.
- **Brute force:** Using a program to generate likely passwords or even random character sets. These attacks start with commonly used, weak passwords like Password123 and move on from there. The programs running these attacks usually try variations on upper and lowercase characters, as well.

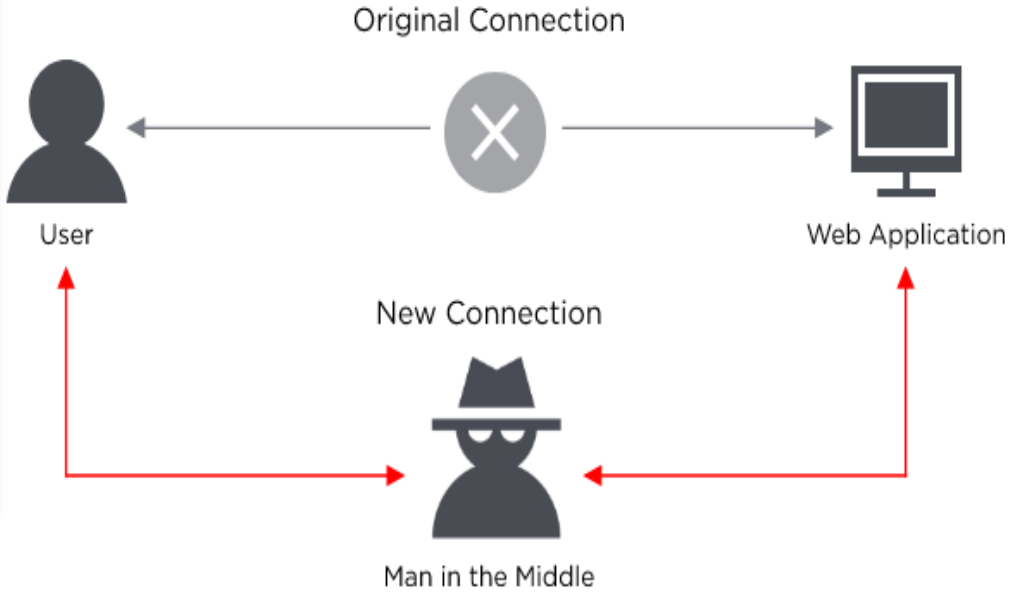




COMMON ATTACKS TO PASSWORDS

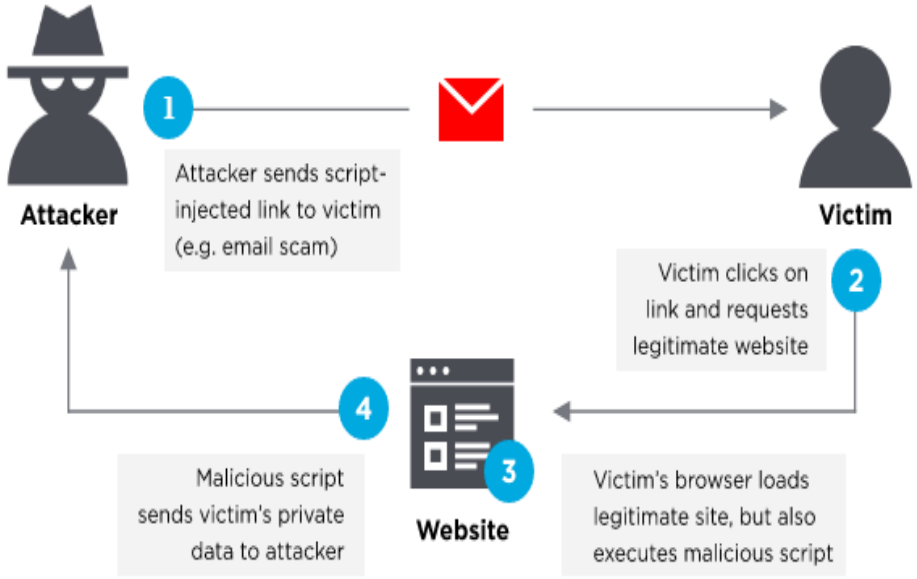
- **Traffic interception/ Eavesdropping:** In this attack, the cyber criminal uses software such as packet sniffers to monitor network traffic and capture passwords as they're passed. In this attack the software monitors and captures critical information. Obviously, if that information such as passwords is unencrypted, the task is easier. But even encrypted information may be decryptable, depending on the strength of the encryption method used.
- **Man In the Middle/ Social Engineering:** In this attack, hacker's program doesn't just monitor information being passed but actively inserts itself in the middle of the interaction, usually by impersonating a website or app. This allows the program to capture the user's credentials and other sensitive information, such as account numbers, social security numbers, etc. MITM attacks are often facilitated by social engineering attacks which lure the user to a fake site.





COMMON ATTACKS TO PASSWORDS

- **Key logger attack:** A cyber criminal manages to install software that tracks the user's keystrokes, enabling the criminal to gather not only the username and password for an account but exactly which website or app the user was logging into with the credentials. This type of attack generally relies on the user first falling prey to another attack that installs the malicious key logger software on their machine. It is also known as electronic monitoring.
- Accessing password file
- Login spoofing (human errors)



COMMON ATTACK PATH

- Outsider → normal user → administrator
- May only require **one** weak password to crash entire system!
- Normally attacker tries to grab one account and tries for privilege escalation.

RESPONSE TO A PASSWORD ATTACK

- If an attack happens with a password what is the proper response mechanism?
- Lock after 3 failed attempts
- How long should system lock?
- 5 seconds(Attacker can easily surge for next round) or 5 minutes(aattacker can cause denial of service attack)

Bad passwords

- **Default passwords** (as supplied by the system vendor and meant to be changed at installation time): *password*, *default*, *admin*, *guest* etc.
- **Dictionary words:** *chameleon*, *RedSox*, *sandbags*, *bunnyhop!*, *IntenseCrabtree*, etc.
- **Words with numbers appended:** *password1*, *deer2000*, *john1234*, etc.,
- **Words with simple obfuscation:** *p@ssw0rd*, *g0ldf1sh*, etc.
- **Doubled words:** *crabcrab*, *stopstop*, *treetree* *passpass* etc, can be easily tested automatically.

Good Passwords

- Allow long passphrases (FSa7Yago, 0nceuP0nAt1m8)
- Randomly generate passwords, though probably inappropriate for most scenarios
- Check the quality of user-selected passwords
  - use a number of rules of thumb
  - run dictionary attack tools

Bad passwords

- **Anything personally related to an individual:** license plate number, Social Security number, current or past telephone number, student ID, address, birthday, sports team, relative's or pet's names/nicknames/birthdays, etc.,: frank, Fido, AustinStamp(names), (10021995birthdates)
- **Common sequences from a keyboard row:** *qwerty*, *12345*, *asdfgh*, *fred*, etc.
- **Numeric sequences based on well known numbers** such as 911, 314159, or 27182, etc.
- **Identifiers:** *jsmith123*, *1/1/1970*, *555–1234*, "your username", etc

Good Passwords

- Give user suggestions/guidelines in choosing passwords
- e.g., think of a sentence and select letters from it, "It"s 12 noon and I am hungry" => "I"S12&IAH"
- Using both letter, numbers, and special characters

Passord Strength

- Three groups of users – each group advised to select passwords as follows
  - **Group A:** At least 6 chars, 1 non-letter
  - **Group B:** Password based on passphrase
  - **Group C:** 8 random characters
- Results
  1. **Group A:** About 30% of pwds easy to crack
  2. **Group B:** About 10% cracked
     Passwords easy to remember
  3. **Group C:** About 10% cracked
     Passwords hard to remember
- Inferences
  - **Assigned passwords sometimes best**
  - **If passwords not assigned, best advice is…**
    - Choose passwords based on passphrase
    - Use pwd cracking tool to test for weak pwds

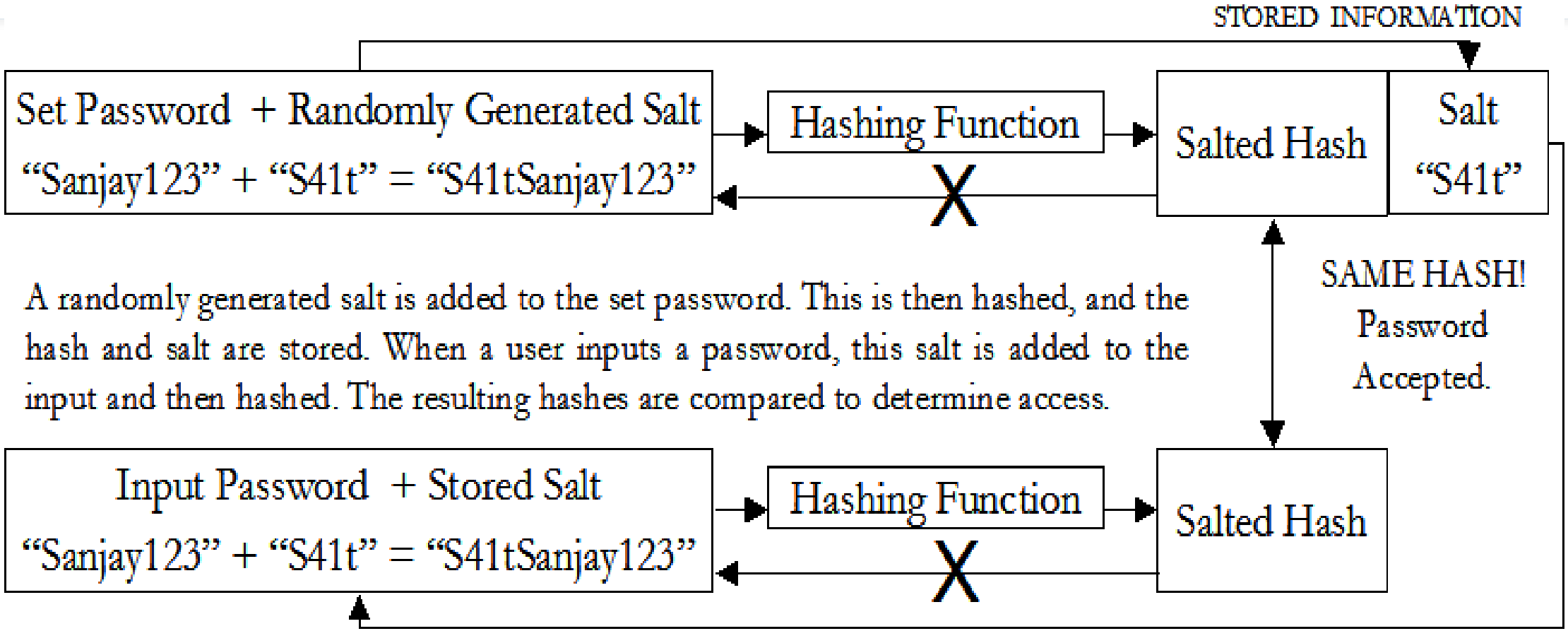
Password Verification

- To determine the validity of entered password machine must have access to the correct password. The passwords can be stored as:
  - **File System:** Clear text
  - **Dedicated Authentication Server:** Clear text
  - **Encrypted Password:** Password + Encryption = bf4ee8HjaQkbw
  - **Hashed Password:** Plain-text is converted into a message digest through the use of a hashing algorithm (i.e. MD5, SHA)
  - Password + Hash function = aad3b435b51404eeaad3b435b51404ee
  - **Salted Hash**(Username + Salt + Password) + Hash function =ad3b435b51404eeaad3b435b51404ee
  - **Create Fake Password files and store it all over the system**

SALT

- Hash password with **salt. Salting is adding a random piece of data to the password before hashing it.**
  - This means that the same string will hash to different values at different times.
  - Users with same password have different entries in the password file.
  - Salt is stored with the other data as a complete hash
- Choose random salt **s** and compute **y = h(password, s)** and store **(s,y)** in the password file.
- Note that the **salt s is not secret**. Still easy to verify salted password?
  - Hacker has to get the salt add it to each possible word and then rehash the data prior to comparing with the stored password.
  - With salt, attacker must compute hashes of all dictionary words once for each password entry.
  - With 12-bit random salt, same password can hash to 2^{12} different hash values.
  - Attacker must try all dictionary words for each salt value in the password file
- Without salt, attacker can pre-compute hashes of all dictionary words once for all password entries.
- Identical passwords hash to identical values; one table of hash values can be used for all password files.

SALT


STORED INFORMATION
Set Password + Randomly Generated Salt
"Sanjay123" + "S41t" = "S41tSanjay123"
Hashing Function
Salted Hash
Salt
"S41t"
A randomly generated salt is added to the set password. This is then hashed, and the hash and salt are stored. When a user inputs a password, this salt is added to the input and then hashed. The resulting hashes are compared to determine access.
SAME HASH!
Password
Accepted.
Input Password + Stored Salt
"Sanjay123" + "S41t" = "S41tSanjay123"
Hashing Function
Salted Hash

Password Security Issues

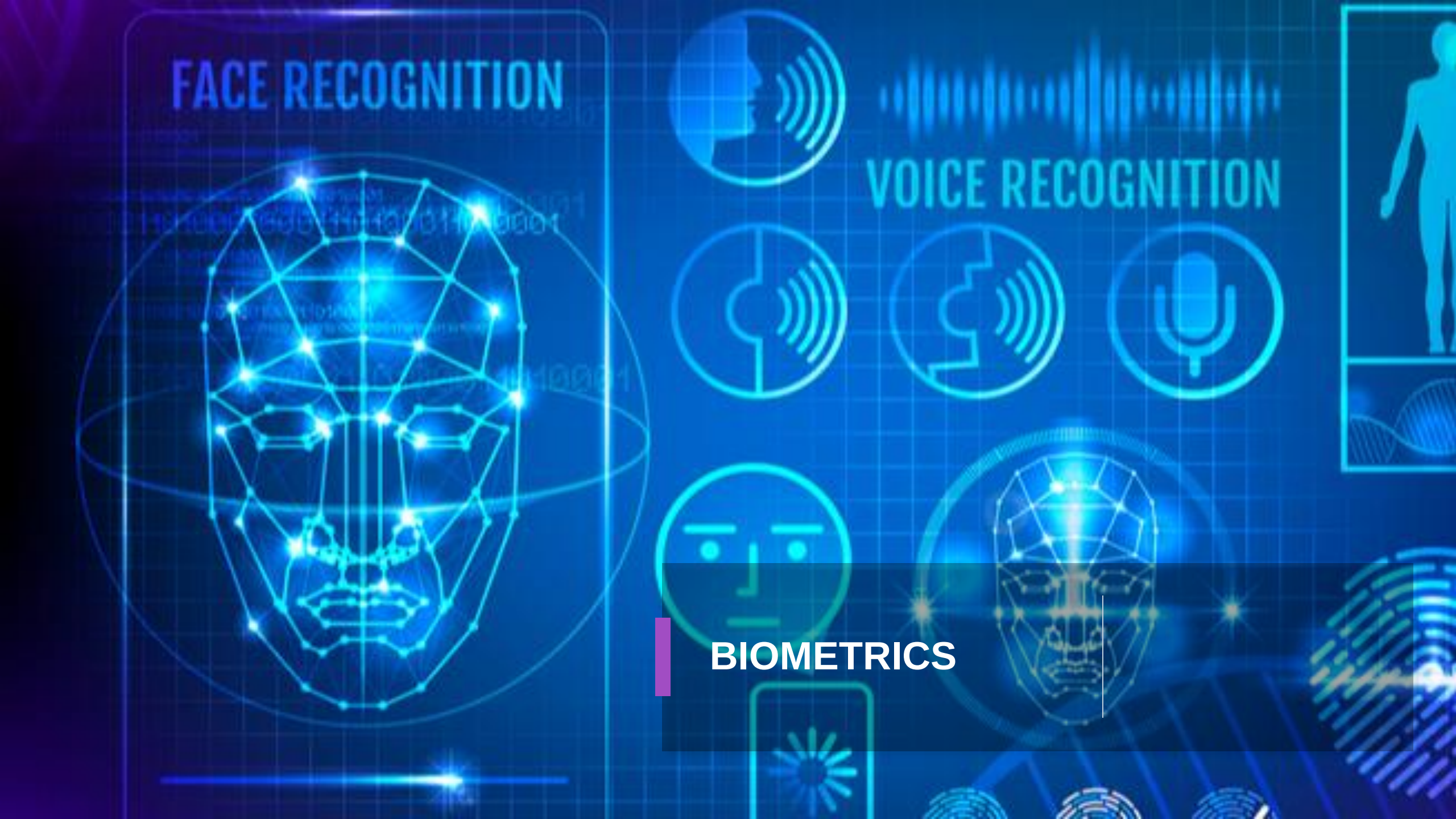
- **Remembering Multiple passwords:** Results in password reuse
- **Social Engineering**
- **Disclosure:** Voluntary disclosure of information and Inadequate guarding of system passwords
- **Inference:** Known pattern to creation of passwords and Use of generated passwords with predictable algorithm
- **Exposure:** Accidental release of password
- **Loss:** Forgetting passwords. Can lead to creation of easy passwords
- **Snooping/Eavesdropping:** Keyloggers and Network sniffing (intercepting of network communication where a password is submitted)
- **Guessing:** Limited amount of choices which can be figured out through process of elimination. Use of blank/common passwords, passwords which can be figured out by knowing name of relatives, pets, etc.
- **Cracking**: Automated "guessing"

Password Cracking Tools

- Popular password cracking tools
  - Password Crackers
  - Password Portal
  - L0phtCrack and LC4 (Windows)
  - John the Ripper (Unix)
- Admins should use these tools to test for weak passwords since attackers will|

Password Security Mechanisms

- Protect stored passwords (use both cryptography & access control)
- Disable accounts with multiple failed attempts
- Require extra authentication (2 or 3 factor) mechanism (e.g., phone, other email account, OTPS etc.)
- Ensure periodic password changes
- The same password can be rehashed many times over to make it more difficult for the hacker to crack the passw ord.
- Enforce strong passwords!
- Set BIOS to boot first from the hard drive.
- Password-protect the BIOS.
- Audit access to important files.



BIOMETRICS

BIOMETRICS: "something you are"

Fingerprint

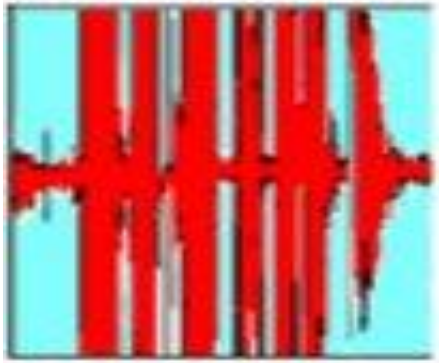
Face

Iris

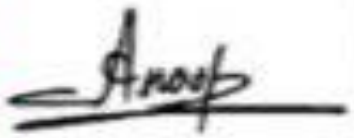
Palm print

Hand vein

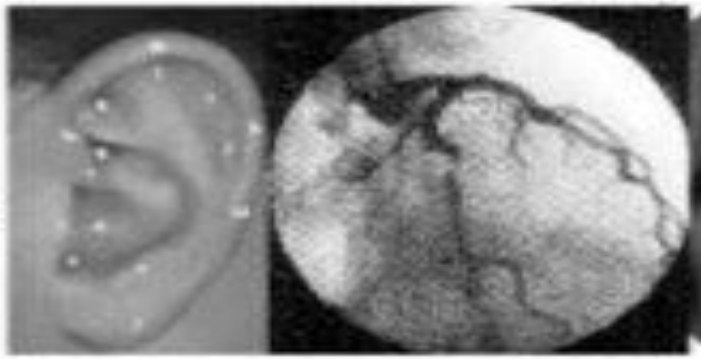
Finger geometry



Voice



Signature



Ear

Retina

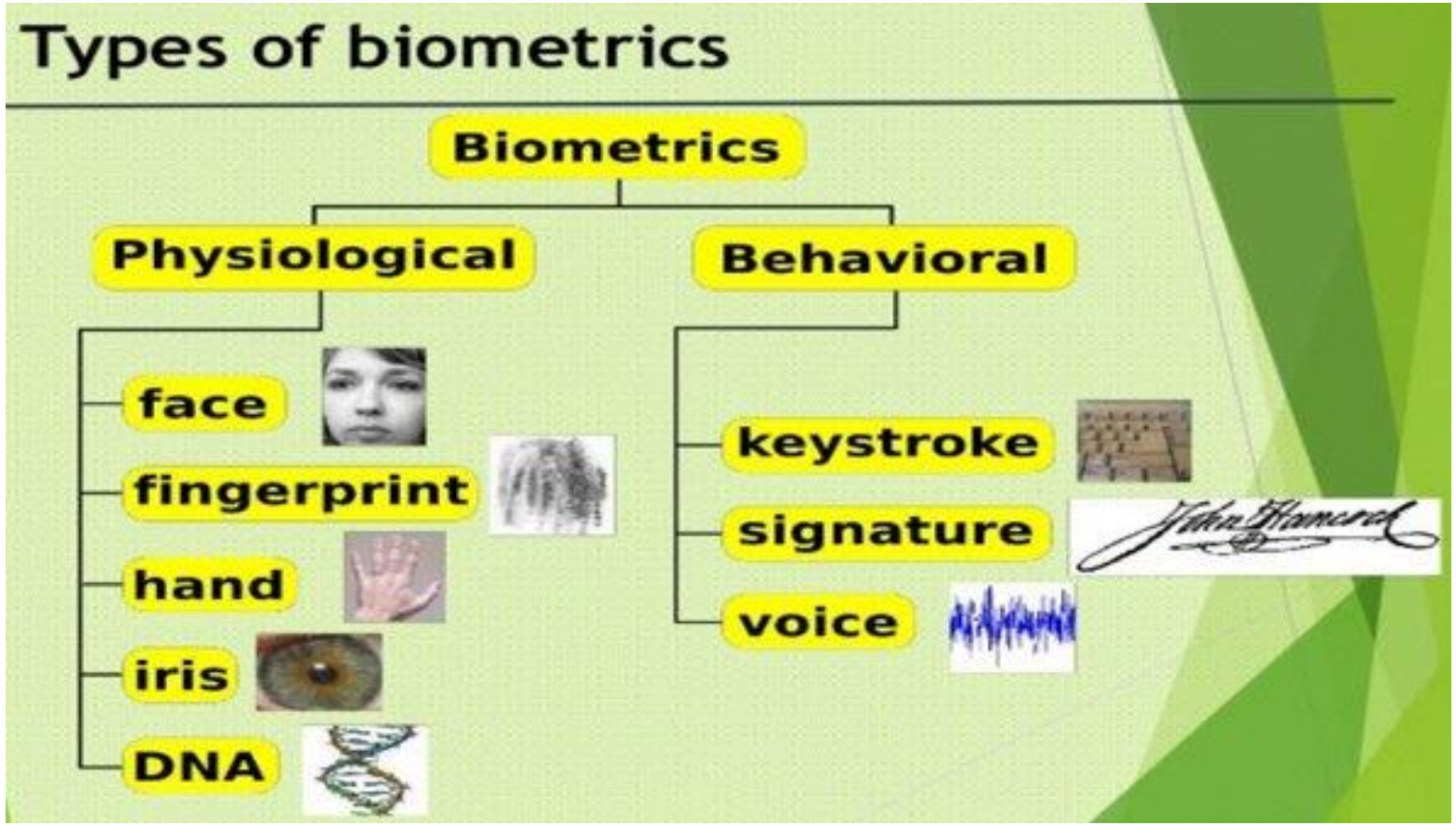
Tooth-shape

Walking gait

BIOMETRICS “You are your key” :Schneier

- Biometrics are automated methods of recognizing a person based on a **physiological or behavioral characteristics.**
- The word biometric is derived from the Greek words bio and metric. Where bio means life and metric means to measure.
- Uses certain biological or behavioral characteristics for authentication
- **Biological/physical Examples**
  - Fingerprint, Iris, Retina, Facial Recognition, & Hand Recognition
- **Behavioral Examples**
  - Handwriting(Signature), Walk/Step(Gait recognition), Typing Rhythm, Mouse Gesture Recognition, digital doggie (Odor recognition), Speech recognition

TYPES OF BIOMETRICS



Why Biometrics?

- Provides **better security and accuracy***. In contrast to passwords, badges, or documents, biometric data cannot be forgotten, exchanged, or stolen, and cannot be forged.
  - Biometrics systems are less prone to attacks
  - Need sophisticated techniques for attacks
  - Cannot steal facial features and fingerprints
  - Need sophisticated image processing techniques for modifying facial features
- Biometrics systems are **more convenient**
  - Need not have multiple passwords or difficult passwords
  - E.g., characters, numbers and special symbols, Need not remember passwords
  - Need not carry any cards or tokens
- **Better accountability**
  - Can determine who accessed the system with less complexity
- **Cheap and reliable**

IDEAL BIOMETRIC CHARACTERISTICS

1	**Universality**	How commonly biometric is found applies to (almost) everyone. In reality, no biometric applies to everyone
2	**Uniqueness/ Distinguishing**	How well biometric distinguishes between others with virtual certainty
3	**Permanence**	Physical characteristics measured should never change. How well biometric resists aging
4	**Collectability**	The physical characteristics should be easy to collect without causing potential harm to subject. How easy biometric is to acquire? Depends on whether subjects are cooperative
5	**Performance**	Reliance, Robust Accuracy, speed, and Userfriendliness of system capturing biometric.
6	**Acceptability**	Degree of approval by the public for use
7	**Circumvention**	How hard it is to fool authentication system

Phases of a biometric system

1. **Enrollment phase**(**Identification**): A raw biometric is captured by a sensing device such as fingerprint scanner or video camera and entered into database. The distinguishing characteristics are extracted from the raw biometrics sample and converted into a processed biometric identifier record called **biometric sample or template or helper data. Its slow phase.** Example: FBI fingerprint database
2. **Recognition phase**(**Authentication**)**:** Biometric authentication works by comparing two sets of data: the first one is preset by the owner of the device, while the second one belongs to a device visitor. If the two data are nearly identical, the device knows that "visitor" and "owner" are one and the same, and gives access to the person. Its fast phase. Biometric sensor devices example: **Fingerprint scanners**: **optical, capacitive** and **ultrasound**

How are your fingerprints stored?

- Google and Apple store your fingerprint on the device itself and do not make a copy of it on their own servers.
- Apple's TouchID won't store the actual image of the fingerprint, but a **mathematical representation** of it. So even if a malicious hacker reaches this mathematical representation, he cannot reverse engineer it to reveal an actual image of your fingerprint. Not only that, but **the fingerprint data itself is encrypted**.

Biometric Errors

- ***Fraud rate***- The rate at which mis-authentication occurs
- For e. g. Bob poses as Alice and the system mistakenly authenticates Bob as Alice
- Occurs when two people have high degree of similarity
  - Facial features, shape of face etc.
  - Template match gives a score that is higher than the threshold
  - If threshold is increased then false match rate is reduced, but False no match rate is increased
- ***Insult rate***- The rate at which the system fails to authenticate the subject.
- For e.g. Alice tries to authenticate as herself, but the system fails to authenticate her
- occurs for the following reasons
  - Changes in user's biometric data
  - Changes in how a user presents biometric data
  - Changes in environment in which data is presented
- ***Equal error rate-*** rate for which the fraud and insult rates are the same. A way to ***compare*** different biometrics
- For any biometric, can decrease fraud or insult, but other one will increase. For example
  - 99% voiceprint match $\Rightarrow$ low fraud, high insult
  - 30% voiceprint match $\Rightarrow$ high fraud, low insult

Fingerprint

- Fingerprints do not change over time unlike other biometrics.
- Fingerprint image of a person is captured and features are enhanced using various image processing techniques. It records its features like **arches, whorls** and **loops** along with the outlines of **edges**, **minutiae** and **furrows**. This is known as fingerprint recognition.

Loop (double)

Whorl

Arch

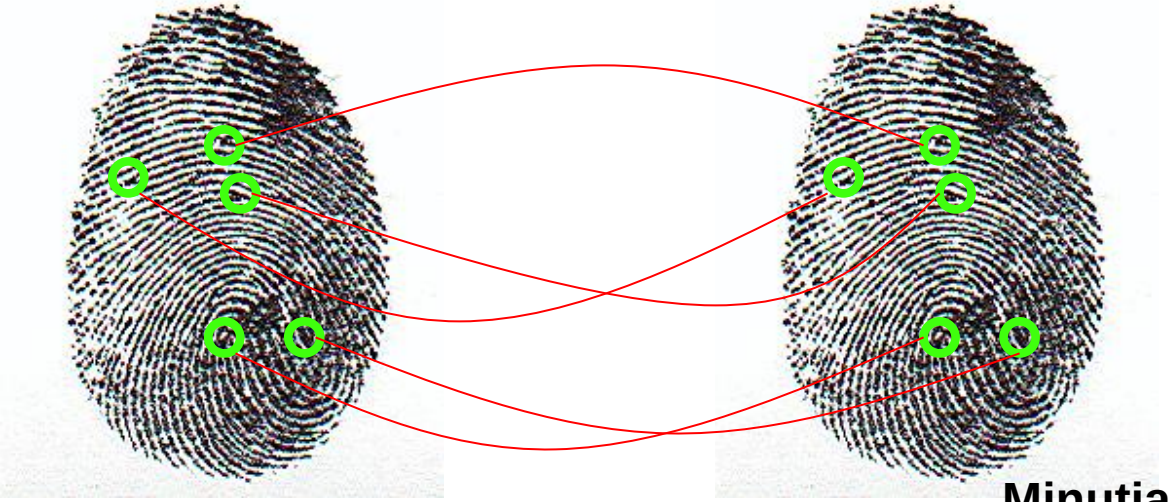
Fingerprint: Enrollment

- Capture image of fingerprint
- Enhance image
- Identify “points”
- Lines that create a finger print is called **ridges** and spaces in between are called **valleys**.

Fingerprint: Comparison

- Fingerprint comparison can be attained in three ways, such as **minutiae**(specific points), **correlation and ridge.**

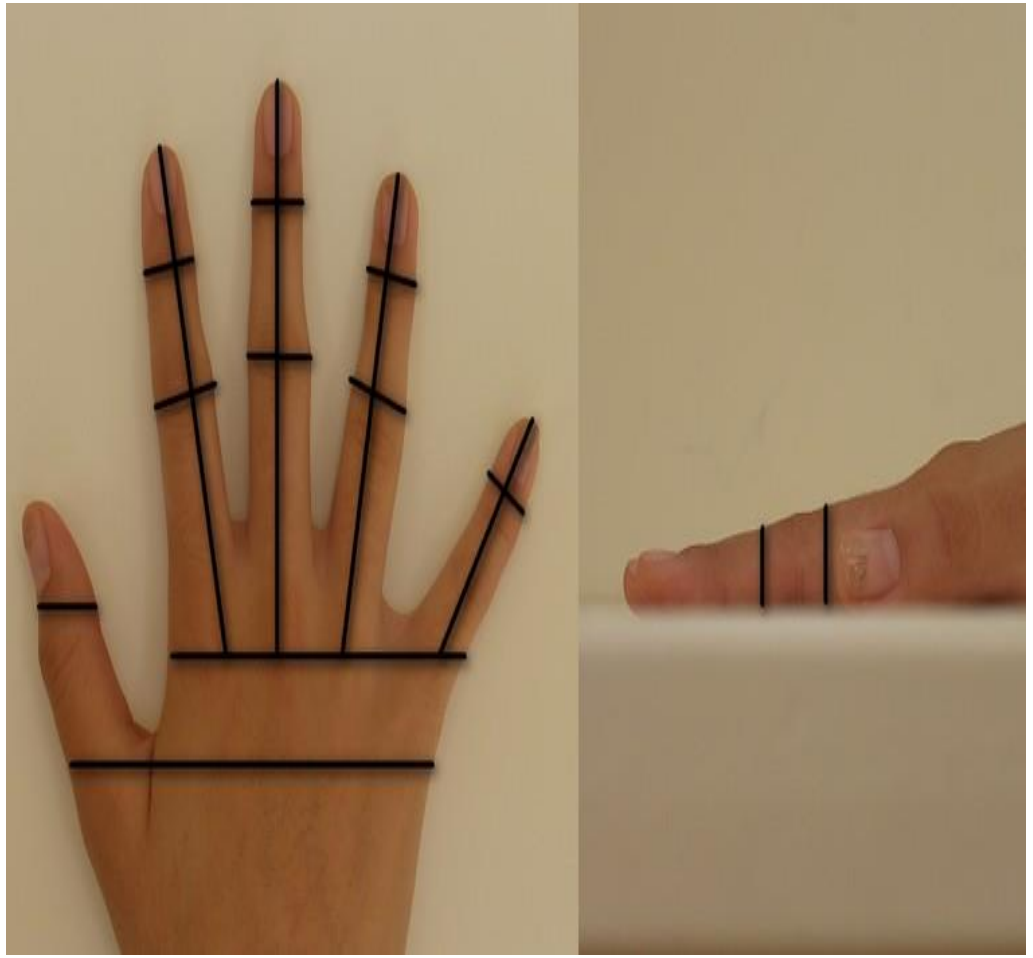
1. **Minutiae based fingerprint** matching stores a plane includes a set of points and the set of points are corresponding in the template and the i/p minutiae.
2. **Correlation based fingerprint** matching overlays two fingerprint images and association between equivalent pixels is calculated.
3. **Ridge feature based fingerprint** matching is an innovative method that captures ridges, as minutiae based fingerprint capturing of the fingerprint images is difficult in low quality.



Minutia comparison

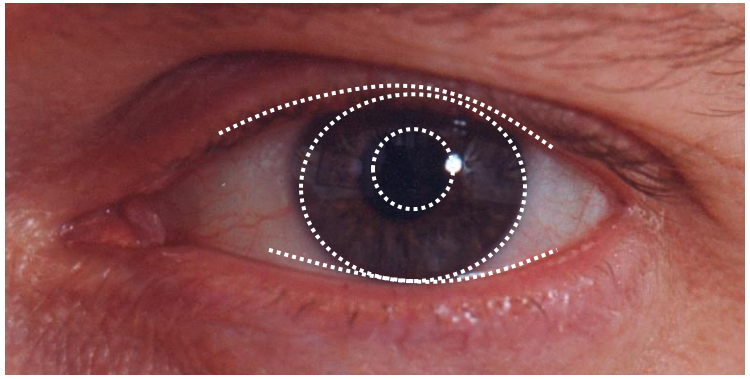
Hand Geometry

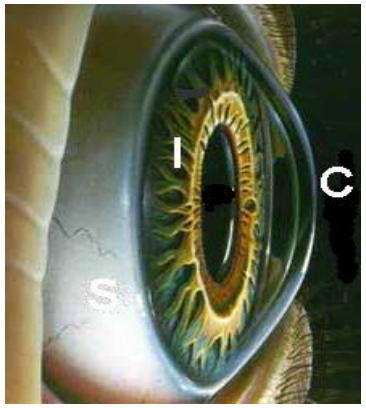
- A popular biometric that measures shape of hand(16 measures)
  - Width of hand, fingers
  - Length of fingers, etc.
- Though human hands not so unique, Hand geometry sufficient for many situations which is quick and robust.
- Not useful for ID problem
- **Advantages**
  - Quick: 1 minute for enrollment, 5 seconds for recognition
  - Hands are symmetric so recognition is fast.
- **Disadvantages**
  - Cannot use on very young or very old
  - Relatively high equal error rate



Iris Patterns

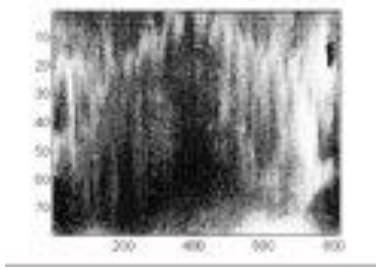
- **Iris recognition** is a one type of bio-metric method used to identify the people based on single patterns in the region of ring shaped pupil of the eye. Generally, the iris has a blue, brown, gray or green color with difficult patterns which are noticeable upon close inspection.
- It is often suggested as ultimate biometric for authentication.
- **Attack possible:** High clarity photographic images of user. To prevent this replay attack, iris scan systems can shine a light on the eye to verify whether pupil is contracting before confirmation.

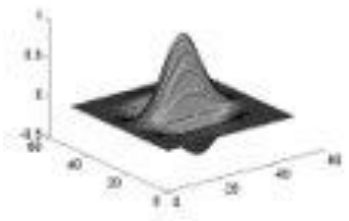


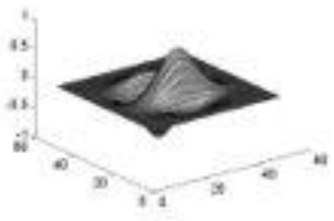


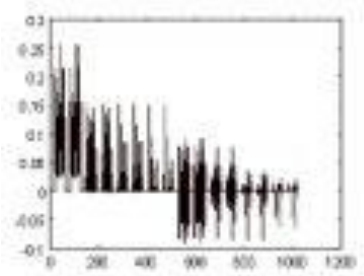
Iris Scan-Recognition Process

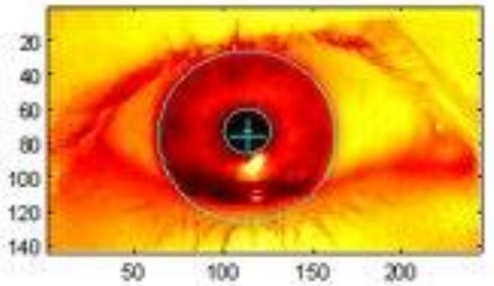
1. An iris scanner first locates the eyes and takes black and white photo of eyes.
2. The resulting image is processed using a 2D-wavelet transform, which results in a 256 byte **"iris code"**.
3. Iris codes are compared based on **hamming distance** between Iris codes.Eg: y-Iris scan of Alice stored in Database and x-iris code computed. Then **d(x,y)=no. Of non-match bits/no. Of bits compared** d(0010,0101)=¾ d(101111,101001)=2/6. A perfect match corresponds d(x,y) to 0. Usually a threshold of acceptance os kept as 0.32 otherwise non-match.





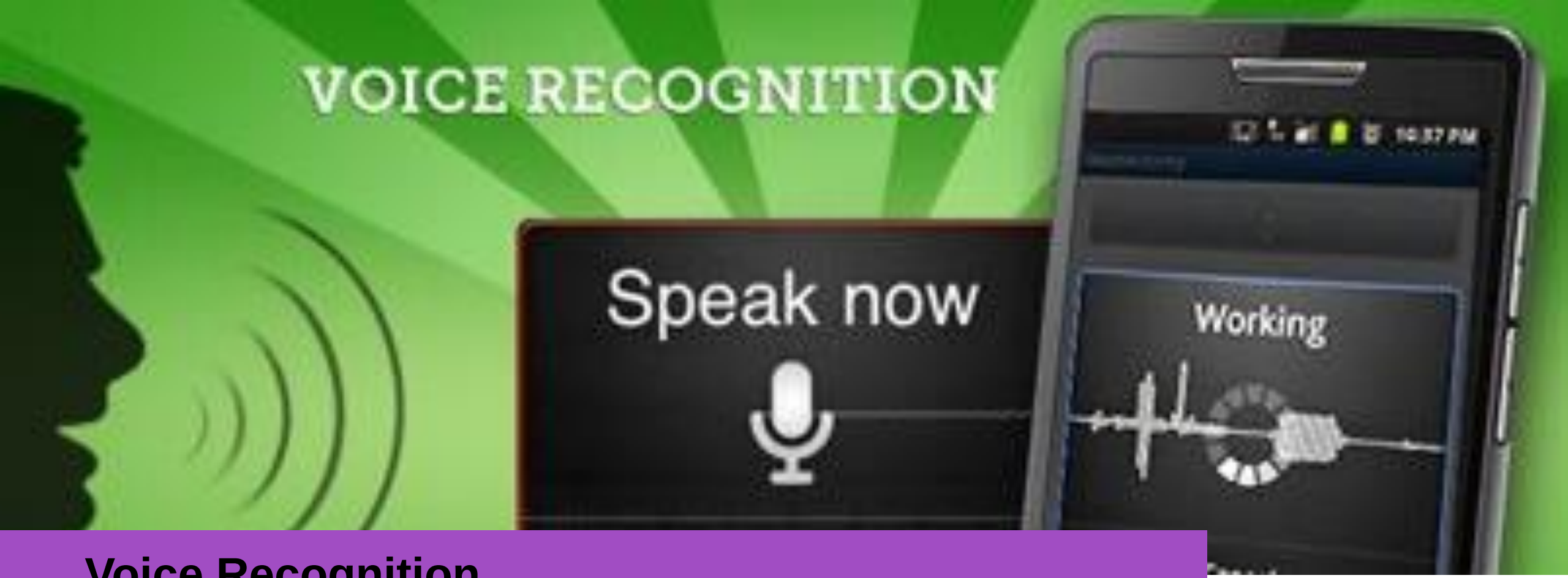






FACE RECOGNITION

- Present facial recognition systems work with face prints and these systems can recognize **80 nodal points** on a human face.
- Nodal points are nothing but end points used to measure variables on a person's face, which includes the length and width of the nose, cheekbone shape and the eye socket depth.



Voice Recognition

- Voice recognition technology is used to produce speech patterns by combining behavioral and physiological factors that can be captured by processing the speech technology. The most important properties used for speech authentication are **nasal tone, fundamental frequency, inflection, cadence.**
- Voice recognition can be separated into different categories based on the kind of authentication domain, such as a **fixed text method**, in the text dependent method, the **text independent method** and **conversational technique**.

Equal Error Rate Comparison

- **Equal error rate (EER):** The point at which fraud rate == insult rate
- **Fingerprint** biometrics used in practice have EER ranging from about 10^{-3} to as high as 5%
- **Hand geometry** has EER of about 10^{-3}
- In theory, **iris scan** has EER of about 10^{-6}
  - Enrollment phase may be critical to accuracy
- Most biometrics much worse than fingerprint!
- Biometrics useful for authentication…
  - …but for identification, not so impressive today

Biometrics: The Bottom Line

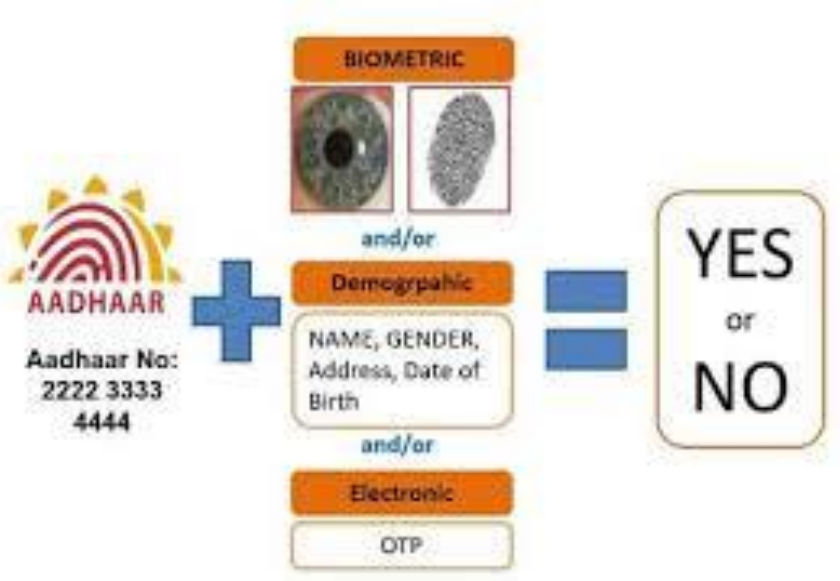
- Biometrics are hard to forge
- But attacker could
  - Steal Alice's thumb
  - Photocopy Bob's fingerprint, eye, etc.
  - Subvert software, database, "trusted path" …
- And how to revoke a "broken" biometric?
- **Biometrics are not foolproof**
- Biometric use is relatively limited today
- That should change in the (near?) future

Something You Have: SMART CARDS

- A smart card looks like credit card but includes small amount of memory and computing resources so that it can store cryptographic keys or other secrets and able to do some computations.
- Other Examples of "something you have " authentication: ATM Card, Password generator, ADHAR

AADHAR

- India's **Aadhaar project** is emblematic of biometric registration
- Aadhaar number is a 12-digit unique identity number issued to all Indian residents. This number is based on their biographic and biometric data (a photograph, ten fingerprints two iris scans).
- Initially the project has been linked to **public subsidy** and **unemployment benefit schemes** but it now includes a **payment scheme**.
- Financial inclusion is expected to be a key application of Aadhaar authentication with **Aadhaar Pay**, a new payment scheme announced in 2017.



TOKEN BASED AUTHENTICATION

- A security token is a peripheral device used to gain access to an electronically restricted resource.
- An authentication token is a small device that generates a new random value every time it is used, which is an alternative to password that can be used for authentication.
- Each authentication token is pre-programmed with a unique number called random seed. It ensures the uniqueness of the output produce by the token.

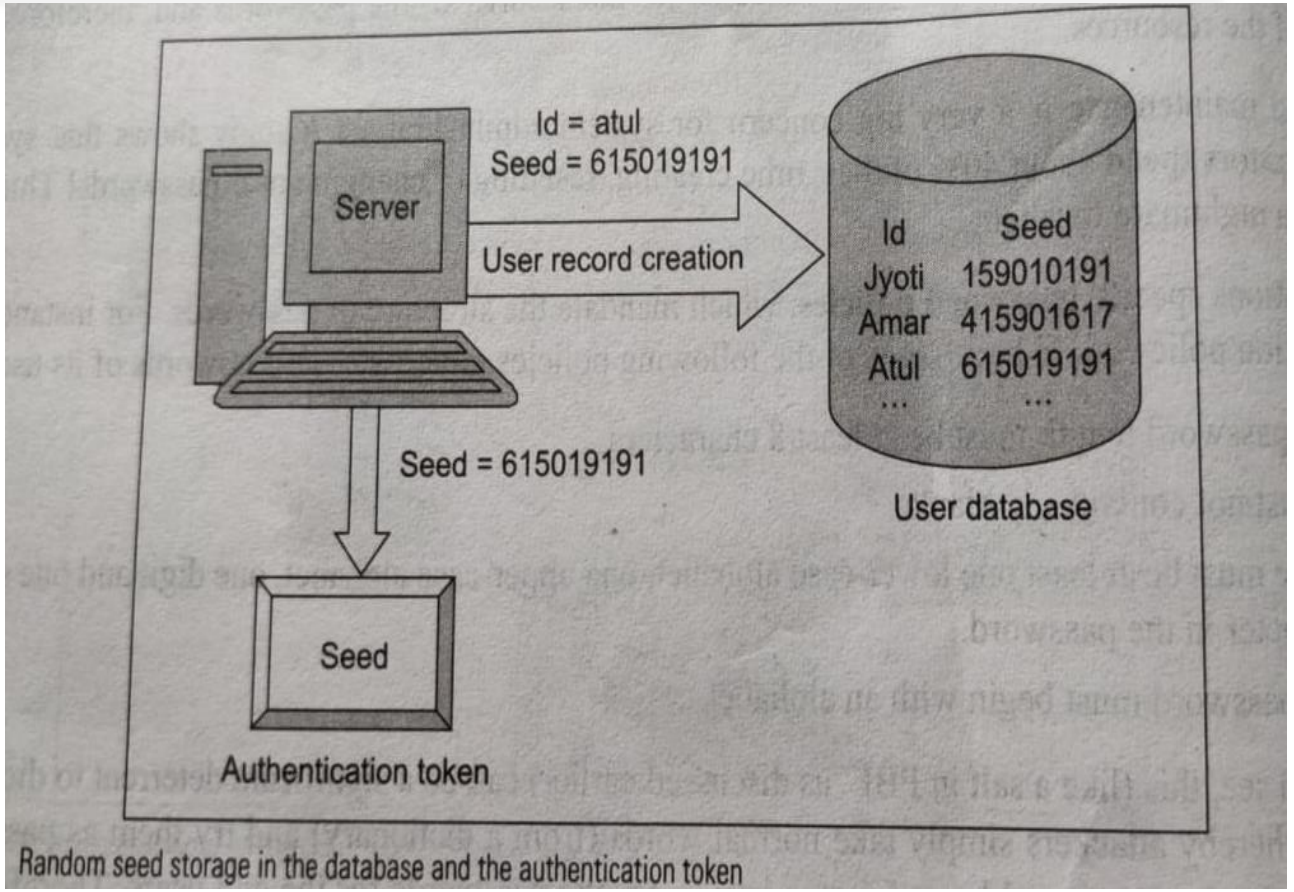
Authentication Token Device

- Authentication Device has the following features:
  - Processor
  - LCD for displaying outputs
  - Battery
  - Optionally a small keypad for entering information
  - Optionally a real-time clock

Working of Authentication Token

1. Creation of a Token

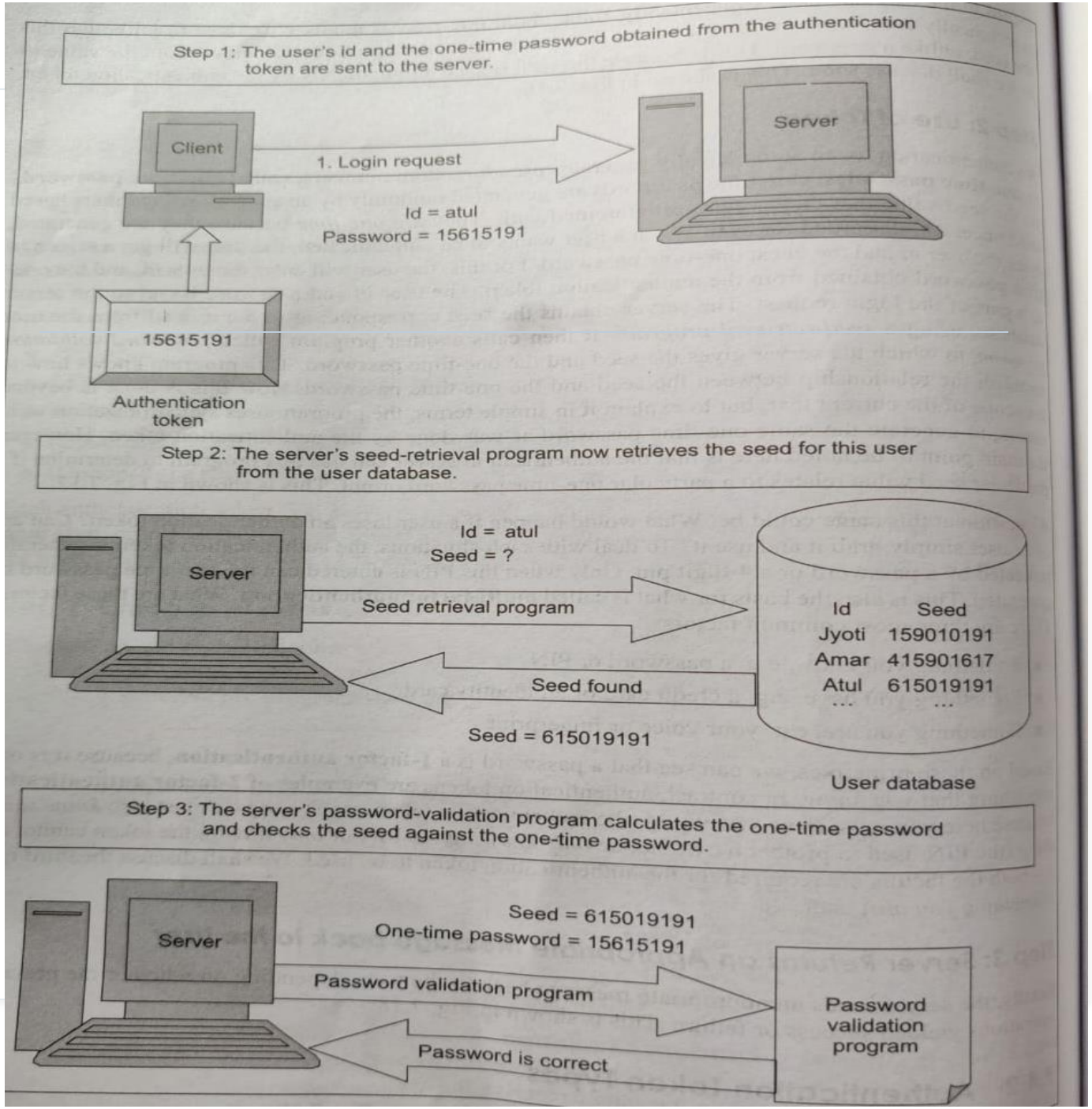
- Authentication tokens are created by Authentication servers along with random seed for token. It is automatically placed or pre-programmed inside each token by the server.
- Server also keeps a copy of the seed against the user ID in the user database.
- Seed can be conceptually considered as a user password. Difference is that the user password is known to the user, seed value remains unknown to the user



Random seed storage in the database and the authentication token

2. Use of the Token

- An Authentication Token automatically generates pseudorandom numbers called one-time passcodes or one time passwords.
- One time passwords are generated randomly by authentication tokens using the seed value.
- When a user wants to be authenticated by any server, the user will get a screen to enter user ID and the latest onetime password. The users enters its ID and gets is latest one-time password from the authentication token. The user ID and password travels to the server as a part of the login request.
- Server obatins the seed corresponding to the userid from database and verifies using some mechanism that this one-time password is created using the valid seed value.



Security

- **What happens if user loses an authentication token? Can another user grab token and use it?**
- Authentication token is generally protected by a password or a 4-digit PIN. Only when PIN is entered one-time password will be generated.
- Hence authentication tokens are 2-factor authentication techniques.

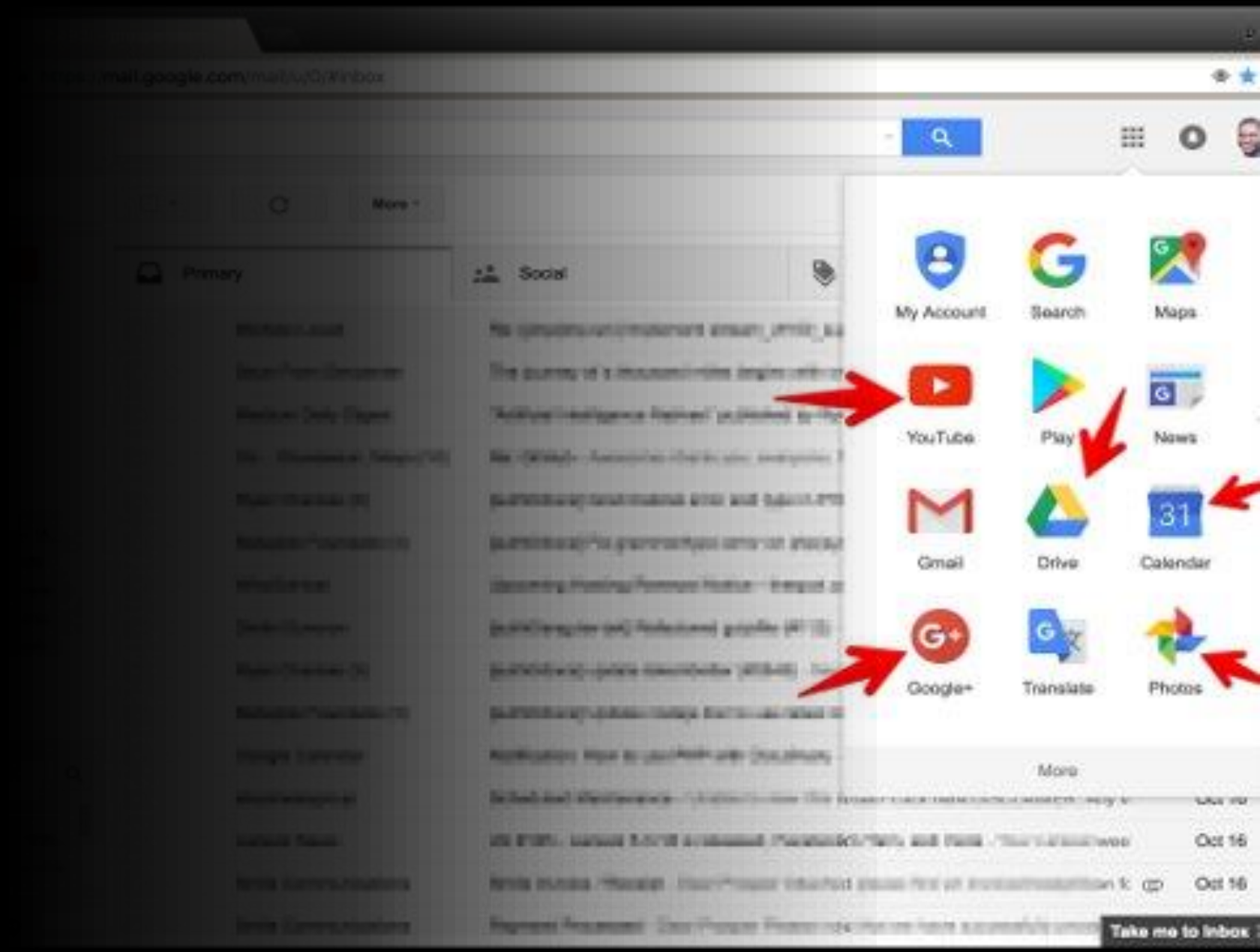
Types of Authentication Tokens

1. Challenge Response Tokens
2. Time based Tokens

2-factor Authentication

- Requires any 2 out of 3 of
  - Ø Something you **know**
  - Ø Something you **have**
  - Ø Something you **are**
- Ø Examples
  - Ø ATM: Card and PIN
  - Ø Credit card: Card and signature
  - Ø Password generator: Device and PIN
  - Ø Smartcard with password/PIN

Single Sign On



Single Sign-On

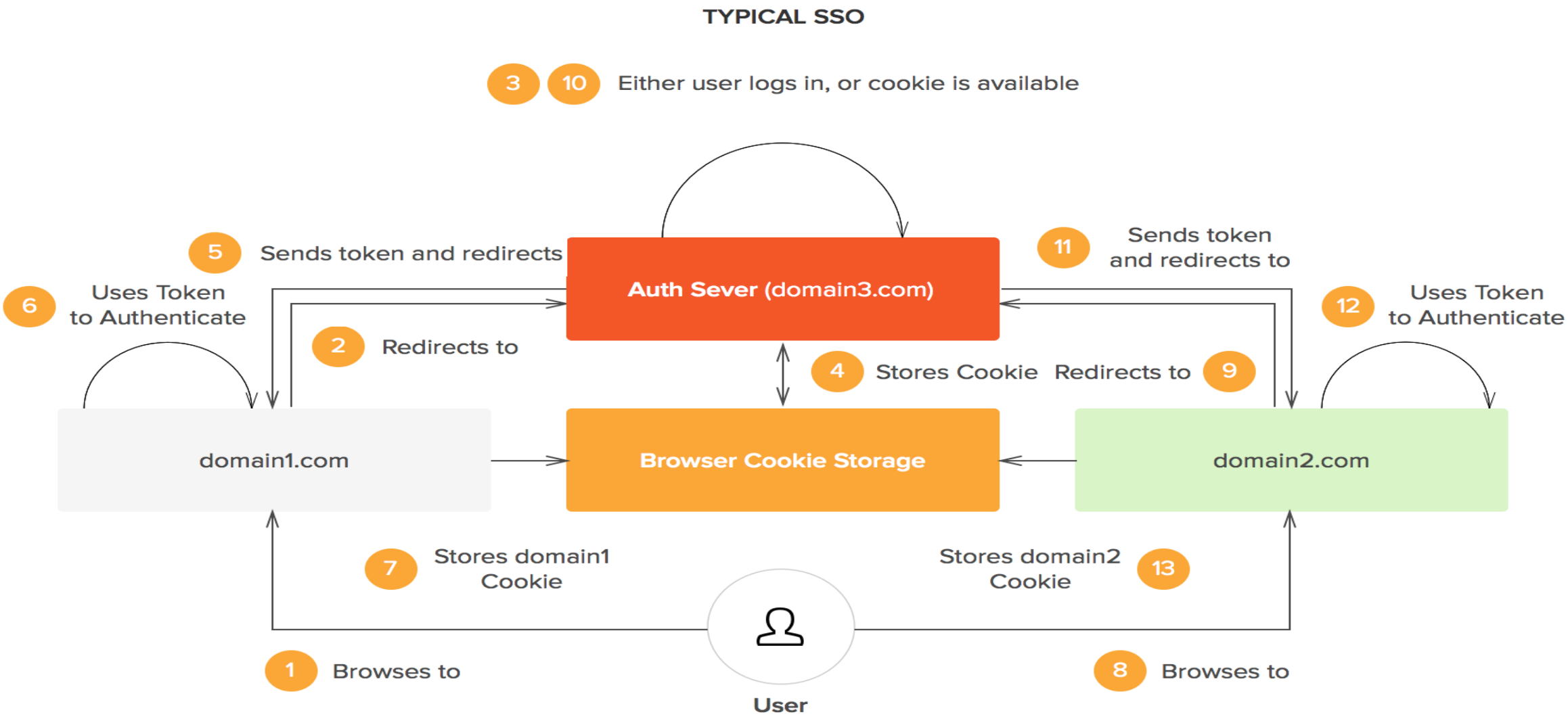
- **Single Sign On (SSO) login** refers to when a user logs in to an application with a single set of credentials and is then automatically signed into multiple applications where the "credentials" stay with them wherver they go on Internet. With SSO login, a user gains access to multiple software systems without maintaining different login credentials such as usernames and passwords.
- Single sign-on (SSO) is a property of access control of multiple, related, but independent software systems. Single sign-off is the reverse property whereby a single action of signing out terminates access to multiple software systems.
- As different applications and resources support different authentication mechanisms, single sign-on has to internally translate to and store different credentials compared to what is used for initial authentication.
- Initial authentication requires person's participation, but subsequent authentications would happen behind the scenes.
- **EXAMPLE:** Google accounts, auth0

Key Benefits of SSO Login

- **Eliminate the time spent** re-entering user credentials, thus improving productivity for users and increasing conversion rates for product owners
- **Eliminate password fatigue** from having to store or remember different usernames and passwords.
- **Reduce complaints about password problems**, thus reducing the costs associated with setting up several helpdesk systems for password-reset issues, invalid credentials, etc.
- **Minimize phishing**, thus improving security.
- Streamlines the local, desktop, and remote application workflows, thus improving users' productive capacity.

How SSO Login Works

- **Session** is shared with other domains in some secure way e.g., a signed JSON Web Token (JWT).
- **GIVEN SCENARIO:** Three apps have been developed separately: **App FOO**, **App BAR**, and **App BAZ**. They are also hosted on three different domains: **foo.com**, **bar.com** and **baz.com** respectively.
- With SSO login, a *central authentication server exists which is* **foobarbaz.com**.
    1. The user accesses **foo.com**.
    2. The user is redirected to **foobarbaz.com**, where an authentication-related cookie is generated.
    3. The user navigates to **bar.com**.
    4. The user is redirected to **foobarbaz.com**.
    5. **foobarbaz.com** checks whether the user already has an authentication-related cookie and redirects the user back to **bar.com**, providing access to its features and content.
    6. The same process applies to **baz.com**.


TYPICAL SSO
3
10
Either user logs in, or cookie is available
5
Sends token and redirects
Auth Sever (domain3.com)
11
Sends token and redirects to
6
Uses Token to Authenticate
2
Redirects to
4
Stores Cookie
Redirects to
9
12
Uses Token to Authenticate
domain1.com
Browser Cookie Storage
domain2.com
7
Stores domain1 Cookie
Stores domain2 Cookie
13
1
Browses to
8
Browses to
User

Web Cookies

- An **HTTP cookie / web cookie/ Internet cookie/ browser cookie/ cookie** is a small piece of data (in the form of an HTTP header that consists of a text-only string.)sent from a website (Web Server)and stored on the user's computer by the user's web browser while the user is browsing.
- Browser stores each data in a small file, called cookie.txt. Data is stored as name-value pairs. a Web site might generate a unique ID number for each visitor and store the ID number on each user's machine using a cookie file.
- browser accesses the cookie file again when you visit the website that created the cookie file. The browser uses the information stored in the cookie file to help ease your navigation of the website by letting you log in automatically or remembering settings you selected during your earlier visits to the website, among many other functions.
- **Who creates cookie?**
- cookie file is generated by the site you're browsing and is accepted and processed by your computer's browser software.
- **Is it a program?** No. They cannot gather any information on their own.
- **NOTE:** Only website that created cookie can read the cookie file it has created.

Cookie Parameters

- **Cookies have six parameters that can be passed to them:**

1. The **name** of the cookie.
2. The **value** of the cookie.
3. The **expiration** date of the cookie - this determines how long the cookie will remain active in your browser.
4. The path the cookie is valid for - this sets the **URL** path the cookie us valid in. Web pages outside of that path cannot use the cookie.
5. The **domain** the cookie is valid for - this takes the path parameter one step further. This makes the cookie accessible to pages on any of the servers when a site uses multiple servers in a domain.
6. The **need for a secure connection** - this indicates that the cookie can only be used under a secure server condition, such as a site using SSL.

Necessity of cookies

- **Tracking**
  - **To facilitate** hassle-free automatic logins and authentication
  - **To monitor advertisements ie** third party ad serving, ad management,
  - **To collect demographic information about who is visiting the Web site:** a web server can gather information about which web pages are used the most, and which pages are gathering the most repeat hits.
  - **a website can infer that the user has already visited.**
- **Personalization/Better User Experience(user preferences):** Servers can use cookies to provide personalized web pages. When you select preferences at a site that uses this option, the server places the information in a cookie. When you return, the server uses the information in the cookie to create a customized page for you.
- **Session management :**Logins, shopping carts, game scores, or anything else the server should remember Cookies are also used for online shopping. Online stores often use cookies that record any personal information you enter, as well as any items in your electronic shopping cart, so that you don't need to re-enter this information each time you visit the site. e-commerce sites, the unique ID can be used to set up the shopping cart for a user. A user can add items to a shopping cart and this preference is linked to the site's database along with the unique id. Later, when the same user visits the site, this unique id from cookie can be used to retrieve the shopping preference of that particular user.

Who earns with COOKIE?

- **Cookie-based ad tracking :** helps in user profiling/website preference tracking.
- Many largest websites online use large-scale third-party ad serving networks which cover many sites.
- Googles Adsense/Adwords ad serving network. Literally, millions of pages run Adsense ads. For every click a valid user makes on a Google-served ad on their site, site owners make money ranging from pennies to dollars.
- Webmasters have always been able to track access to their sites, but cookies make it easier to do so. In some cases, cookies come not from the site you're visiting, but from advertising companies that manage the banner ads for a set of sites (such as DoubleClick.com). These advertising companies can develop detailed profiles of the people who select ads across their customers' sites.
- Accepting a cookie does not give a server access to your computer or any of your personal information (except for any information that you may have purposely given, as with online shopping). Also, it is not possible to execute code from a cookie, and not possible to use a cookie to deliver a virus.

Types of Cookies

1. **Transient Cookies**
   - This type of cookie is stored in the computer's memory, i.e. RAM, and contains a session ID that stores information on the user's browsing session. These cookies are also known as **session cookies** and are deleted from the computer as soon as the browser is shut down. The session id in a session cookie acts as the user id that allows navigating from one page to another without logging in repeatedly. These session cookies will become inaccessible if the session becomes inactive for a specific period of time.
2. **Persistent Cookies**
   - These types of cookies are stored on the computer's hard drive and contain information on user preferences for a website. These preferences can be used for further browsing sessions and are retained as long as the user allows them. Unlike transient cookies, these cookies are permanently stored and are not deleted when a browser is shutdown.
3. **Flash Cookies**
   - These cookies are small Flash files such as video clips or gifs stored on a user's computer by websites. They contain the same data as normal cookies and they function in the same way. Unlike other cookies, Flash cookies cannot be deleted by any mechanism from the browser level. They can only be deleted manually or by using some special add-ons. So, if a Flash cookie is retained on a computer's hard disk, it will be used as a backup for the normal cookie and hence can be used to recall the preferences of a particular user to a website.

Risk Associated with Cookies

- Common risks of cookies are:
  1. **Cross Site Request Forgery Attack (XSRF):** When a website receives a request, it cannot distinguish whether the action is initiated by the user or not. user visits a legitimate site and receives a legitimate cookie. The user then visits a malicious site that instructs the user's browser to perform some action targeting the legitimate site. The legitimate site receives the request along with the legitimate cookie and performs the action since it appears to be initiated by a legitimate user.
  2. **Session Fixation:** an attacker impels the user to use the attacker's or another's session ID. This can be done by using the cookie's browser directive path, hence the user pretends to be someone else. Using this method, an attacker can urge the user to log in as the attacker on various application levels. a user receives a malicious cookie that contains the cookie issuer's session ID. When the user attempts to log into a targeted domain, the issuer's session ID is logged in instead of the user's session ID. In this way, it looks to the targeted domain like the issuer is performing actions that the user is actually performing.
  3. **Cross-Site Scripting:** a user visits a malicious website and receives a cookie that contains a script payload targeting a different website. The malicious cookie is disguised to look like it originated from the targeted website. When the user visits the targeted site, the malicious cookie, including the script payload, is sent to the server hosting the targeted site. This occurs when an attacker takes advantage of a website that allows its users to post unfiltered HTML and JavaScript content. By posting malicious HTML and JavaScript code, the attacker can cause the victim's web browser to send the victim's cookies to a website the attacker controls.
  4. **Cookie Tossing Attack:** a user visits a malicious site that provides a cookie designed to look like it originated from a subdomain of a targeted site, such as *http://subdomain.example.com*. When the user visits the targeted site, *http://example.com* in this case, the subdomain cookie is sent along with any legitimate cookies. If the subdomain cookie is interpreted first, the data in that cookie will overrule the data contained in any subsequent legitimate cookies.**Cookie Overflow Attack**
  5. **Session Hijacking/Cookie Hijacking**
  6. some viruses and malware may be disguised as cookies: "supercookies","zombie cookies"(a cookie that recreates itself after being deleted**)**

COOKIE LOCATION

- **Where is the location of the Cookies folder?**
- **Firefox cookies are stored in "cookies.txt"**
  at C:\Users\Default\AppData\Roaming\Microsoft\Windows
- **IE**
- C:\Documents and Settings\<User name>\Local Settings
- **Chrome**
- C:\Documents and Settings\<user name>\Local Settings\AppData\Google\Chrome\User Data\Default\Cookies
- C:\Users\sdk\AppData\Local\Google\Chrome\User Data\Default

Attack Prevention

- **Browser setting controls:** Every browser gives you a range of options for handling cookies. It lets you delete existing cookies in a single click and choose how future cookies are collected or stored.
  - **Banning all cookies**? **NO**. makes some websites difficult or impossible to navigate. However, a **setting that controls** or limits third-party and tracking cookies can help protect your privacy while still making it possible to shop online and carry out similar activities.
- **Use https method of authentication** wherever possible: Network traffic analysis can be done by intruder, which includes cookies sent on ordinary unencrypted HTTP sessions. MITM attack, Evesdropping
- Even if the cookies are stolen, developers should come up with codes that **check the source address and referrer of the authenticated user periodically**. If there is a **mismatch, the session need to be closed suddenly.**
- The **expiry time of the cookie** that is allocated from the web server needs to be **shortened.**

Thank you

anooja@somaiya.edu