Student name: Yashasvi Chaurasia Student ID: 2020159

CSE 345/545 Foundations to Computer Security

Mid-Sem Exam

Deadline: 1500hrs, October 17 2022

Q1.

a) Assuming Country and number of Items purchased to be Sensitive Data while the rest to be insensitive Data we can k-anonimize the data as follows:

Suppress Name ,generalize price and customer id for k=2

				S	S	
Customer ID	Name	Place	City	country	No Items purcha	price
c00100<	*	***	***	uk	2	1000-7000
c00100<	*	new york	new york	us	2	1000-7000
c00100<	*	new yok	new yok	us	3	1000-7000
c00100<	*	***	***	in	2	1000-7000
c00100<	*	***	***	ca	1	7000-14000
c00100<	*	new york	new york	us	3	1000-7000
c00100<	*	***	***	aus	2	7000-14000
c00100<	*	***	***	aus	1	7000-14000
c00100<	*	chenn	chenn	in	1	7000-14000
c00100<	*	mum	mum	in	1	7000-14000
c00100<	*	chenn	chenn	in	1	7000-14000
c00100<	*	mum	mum	in	2	7000-14000

For k=3

We suppress price and name .

				S	S	
Customer ID	Name	Place	City	country	No Items purcha	price
c00050<	*	*	*	uk	2	*
c00050<	*	new york	new york	us	2	*
c00050<	*	new yok	new yok	us	3	*
c00050<	*	*	*	in	2	*
c00050<	*	*	*	ca	1	*
c00050<	*	new york	new york	us	3	*
c00050<	*	*	*	aus	2	*
c00050<	*	*	*	aus	1	*
c00050<	*	*	*	in	1	*
c00050<	*	*	*	in	1	*
c00050<	*	*	*	in	1	*
c00050<	*	*	*	in	2	*

b)

Two methods of increasing utility of data :

1. Try to generalize rather than suppressing the data as it helps in increasing the utility of the data.
   For example I used generalization in for table with k=3 and have the following anonimization still with k=3 but with generalized price column.

				S	S	
Customer ID	Name	Place	City	country	No Items purcha	price
c00050<	*	*	*	uk	2	1000-7000
c00050<	*	new york	new york	us	2	1000-7000
c00050<	*	new yok	new yok	us	3	1000-7000
c00050<	*	*	*	in	2	1000-7000
c00050<	*	*	*	ca	1	7000-14000
c00050<	*	new york	new york	us	3	1000-7000
c00050<	*	*	*	aus	2	7000-14000
c00050<	*	*	*	aus	1	7000-14000
c00050<	*	*	*	in	1	7000-14000
c00050<	*	*	*	in	1	7000-14000
c00050<	*	*	*	in	1	7000-14000
c00050<	*	*	*	in	2	7000-14000

2. For the suppressed data we can carefully change the data value to something random rather than removing the data completely. This can help us to balance the tradeoff between the level of anonymization and the data utility.

For Example:

				S	S	
Customer ID	Name	Place	City	country	No Items purcha	price
c32	sdhkahds	*	*	uk	2	1000-7000
c34	asd	new york	new york	us	2	1000-7000
c36	hgfs	new yok	new yok	us	3	1000-7000
c44	rydv	*	*	in	2	1000-7000
c45	asd	*	*	ca	1	7000-14000
c47	axscc	new york	new york	us	3	1000-7000
c48	sdc	*	*	aus	2	7000-14000
c49	ad	*	*	aus	1	7000-14000
c22	cdsfdc	*	*	in	1	7000-14000
c11	adfc	*	*	in	1	7000-14000
c13	adfaaas	*	*	in	1	7000-14000
c42	sdax	*	*	in	2	7000-14000

c) Sample Dummy Data: Education Quality and Job Opportunity are sensitive data which can be used to analyse the institutes performance.

According to GDPR

					s	s	
Name	City	Batch	Branch	Gender	Education Quality	Job Opportunity	Income
Raj	Delhi	2012	CSB	m	9	10	20L
Ram	Bangalore	2014	CSE	m	8	8	3Cr
Sunita	New York	2009	CSB	f	7	7	42L
Priyam	Lucknow	2018	CSE	m	8	9	1Cr
hiren	kolkata	2009	CSAM	m	10	10	67L
Rani	Delhi	2012	CSAM	f	7	8	51L

					s	s	
Name	City	Batch	Branch	Gender	Education Quality	Job Opportunity	Income
*	**	***	CSB	*	9	10	Below 50L
*	**	***	CSE	m	8	8	Above 50L
*	**	***	CSB	*	7	7	Below 50L
*	**	***	CSE	m	8	9	Above 50L
*	**	***	CSAM	*	10	10	Above 50L
*	**	***	CSAM	*	7	8	Above 50L

I applied suppression and generalization to remove all personal identifiers from the data and used 2 k anonimity. GDPR regulations require us to remove any personal identifiable information like name ,location, address,etc

The dummy data is useful to help organization learn about the progress of the institute and how the students feel about the institute about the quality of education and the impact the institute has had on their lives by providing them opportunities

Q2.

a) Ipv4 Addresses are 32 bits long using numeric characters. Hence, the total number of Plausible Ip addresses generated by Ipv4 standards is insufficient to map all the devices in the current network architecture. The Ipv4 addresses were designed just to identify addresses over a network and not designed to implement security features hence security is dependent on the application.

Ipv6, on the other hand, uses 128-bit addresses, which use hexadecimal characters hence they can address a very large number of addresses which are enough to map all the devices for the current and future network architecture. Ipv6 addresses have mandatory security features implemented with the addresses.

The current Network Architecture uses NAT and many legacy protocols, such as legacy FTP which cannot support Ipv6 addresses as they are longer than the Ipv4 addresses and may cause conflicts with current protocol syntax. Some Internet applications are built to work with only ipv4 addresses. Ipv4 addresses cannot be routed directly to Ipv6 addresses; hence, we need the entire network structure to be Ipv6 before we move on to Ipv6. Once Ipv6 is adopted completely, all the currently NAT-based network architecture will be redundant.

Ipv6 is more secure than Ipv4 as Ipv6 follows IPsec protocols by default; security in ipv4 is optional, and in ipv4, packets need to be modified, and their source or destination Ip address is changed at NAT before they are sent to the server for processing the data. A corrupted NAT device could transmit packets or retransmit packets to incorrect Ip addresses, posing a security threat as the packets are being modified before they reach the servers. In Ipv6 addresses, we don't need NAT devices, and the packets also cannot be modified while they are being transmitted hence Ipv6 comes out to be better at security than IPv4.

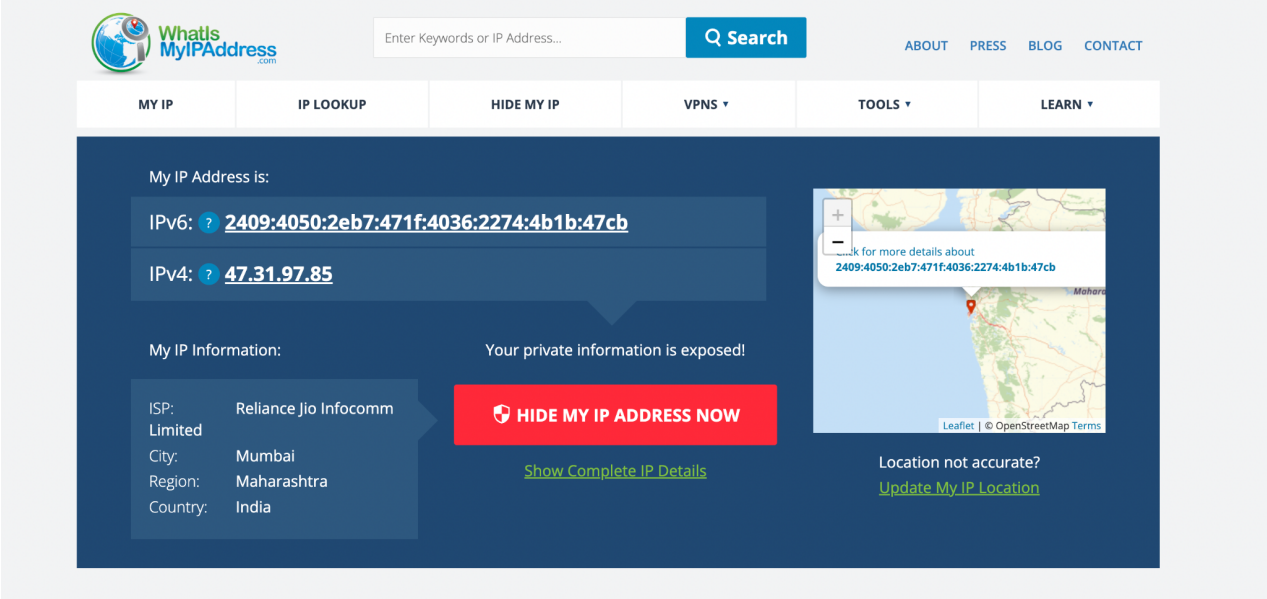
b) IPSec is the set of security protocols that securely transmit data over the internet. For the transportation of data, the IPsec follows two protocols, AH and ESP, where AH secures the entire packet along with the packet header. Hence, no part of the packet can be modified during transmission, while the ESP protocol secures the packet, excluding the header of the packet. For Ipv4, we cannot even use ESP protocol even if it does not encrypt the header as after changing the header by NAT , the NAT cannot modify the checksum inside the encrypted payload and hence the packet will be dropped as the checksum fails hence we cannot use IPsec protocols with Ipv4.

c) Assuming that all the devices have a unique Ipv6 address, we can set a specific threshold for the number of packets per second an IP address can send to a device; once the number of packets per second has crossed the threshold limit, we can either blacklist that IP from further communication for a certain period or we can simply drop extra packets once the threshold is reached. As we are using ipv6, all the incoming packets can be identified and authenticated to the source from where

the packet has arrived, and hence we can blacklist that IP from sending any further request; we would not have been able to blacklist IPs if the network was using ipv4 addresses as then we would be blocking the entire NAT device or a public IP (blocking many innocent clients rather than a single attacker)and not the exact device generating the flood attack.
If the network used Ipv4 addresses, we could follow the above approach, but instead of blocking an IP, we could drop packets once the threshold is reached.

d) Assuming we are using Ipv6 addressing for all our connections.
Between Wifi Router and Router I will use IPsec in tunnel mode to create a secure encrypted tunnel.
Also for Router to the internet all the communications will be done using Ipsec in tunnel mode to ensure no on is able to listen on the packets during transmission .
To establish the secure connections I will use IPsec IKEv2 which uses a version of Diffie Hellman key exchange to securely transmit the keys.
For all the devices within the network we can use Ipsec with AH along with ESP protocol to authorize and encrypt the packets in transmission so that we can verify if the packet was modified or tampered during its transmission period.
For all the key exchanges we will use IKE or Diffie hellman key exchange so so that we donnot need to perform a offline key exchange or use any third party service.

e) Both the ipv6 and ipv4 addresses are shown. As for current scenarios, we follow a dual-stack architecture where we follow both ipv4 and ipv4 addresses to avoid compatibility issues between ipv4 and ipv6 addresses.



Ipv4 addresses are hidden behind the addresses of a NAT device, and hence in ipv4, the actual private IP is abstracted behind the NAT device.
The current default IP addresses being used are the ipv4 addresses and ipv6 as we still rely on NAT devices and DNS, which need to accommodate ipv6 addresses, so sticking with the current stack of ipv4 addresses could be beneficial for continuing legacy applications.

In ipv6, each request can be retraced back to the source of information, which provides authentication for each piece of information.
Also, Ipv6 addresses have security protocols enabled to avoid packet tampering, and packets cannot be modified while in transit.
Ipv6 protocols address the problem of limited IP addresses for each device on the network.

Q3.

a) No, the protocol mentioned above is not secure, as it is vulnerable to MITM attacks.
A third user can intercept the challenge and send it to Alice, impersonating a server.
Alice then encrypts and sends back the challenge to the third person, who authenticates with Alice and sends Alice's encrypted message back to the server.
In this way, the third person authenticates itself with Alice and continues to act as a server for Alice while it acts as Alice to the server.

b) Bob should send a certificate with its digital signature encrypted with bobs private key on it so that anyone who tries to tamper with the communication or perform a MITM attack can be easily avoided, and the authentication would fail as the hash for the document won't match with the digital signature. Yes, this will impact the protocol as this R would secure the communication between bob and Alice and prevent MITM and IP spoofing attacks.

c) Yes, it is possible to perform MITM and Ipspoofing attacks even on secure SSL/TLS connections as if initially a third party IP spoofed and pretended to be Alice and the server then created a secure connection with the third person. The third person can further generate a secure connection with Alice and be a part of a MITM attack pretending to be the server for Alice, even on an SSL/TLS connection.
This 3rd person performs a MITM attack when the connection is being transferred from HTTP to HTTPS in other words; the third party is able to get between the communication not by breaking into the communication but by being part of the communication by faking an identity.
SSLstripping generates a fake certificate of the server we want to mimic, which we use to mimic as server and generate a secure ssl/tls connection with Alice.

d) Yes, Mutual authentication can be performed if one party has a verifiable certificate, then the verifiable certificate can be shared with the other party along with a digital signature where the other party can then verify the correctness of the certificate.
The digital certificate contains the transmitter's public key and other details. The sender then hashes the certificate, encrypts it with its private key, and sends it to the receiver, then the certificate is hashed, and the generated hash is verified with the attached encrypted hash to check the server's authenticity. The certificates use X.509 public key infrastructure
But this can still be compromised with SSLstripping and many other attacks as only one side of the communication is being verified. The client is not being checked for authenticity.

Q4.

a) Given persistent resources and transient users, we should use ACL as files won't change over some time hence we can quickly lookup if a user has permissions by looking up in the ACL dictionary and then quickly add or remove permissions to that user. ACL revokes permission from the users; hence, it is preferred for the current use case as we would need to add and remove permissions quickly for many transient users.

b) Yes, an unauthorized user can misuse privileges of Alice. File Y has only read access to X and hence it cannot have contents from X written into it as even X has only read access to file Y but an unauthorized user can use the read privileges of Alice via Z from file X and then use Alice's write privileges to write into file Y.Assuming that By unauthorized user is a user who is accessing Alice's permissions without being authorized as Alice.

c) Such unauthorized access of files can be prevented by using the bell LaPadulas model where write-up is allowed but read-down is prohibited hence in the above example File Z can write to Y but it cannot read from file X. For this, to work we have to assign levels of accessibility to files and assign them roles of access as top secret, secret and unclassified.

   Read-down is allowed in Bell LaPadula Model and while write-down is not allowed. Similarly, Read up is not allowed while write-up is allowed in this model hence we can easily avoid the issues caused by access of unauthorized users.

Q5.

a) Perfect Anonymity depends on the user and how they use it. A user could be anonymous on the internet. Still, there are always methods that can deanonymize a user, like analyzing traffic on the tor network if we have control over multiple tor nodes.
   A user needs to know the correct know-how/knowledge to use tools to anonymize themselves. The user should have the correct tools, and the user should have a secure method of payment on the network, which can help them stay unidentifiable on the network.
   A user can always ensure safety, but there can be instances where the user can be at fault for human errors and hence might reveal their identity a user using tor to access Twitter is anonymous during the access, but the moment the user logs in to their account they are identified and hence mark their footprint on the network. Similarly, corrupted tor exit nodes can identify traffic details along with intermediate nodes to de-anonymize a user hence the concept of perfect anonymity is user and use-case dependent.
   Also, it is said not to be completely anonymous as NSA has embedded backdoors in TOR to deanonymize any user on tor as per:https://restoreprivacy.com/tor/

b) Tor relies on three principles of anonymity:
1. User should have access to correct tools like TOR to access the web securely as tor uses layers of encryption to prevent anyone in between the network from eavesdropping on the communication, normal browsers don't perform onion routing and hence they reveal a lot of information which can be used to deanonymize a user.

2. User should have a secure method of payment over the network which does not track back to the user so that the user can safely purchase or make transactions on the internet without being traced back. Cryptocurrencies like bitcoin and eth can help users transact on the internet securely and in an unidentifiable way.

3. Users should know the correct knowledge on how to use the tools correctly and the user should avoid leaving footprints or traces on the internet when they are trying to be anonymous. Example a user should not post anything on social media from a logged-in account if they are trying to be anonymous on the internet as it reveals their identity.

c) Yes, anonymity in tor can be compromised if the user uses blockchain addresses for transactions which they have also revealed on any publicly available domain like Twitter or Reddit for normal transactions as then the user can be traced back to their publicly available media accounts which would reveal loads of information about the user.
If a user is not using tor DNS server than then all the user requests can be monitored at the DNS resolution as then the requests have to go through the DNS server which can read and log details about all the activity of a user.
A corrupted or malicious tor node can analyze user data and governments owning tor nodes can easily use the huge numbers of tor nodes to analyze traffic and deanonymize users.

Tor exit nodes are publically available, and hence if we have enough resources we can limit the bandwidth of these tor exit nodes by performing a DDoS attack at these exit nodes which may cause availability issues of websites as exit nodes will be busy.

Zero Day Exploit mAybe:
If we can get access to a malicious tor node we can encrypt all packets passing through that node and pass it on, which will cause the further nodes not able to access the inner layered packet, which can further cause issues in the tor network as all these packets with extra encrypted layer will never be opened as they are encrypted with the key of a malicious node.
Similarly, instead of encrypting the packet during the request phase, we can encrypt the response packet, which will cause the client unable to read the contents of the packets as the packets are encrypted with a key that is not available to the client.