FCS Mid - sem EXAM

Q1)

Q1 a)

k = 2

Customer ID	Names	Place	City	Country	No items purchased	price
C000*	*	*	*	*	2	6000.00
C000*	*	New York	New York	USA	2	3000.00
C000*	*	New York	New York	USA	3	5000.00
C000*	*	*	*	India	2	5000.00
C000*	*	*	*	*	1	10000.00
C000*	*	New York	New York	USA	3	5000.00
C000*	*	Brisban	Brisban	Australia	2	7000.00
C000*	*	Brisban	Brisban	Australia	1	7000.00
C000*	*	*	*	India	1	8000.00
C000*	*	Mumbai	Mumbai	India	1	7000.00
C000*	*	*	*	India	1	7000.00
C000*	*	Mumbai	Mumbai	India	2	7000.00

For k = 3

Customer ID	Names	Place	Country	No items purchased	price
C000*	*	*	*	2	6000.00
C000*	*	New York	USA	2	3000.00
C000*	*	New York	USA	3	5000.00
C000*	*	*	India	2	5000.00
C000*	*	*	*	1	10000.00
C000*	*	New York	USA	3	5000.00
C000*	*	*	*	2	7000.00
C000*	*	*	*	1	7000.00
C000*	*	*	India	1	8000.00

Customer ID	Names	Place	Country	No items purchased	price
C000*	*	*	India	1	7000.00
C000*	*	*	India	1	7000.00
C000*	*	*	India	2	7000.00

Q1 b) What techniques (at least two) would you do to increase the utility of the anonymized data?Demonstrate.

ANS)

Differential privacy

- Differential privacy (DP) is a system for publicly sharing information about a dataset by describing the patterns of groups within the dataset while withholding information about individuals in the dataset.
- Differentially private algorithms are used by some government agencies to publish demographic information or other statistical aggregates while ensuring confidentiality of survey responses, and by companies to collect information about user behavior while controlling what is visible even to internal analysts.
- Take a entry from a dataset
- Flip a coin
- If heads, return the answer
- If tails, then flip a second coin and return “yes” if heads and “no” if tails
- If we want to find the innocents from that above transform data.
- using conditional probability to find out probability from tranformed data.
- P("yes" | not innocent) = $p_{head} + (1 - p_{head}) * p_{head}$
- P("yes" | innocent) = $(1 - p_{head}) * p_{head}$
-

$$p_{innocent} = \frac{(P(\text{“}yes\text{”}) - P(\text{“}yes\text{”}|not\ innocent))}{(P(\text{“}yes\text{”}|innocent) - P(\text{“}yes\text{”}|not\ innocent))} = \frac{(1 - (P(\text{“}yes\text{”}) - (1 - p_{head}) * p_{head})}{p_{head}}$$

- $P(\text{“}yes\text{”}|not\ innocent) - P(\text{“}yes\text{”}|innocent) = p_{head}$
- Thus, when p_head is 0, the distribution of returned result is identical, no matter an individual is innocent or not .

K - anonymity

- The k -anonymity is used such that there is no relation between sensitive information and non senstive information.
- There is sensitive label and non -sensitive label.
- We will not do anything to senstive data because it is important.
- Now we have to make same row sucht that k rows have same data using 3 techniques
  - Suppression
  - Generalization

- Suppression : - Delete an entry from the row and column.
- Generalization :- replace the entry with less specific info.
  eg : - for k = 2

Age	Test Score
12	79
12	98
12	79
23	98
23	79
34	98

to

Age	Test Score
1*	79
1*	98
1*	79
2*	98
2*	79
3*	98
3*	52

Q1 c)

Fname	Dept	experience
Yash	CSE	Good learning while i was in iiit d
Yash	CB	I enjoy learning in the college
Manthan	ECE	Very hectic schedule
francis	ECE	overall great learning

- Now according to the GDPR law if you anonymise data such that there is no method to reverse the process and identify user directly or indirectly then gdpr law doesnt apply to that data ,and now that data can be used freely.
- So I have choose k -anonymity because in these i have control how much data to anonymise according to my utility of the data such that the information remain anonymous but data is still usable for business purposes.
- eg :- for k =2
- expirenece is sensitive data

Fname	Dept	experience

Fname	Dept	experience
Yash	C*	Good learning while i was in iiit d
Yash	C*	I enjoy learning in the college
*	ECE	Very hectic schedule
*	ECE	overall great learning

and for k = 3

Fname	Dept	experience
*	*	Good learning while i was in iiit d
*	*	I enjoy learning in the college
*	*	Very hectic schedule
*	*	overall great learning

all the detail become anonymous here i cannot utilise do anything with the data now , because there is no data to infer anything now.

Q2)

Q2 a) How does IPv6 address some of the limitations of IPv4 addressing. What stops wider adoption of IPv6 addressing scheme. Is IPv6 more secure than IPv4?

ANS)

- When ipv4 was designed ,the developers / researchers did not know the scale at which internet devices will grow such that in next decade or so, there will be depletion of the ip address.
- Ipv4 is of 32 bits and is able to give unique ip only to 2^{32} internet devices which is approx around 10^9 devices
- Now there are so many internet user, servers and also users have atleast 2 devices which use internet.
- Ipv4 cannot map all the device with unique ip address.
- And to avoid that ipv4 depletion , ISP uses NAT(network address translation).
- It maps private ip to public ips using NAT table.
- So in NAT we have to change source ip (private ip to public ip), to server can able to contact the NAT and NAT will forward the data to corresponding private ip from wwhich connection was initiaited.
- Now to solve one problem we introduced other problem.
- It is easy to change source ip in packet then any attacker will change source ip and can launch some attack .
- Also we cannot identify ip change in packet header , because all the packets has changed ip.
- Now to address this Internet Engineering task force devloped IPV6 to deal with this problem.
- Also introduced the protocol IPSec which encrypts the data above the network layer so that no attacker able to change ip or able to see the packets header above network layer.

- Now because of this attacker cannot change the ip because after establishing the secure ssl /tls / https tunnel the internet key exchange protocol will exchange the symmetric key to encrypt the data/pakcets at the network layer.
- This is how IPV6 address limtation of IPv4

What stops wider adoptation of IPv6 addressing scheme.

- The reasons is very simple, investment
- There are 10 of millions of router the company has to change to get benifits of IPv6.
- Because there is band-aid workaround exists for IPv4 (NAT), there is not compelling argument for the companies to invest in infrastructure to change all the devices and router .
- Also IPv6 is not backword compatible , so if devices is changed to native IPv6 then it cannot communicate with device having ipv4 protocol.

Is IPv6 more secure than IPv4?

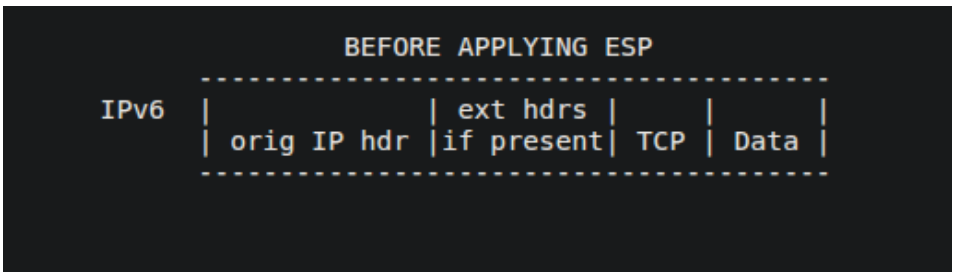
- As explained above the benifits of IPv6(+ Ipsec), many attack like packet anayisng , ip spoofing , is prevented, also maybe we can also eliminate ssl / tls third party authenication, as source ip cannot change then in the certificates we can verify using source ip only, hence also avoiding CA compromises.

Q2 b) What is IPSec? Can it be implemented over IPv4?

- Ipsec is IP security protcol which uses ESP, AH, IKE components that authenicate and encrypts data above network layer to form secure point to point tunnel.
- We will briefly discuss ESP, AH and IKE

ESP

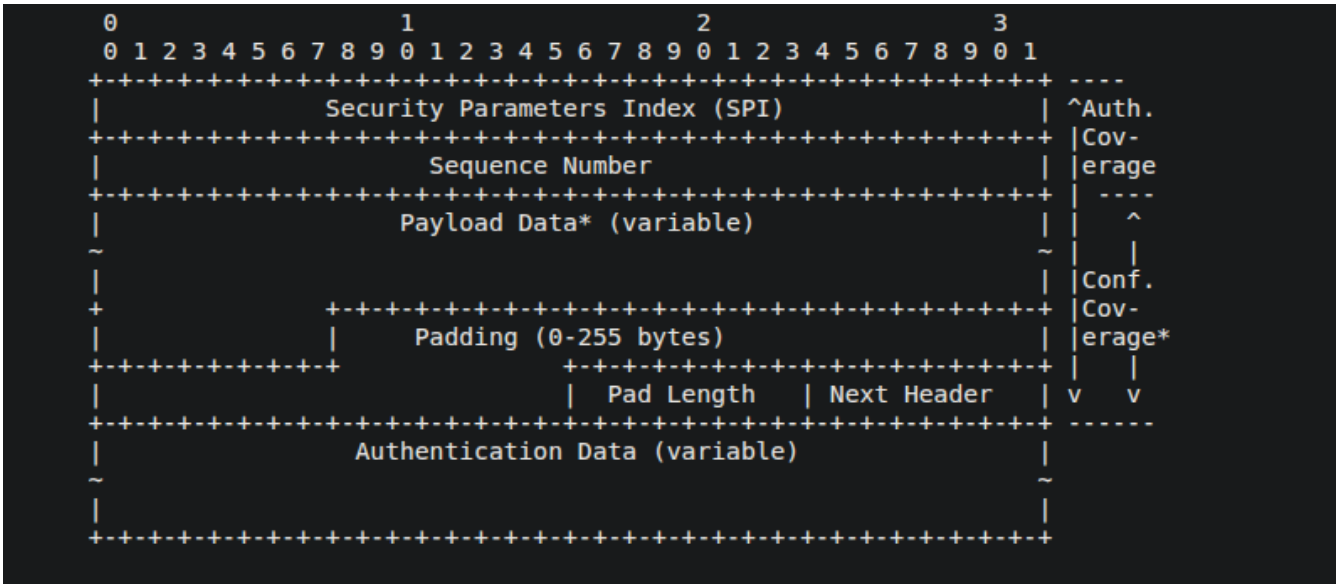
- The Encapsulating Security Payload (ESP) header is designed to provide a mix of security services in IPv4 and IPv6.
- It can be used with Ipv4 and Ipv6
- 

- 

```
                  AFTER APPLYING ESP
        ------------------------------------------------------
IPv6    | orig |hop-by-hop,dest*,|   |dest|   |     | ESP   | ESP|
        |IP hdr|routing,fragment.|ESP|opt*|TCP|Data|Trailer|Auth|
        ------------------------------------------------------
                                    |<---- encrypted ---->|
                                  |<---- authenticated ---->|

              * = if present, could be before ESP, after ESP, or both
```

- Its generally implemented at the default gate way where the ip source is gateway source hence any changes after that in ip header the packet will dropped after esp authentication fails.



AH

- AH protocol provides data origin authication, data integrity , using the symmetric key exchange by IKE.
- We will store the encypted header in AH field , such that it contains original Ip
- If there is change of ip in packet header then header checksum fails and we can drop the packet.

IKE

- This is protocol used to exchange symmetric keys after ssl/tsl/https secure connection.
- It is using RSA protocol for key exchange. since we have already established ssl connection.It is less likely MITM attack occurs.

Can it be implemented over IPv4

- Yes we can implement it over tunnel mode if ipsec

Q2 c) How would you defend the network, with limited overhead, against packet potential flooding attacks? Briefly justify the rationale

for your solution

ANS)

- To defend from the flooding attack we have to user firewall, because attack will not try to flood application like performing dos or ddos on website , but also can do ping attack or icmp attack at network layer.
- To prevent from this attack we must have stateful firewall.
- Because attacker may use ddos attack using proxy , now if the ip address is different then we cannot say if genuine ip or attacker proxy ip.
- To do that we need packet inspection, but since we have limited overhead we can use only store only some recent packets since firewall memory is llimited.
- Stateful packet inspection, also referred to as dynamic packet filtering
- Now stateful firewall able to detect most of the packets .
- Field in packet header is inspected.
- And because attacker can do SYN flooding attack , ping attacks , ddos on application level all those can be prevented.

Q2 d) How would you set up a secure communication channel for each device in the above architecture?

- Assuming we are using Ipv6 architecture in local network and all device support ipv6.
- We will be using SEND (Secure Neighbor Discovery Protocol) , send protocol is designed to estalisbhed secure communication where we cannot say for sure , that enviroment is safe , attacker will be not able to reacher this interface or part of the network.
- Send use NDP Neighbor Discovery Protocol which discovers local ipv6 nodes on local link and maintain active neighbour nodes.
- NDP is not secure because attacker can mitigate as device in local network.
- Send secures NDP by CGA (Cryptographically Generated Addresses) , it make sure that sender of NDP which contain ip address is the owner .
- We first generate public- private key pair for all of the nodes.
- CGA is formed by replacing the least-significant 64 bits of the 128-bit IPv6 address with the cryptographic hash of the address owner's public key.
- The Secure Neighbor Discovery (SEND) Protocol prevents an attacker who has access to the broadcast segment from abusing NDP or ARP to trick hosts into sending the attacker traffic destined for someone else, a technique known as ARP poisoning.
- The SEND protocol requires no public-key infrastructure.
- A new NDP option, the RSA Signature option, is used to protect all messages relating to Neighbor and Router discovery.
- Now for remaining non ip device like printer can be shared as service provided by the associated pc , can be accessed by ip address of the PC/device and port number.

Q2 (e) Turn on your phone's hotspot and connect your laptop to the network. Check your IP address. Is it Ipv4 or Ipv6? Share a screenshot of the IP address as well. Also, state the advantages anddisadvantages of the type of IP address you get

```
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

wlp0s20f3: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
        inet 192.168.57.195  netmask 255.255.255.0  broadcast 192.168.57.255
        inet6 fe80::b08:baa6:8f7:3da7  prefixlen 64  scopeid 0x20<link>
        inet6 2405:204:328c:9ddc:beea:6563:6521:b984  prefixlen 64  scopeid 0x0<global>
        inet6 2405:204:328c:9ddc:56f:d6e6:e2b8:352f  prefixlen 64  scopeid 0x0<global>
        ether 14:4f:8a:e0:b5:41  txqueuelen 1000  (Ethernet)
        RX packets 17649  bytes 9580440 (9.5 MB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 5885  bytes 2766789 (2.7 MB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
```

- public ip address :-

- I got both ipv6 and ipv4 address.
- important thing to note here is in ipv6 public ip is same as private ip, hence it can be accessible from anywhere , while ipv4 private cannot be accessible unless we do port forwarding.

ipv4

- Every internet device can has inbuilt ipv4 support , we can communicate with any ip device if it accessible and available.
- ipv4 can ensure security and privacy
- ipv4 has serious problem it is depleting because internet explosion , many device are now internet device , every ip device cannot mapped uniquely.
- due to this we have to use NAT but because of the we have to change source ip in the packet due to NAT, hence arp spoofing and ip spoofing attack can be possible.

ipv6

- ipv6 format able to mapped all the ip device easily because 128 bits.
- has better performance than ipv4.
- has better security using ipsec by encrypting the network layer and hence no one able to change ip in the header , so no arp/ip spoofing attack possible.
- IPV4 and IPV6 devices cannot communicate directly with each other, even in the most unusual of circumstances.

- IPV4 is still widely used & the world is slow to convert to IPV6

Q3 [Authentication]

Q3 a)

ANS)

- No it is not secure.
- Bob and alice have pre-shared key.
- Suppose Z is attacker , and it hijacks the connection .
- So if alice send "I am alice " to Bob but it will intercepted the message , delete it from the medium.
- Now Z send "I am Z" to Bob.
- Now send challenge R to Z .
- Now Z will send different challenge R2 to alice.
- Alice will reply R2 to attacker and send encypted challenge with encrypted key to Z.
- Now since Bob is stateless server it will send R2 challenge and encrypted challenge to Bob, now bob will decrypt and verify the challenge and it will authenicate Z .
- Now communication happening between alice to Z and Z to Bob , hence there is MITM attack is possible , so bob and alice cannot communicate securely.

Q3 b)

ANS)

- R could be anything , it could be simple message or hash function or public key,or digital cetificate , it can be anything
- I would like to implemeted simple string message to send to the alice.
- Since Bob is stateless protocol we would require to send R , encrypted R message , all Bob has to do decrypt , encrypted R and comapare with given r message if the key used to encrypt R is correct then the after decrypting with same pre-shared (bob does that) we wiil get original message , hence we can compare. if both value are equal then we can say Alice is verified. otherwise it is not.
- Hence we can say that both design of R doesnt have any impact on protocol.
- Design of R was otp , then also it same thing because bob is stateless server.

Q3 c)

ANS)

- Assuming that we are using asymmetric key for mutual authentication because of SSL/TLS
- Assuming that alice has already has public key of Bob.
- Now in this design R is certificate, so suppose Bob send his cerificate as R.
- And attacker Z has trying to MITM , Z is sending his certifcate , but because TLS/HTTPS or CA we will validate that Z is trying to intrusion between the communication.

- This is possible because the public key is already been shared to alice.
- Now suppose public key is not shared to alice before, the attacker shared his/her public key to alice masquarding as bob.
- Now alice send message to bob but intercepted by Z (attacker). trying to do same MITM attack described above.
- But now Bob is sending its digital cetificate , now if Z attacker send his own certificate and to alice ,alice send encrypted certificate of z , and now if z send encrypted certificate of from alice which is z cerficate , assuming the bob also verify its own certifcate using public key or CA . now MITM is not possible.

Q3 d)

- Yes we can if consider bob as server and alice as client and bob has digital certificate.
- suppose after generating public - private key pair bob ,send his digital certifcate which contains the public to alice , after alice had send first handshake message.
- Now using CA alice will verify bob's identiy/ valid server , now after that alice uses bob public key to share the symmetric key to bob and bob decodes that and after the alice and bob use that key to communicate .
- It is secure because we are using certificate to verify identity of the server .
- If Attacker try to spoofing or try MITM then he/she has send his own certificate which is not matching with public key of bob hence comm is secure.

Q4 Access Control

Q4 a)

ANS)

- Given transient users the ACL is more efficient to manage persistent protected resources
- Because if I want to remove certain user group from the ACL list , to revoke the permission to access those resources, then we can do in ACL , while there is no equivalent operation in capabilities.

Q4 b)

ANS)

- Bob has read , write and execute access of file z.
- Hence bob can write some trojan horse code to file z .
- Now if bob invokes alice via file z.
- If bob wrote trojan horse code such that it will copy the file of X (because Alice has read access to X) and write to Y (because ALice wrtie access to Y), so that bob can read what is written in file X .
- Because of permission is handle via ACL / capabilites it is difficult to manage the permission to process to get access to resources , if commited one mistake and attacker will try to use that to gain access to file to which process has not access to.

Q4 c)

ANS)

Bell-LaPadula security model

- It is a model in which permission is assigned to position or rank (used in military) a level of privilege to access/modify information and that position or rank is assigned to the person.
- There 4 level of access to information :
  - Unclassified
  - confidential
  - secret
  - top secret
- There is no read up policy formally defined as :-
  - A subject s has read access to an object iff the class of the subject C(s) is greater than or equal to the class of the object C(o).
- Means if the persion is a rank such that he/she can acesss only confidential information then he/she cannot access secret or top secret level of information.
- But person can read confidential information (same level) or unclassified level information (lower level).
- 2 nd property is no write down
  - While a subject has read access to object O, the subject can only write to object P if C(O) ≤ C (P)
- No process may write data to a lower level
- Why we are enforcing such thing is to prevent trojan horse , the process will not able to use higher level process to get access to confidential data and write to lower level so that lower level process (attacker) can read that.

Q5 Network Anonymity

Q5 a). Is perfect anonymity possible on Internet? Provide justification for either of your answers.

ANS)

- yes, using certain tools we can achieve lets discuss about them in details.
- So what takes to be anonymous , no one able to tell or traceback that the particular task is done by this internet user (means no accountabilty).
- Now for that how anyone (Internet service provider or government or vpn comapnies or website) able to track user.
- Website uses login information that is based on either soft authentication (containing email, etc) or hard authentication (aadhar card , pan card).
- So when we want to be perfect anonymous either we cannot access those service which we can access while login
- Also website tracks ip, so to be anonymous we have hide our ip, because isp knows our real identity associated with the ip.

- So for that we must use vpn to hide ip, but it really hides your ip, yes it does , but for only website and isp only, vpn company still knows who you are, because they are information about email id, or bank id or card id from which you did pay them for the service , so if user did something bad then the goverment or law agencies will force company to share user info.
- So from that we are not fully anonymous.
- Now we know vpn are not able to provide fully anonymity, but we get the idea what it need to be full anonymous (for now), hide your ip behind some proxy.
- There is one tool to achieve this , tor network which is open source.
- What tor does is make circuit for among different nodes , eg entry node, middle nodes , and exit nodes.
- So when we want to communicate with website , we use this circuit to communicate with the wesbsite.
- So when we are sending data to the website we will encrypt the data usign all the nodes public ip so no intermediates nodes can read the data and get the information about the user.
- So before sending we use encrypt in this order first encrypt with the exit nodes public , then intermediate nodes public and at last entry node or guard node public key and because of that no nodes except exit nodes able to get original message
- If we are using https/tls then exit node will not able to read message
- Hence it is calling the onion routing because like peeling the onion and at the core you find something meaningfull , here message is encrypted mutiple times only after decrypting n time you get original message.
- Now we talk about encrytion what about anonymity
- entry node only knows from where message is coming , doesnt know what is inside that.
- all the intermediate nodes doesnt know from where message is coming, and exit node also doesnt know and if it tls connection exit node also doesnt know from where message had come.
- and we are selecting nodes from tor diectory which has insane amount of nodes, hence we can say that using tor atleast our ip is anonymous.
- Till know we know that how to remain anonymous by avoiding
    - any authentication involving to accesss service which is soft authentication or hard authentication , that is able to traceback to your hard/physical identity.
    - and by hiding the ip.
- Is this still enough? No
- Suppose you want to do some shopping via certain website that allows user to buy thing anonymously, then if we pay to the wesbite if goverment or agencies want to track you , they will track flow of money and able to reach you using bank transaction.
- Now to eliminate certain tradition bank which have your hard identity, user can use cryptocurrency.
- But still goverment or law agencies will track flow of money via this cryptowallet able to track you.
- Hence In summary it is possible but you have to give up using these services , which is very difficult to do, but still you can be anonymous while avoiding these services.

Q5 b) What are the principles and technique that tools like TOR rely on to provide anonymity for users? Justify.

ANS)

- Tor network contains onion router which maintain tls connection with each other
- When the user want to use tor nodes as intermediate node then user gets the of onion router listed on the tor directory, first it will select exit node and then choose intermediate nodes and then guard node.
- While selecting the nodes no two router belongs two same organisation or selection two same router.
- So when we want to communicate with website , we use this circuit to communicate with the wesbsite.
- So before sending user will encrypt in this order first encrypt with the exit nodes public key, then intermediate nodes publickey and at last entry node or guard node public key and because of that no nodes except exit nodes able to get original message because only those can decrypt the encryption with the private key.
- If we are using https/tls then exit node will not able to read message
- Hence it is calling the onion routing because like peeling the onion and at the core you find something meaningfull .
- now tor principle while selecting the nodes is randomisation, because tor has many router spread across all the continent, so while selection router based on randomisation, it will be very diffiult to get two malicision nodes in same circuit, to do traffic analysis attack.
- Also tor ensures the guard should have some criteria and reputataion , some other parameters , to become guard nodes.
- Same for exit nodes , because these nodes are two nodes where there is high chances user identity might reveal, because guard nodes knows your ip and exit nodes your message if your are not using https/tls.
- So tor uses asymmetric key and probabiltity of malicious nodes.
- Hence we can say tor relies on randomisation principle annd ip masking.

Q5 c)

- One thing can be done , using google ads we can store some cookies . now suppose we are exit node and we are doing dns resolution we can redirect such that our own website to get the cookie and again redirect before user knows to their original website .
- This method works only if user doesnt clear his cookies regularly.
- And also we will try ddos to reduce the other tor router , or hammer connection so that , probabilty of selection our malicious nodes become higher.
- We can also packet analysis , we cannot add some pattern to the packets such that , other nodes (maybe ours) or someone who is ready for sharing data , while using 100 - 1000 s of nodes in different countries and hope for if your two nodes get selected, as guard nodes and exit nodes then you can do add some induced pattern while towards the other end .
- It will not work pratically in real environment because tor project notices this type of remove or ban them from directory. hopefully till then our task should be done.
- Rogue exit nodes if connection is not https get all the information about the user.
- Tor is slow and we can use that fact and try to induced time delay and observe traffic on the other side.
  we can get the pattern or correaltion .