What is considered to be an anomaly in the network caused by the covid events?

- Increased internet usage: With people staying at home and practicing social distancing, there has been a significant increase in internet usage worldwide. This is due to more people working from home, attending school online, and using video conferencing tools to stay in touch with friends and family.

- Increased demand for online services: As more people turn to the internet for work, education, and entertainment, there has been a corresponding increase in demand for online services such as video streaming, online gaming, and e-commerce. This has put a strain on the infrastructure supporting these services, leading to slower internet speeds and service disruptions in some areas.

- Cybersecurity risks: The COVID-19 pandemic has also created new cybersecurity risks, such as an increase in phishing attacks and ransomware attacks targeting remote workers and healthcare organizations. This has put additional pressure on internet security teams to ensure that their networks are protected against these threats.

What it illegitimate anomaly in the network that is caused by attack and how can we detect it?

The sites that will be in our scope are:
The main search engines such as google and bing
Worldwide popular news sites
Government sites of major states in the world (USA, Russia, China, Germany)

The metrics that we will measure will be:
The rtt of the ping to these sites, source and destination of the packets.

The route to the destination according to the traceroute data, and also the is_success and last_hop_errors to monitor if it was successful or encountered errors on the way. We can also see the end time and see how much time it took for the whole traceroute process, and although alone it does not tell a lot, we can still watch major variance in the times to monitor anomalies.
Also we are given the median rtt between each hop that can point to problems in certain links in the network.
*In all the packets in all the measurements, we can see the ttl field that indicate about the general distance the packet traveled before reaching, which give us some more information about the topology of the network.

DNS: we are given the response list which can show us the following:

- A server that was once mapped to this hostname is no longer received in the output, can indicate about anomaly in the specific server.
- The dns response time to indicate about how burdened the server is.
- The AS which can show us events like the pakistani BGP attack on youtube.

- The authoriatives servers that can indicate how many alternate servers can be used - reflecting how safe the site is from being completely taken down.
- Beside seeing only the mapping of the domain name to the servers ips, we can also see cname dns response that map to another domain name which can potentially be used in attacks or other events (although its very common all the time in the regular flow of the network as well)

NTP:
We can see from the ntp the target time offset, which can tell us when there is a large offset, that there is a network problem that is preventing the target's clock from synchronizing with the network time. This can be from either a technical problem, overload or an attack. Because the time offset from the ripe probe is not necessarily caused because of the server, we can use the precision and stratum (indicating about how close is it to the reference clock in the NTP hierarchy), to understand if the targets clock is the one thats reliable or not.

HTTP:
Contains all the common properties for all the measurement, will give us the same general information about the network - reachability, time etc.

SSL certificate:
Beside the general common information, by analyzing the SSl certificates we can detect the risk for man in the middle attack and spoofing, as this is the thing its responsible to protect against.

What we expect to see in case of each event:

- Normal effects that were caused by the covid - We expect to see a rise in the rtt's and problems with reachability that are caused by the spontaneous rise in the internet usage
- Denial Of Service attacks - Although we cant see the general rtt alone, as an indication for a DoS attack, as the usage of the internet during the covid pandemic was increased and with unusual patterns.

  By analyzing the median RTT between each hop in the traceroute, we can detect any anomalies in a specific hop that deviates from the expected behavior. Such irregularities may signify a problem in the link leading to the destination hop. Although the RTT measurements may contain noise and lack a reference baseline, a consistently high RTT observed solely at the endpoint, rather than throughout the entire route, suggests that the issue is concentrated within the final server and not caused by excessive network usage as a whole. Also, we can check according to the NTP as described in the technical part above.
- IP spoofing, Man in the middle and general impersonation:

  By using the SSL certificates we can check the authenticity of the servers. Because we are only checking historic data, we will measure the risk that was during the covid pandemic according to illegitimate/out- certificates

  We are not yet sure about how to use the DNS and http data w.r.t cybersecurity, and will further check how the information can help us when we'll start the technical analysis of the data.