Candidate Number: 2031B
Paper 2
Question 6

a. Performance:
   - If there is a large number of WiFi access points or a lot of users, then there will be a lot of data to log. If this is all processed by a central server or stored in a central database then the system could be under a lot of strain. This might become a performance bottleneck.
   - Furthermore a large quantity of data might take a long time to process, especially if the analysis is being done on hardware which is also used for other tasks. This might result in the analysis of the data being completed too late to act on the information
   - How long is the information stored for? If a historical record is required then the database might become very large.

   Security:
   - Where is the data being stored? Is it in a secure database. If it in the cloud then it might be vulnerable if it has a weak password. If it is on-site then it might be subject to physical tampering
   - How is the data being transmitted? Could somebody who has physical access to the access point (or indeed to any of the transmission infrastructure) tap in to the log data during transmission?
   - If somebody is connected to the WiFi, could they impersonate the access point and send fake log packets?
   - Could a malicious member of staff or a technician who installed the system get access to this data? If so they could delete, analyse, or modify it, which could lead to incorrect conclusions and/or violate the privacy of the WiFi users.

   Privacy:
   - WiFi users may object to the university logging user logons, as someone with access to this data may be able to track an individual user's movements.
   - Residents may object even more strongly to the logging of the WiFi access points in residence buildings than in labs/lecture halls
   - Even if user logons themselves are not logged, the network addresses could be cross-referenced with e.g., a user's lecture timetable to figure out which network address belongs to a device owned by an individual person.
   - Even if users are informed that the WiFi data is being logged, they might not understand/be able to tell exactly what information is being logged and to what extent it can be linked to them as an individual.

b.
   - Have the data be end-to-end encrypted between the WiFi access point to the analysis server.
   - Try to anonymise the data (e.g. try to just collect the volume of data rather than the individual user logons if possible)
   - Have the analysis be done automatically (i.e. not by a human)
   - Do not store the raw data, only store the results of the analysis
   - Use a cloud computing platform such as GCP or AWS to do the analysis to avoid physical tampering/performance bottlenecking
   - Use a secure password for databases/cloud services

- Send a clear mass email to everyone who might use the WiFi to tell them exactly what data is being collected and why. This should be accompanied by clear signage in areas with WiFi
- Restrict who can see the results of the analysis to only high-level trusted staff just in case a user's private data can still be retrieved from it.

c. Instead of monitoring the WiFi usage, instead install infra-red sensors on both sides of doors to lecture halls, labs, and residence buildings. These would consist of a beam of infra-red light stretching across the door to a sensor on the other side. It can be detected if the beam is broken. If someone enters the room, the beam on the outside of the door will break first, then the inside. The beams will be broken in the opposite order if someone leaves the room. This will be able to give a rough estimate of how many people are in the room at a given time. This data could be logged and sent to a central server for analysis.

   For rooms with only a single door, the entire counting analysis can be done within the sensor device, rather than on a central server. (Likewise if a room has multiple doors which are close together, then the sensors could communicate through Bluetooth to count the people in the room. Either way the counting is done on dedicated hardware rather than on a central server). This will improve performance and does not require sharing hardware which was already being used for other tasks.

   Some issues with this approach are if multiple people enter/leave at once, or if someone somehow avoids breaking the beam. This method would also be subject to physical tampering of the devices, just like WiFi monitoring.

   However, a benefit of this system is that the only data which is stored or transmitted are the counts, which is much more anonymous and less privacy-violating than monitoring people's network addresses/user logons.