Security and Privacy in IoT (COMP7048)

Table of Contents

1.0 Introduction

IoT Innovate was approached by a client to design an IoT system architecture for a cashier-less store located in the UK like 'Amazon Go' stores in the USA. We will be designing the architecture of the entire system including the technology, techniques, and network protocols that we will implement into the store. In addition, we will also be highlighting the needed functionalities that a cashier less store like this will have as well as considering the potential security and privacy risks that could occur with a store with this much technology connected to the internet.

2.0 Functional requirements of the respective IoT infrastructure

The following are functional requirements needed to be implemented for this store to work.

The first functional requirement needed for a cashier less store like this is a way to link a customer to an account that stores their payment details. Customers should be able to walk straight in and straight out when they're done. Once a customer has finished shopping, their account will be automatically billed for the items that they have picked up and walked out of the store with.

A second functional requirement is the ability for items to be automatically added and removed from a customer's virtual shopping card. One of the reasons why a store like this is popular is that there is no need to line up and get the items scanned, which is why it's important that a system like this can tell the difference between items and can detect when a customer has picked up the item and when they've returned it. Once the customer has left the store, all this information will be sent and stored on the customer's history on their account.

Another functional requirement is the ability to track and distinguish between customers. An IoT infrastructure like this would only be able to tell what customer is picking up or returning what items but recording and using that data in machine learning algorithms to return a high level of accuracy.

The next function would be the ability for cameras to detect what item has been picked up. This would be done by all the cameras mentioned in the previous requirement as well as cameras stored on shelfs and weight sensors. All the data collected from these data sources will be fed into an AI to determine the item.

Another functional requirement that is needed is knowing when shelves need to be restocked. We would want a technical and efficient way of determining this so there would be no need for any employees to be walking around the store checking stock.

3.0 Technical security and privacy requirements

3.1 Security

A security issue to consider is keeping any data relating to the facial recognition aspect of this store secure. It is becoming increasingly easily to access IoT devices if they fail to have adequate security standards using IoT browsers such as Shodan which makes searching for devices connected to the internet like our cameras for example easily easy to spy; In the United Kingdom there are roughly 14 million ports open on IoT connected devices that Shodan has indexed (Shodan Internet Exposure Dashboard, n.d.). Therefore, having either a complex username and password rule or convention for anyone who has access to the servers where these cameras are streaming to and having access controls to restrict anyone without authorisation accessing.

Due to the nature of a store like this with the increased usage of sensitive data such as payment details are now being done over smart devices and being stored in cloud-based software platforms it increases the attack surface allowing hackers to have a field day of data that they can try to steal. Which is why it's vital to have adequate data protection protocols in place such as point-to-point encryption (P2PE) which means all the transactional data that flows through this store is fully encrypted from the time customers account is automatically charged when they leave the store to where that payment is received such as a business account at a local bank (Fernando, 2021).

To access any QR code that is generated, the customer first needs to login into their account and provide the correct multi-factor authentication (MFA) code using the method they selected when creating their account. We will put into practice a complex password policy that ensures that customer's passwords aren't easily guessable or can't easily be cracked using either a brute force or dictionary attack. The password must be more than 8 characters long including both upper-case and lower-case characters, symbols, and numbers. It is recommended that a customer has a longer password as a password that meets the minimal requirements of this policy would only take eight hours to crack (A Secure Password Is one of the Biggest Steps In Protecting Your Privacy, 2020). However, we understand that everyone who uses this system might not utilise a password manager, hopefully eight hours is long enough to discourage any malicious threat actors. In addition, we can combat brute force or dictionary attacks by implementing a time out feature where if a customer gets their password wrong more than 3 times then their account will be locked.

Our backend database will use structured query language (SQL) to help manipulate the data if needed to provide easy data management. However, this introduces additional risks including SQL injection (SQLi). Malicious threat actors could potentially expose confidential data by passing a string that includes either a wildcard or a statement that our database will register as true. A statement like this `'' UNION SELECT username, password FROM users--'` This statement if passed might cause our application to return all usernames and passwords if the user input is not properly sanitised.

3.2 Privacy

One of the alarming things about a store like is the use of facial recognition and AI which is why it's important that the data controllers and processors consider the GDPR and make sure that the data collected in the store will only be used for relevant purchases and kept secure. If any data was unlawfully and unethically processed such as selling facial recognition and names to a 3rd party specialising in facial recognition software then the cashier less store take responsibility and pay the consequences (Guide to the General Data Protection Regulation, 2018).

We should also consider the Data Protection Act 1998 (DPA) in which the Information Commissioner's Office (ICO) issued their code that covered the use of CCTV cameras. Following the recommendations outlined by the ICO will help ensure that no one is breaking any laws which in itself reduces risks to reputations and avoid regulatory action and potential penalties as well as ensuring that the deployment of our cameras are efficient and the data we collect can be used to their correct purpose and help re-assure the data subjects (In the picture: A data protection code of practice for surveillance cameras and personal information, 2017).

The data that we will be processing could be considered Personally Identifiable Information (PII) so it's important that we take the necessary steps to ensure we aren't in breach of any data protection regulations and if needed outsource to look for a solution from any lawyers with expertise in data protection compliances.

It's vital to keep any data transferring over the network secure and encrypted. This would mean implementing necessary network protocols such as Hypertext Transfer Protocol Secure (HTTPs) so in the case that a threat actor was to implement a man-in-the-middle attack and were sniffing on our network they would not be able to breach confidential information. This would ensure as well one of our privacy requirements with General Data Protection Regulation (GDPR) compliance.

To get a detailed and compressive idea on risks to privacy for this system we were able to identify some of the top frequented privacy threats that OWASP has recorded. Below is a table describing the OWASP risk, threats and remediations (OWASP Top 10 Privacy Risks | OWASP Foundation, n.d.).

OWASP Risk	**Threats**	**Remedies**
Web Application Vulnerabilities	- Broken access control - Server-side injections - Server-side request forgery	Penetration Tests to manage attack surface and comprehensive security designs
Operator-sided Data Leakage	- Inadequate access management - Inadequate data protection	Encrypted data storages and implement policies
Insufficient Data Breach Response	- No procedures in place - No breach transparency	Policies in place in the case of breach
Consent on Everything	- Inappropriate use of content	Terms and services for data collection

Non-transparent Policies, Terms and Conditions	- No information on how data is processed - Information to technical	Ensure that there is clear and concise information on how data is used and understandable by non-lawyers.
Insufficient Deletion of User Data	- Failing to delete personal data upon request effectively and/or timely manner	Data deletion plans
Insufficient Data Quality	- Using outdated, incorrect, or bogus data. - Failure to update to correct data.	Implementing the mandatory data confirmation after a period when the account was created
Missing or Insufficient Session Expiration	- No session termination - collection of additional user-data without the user's consent or awareness	Implementing session timeouts after a certain time of the account being logged in
Inability of Users to Access and Modify Data	- No ability to access, change or delete data	Implementing secure and well design update mechanisms
Collection of Data Not Required for the User-Consented Purpose	- Collection of any data that is not necessary	Clear and concise policies that define how and why we collect what data

OWASP's top privacy threat was web application vulnerabilities, the threats for this could include broken access control, remote injections, and server-side request forgeries so to combat this we will ensure that our security designs are properly searched as well as hiring penetration testers so we can manage the attack surface.

4.0 Associated and unique challenges

The first challenge that a system like this will run into is when it comes to identifying a customer on the shop floor. Plenty of cameras will be used alongside some use of recognition such as facial or body types. This was an original challenge that Amazon themselves faced that delayed the launch "due to various technical difficulties, such as getting the system to watch more than 20 people at once or distinguish between customers of similar body types." (Xle, 2018). We also need to consider calibration of these cameras; it's great to have all these cameras but they need to be able to identify where they're in the store.

No matter how advanced this store will be or how much IoT technology we implement we still face the issue of restocking shelves and cooking food; this is combated by the least technological solution possible, human workers. We will need actual people on the floor cooking and packaging the food and restocking shelves when needed. We may also need human intervention to check IDs when someone tries to purchase alcohol or the potential use of age recognition which begs the question that could it get it wrong?

Another challenge that will occur is knowing what item which and this is done by all the camera and sensor technology implemented. This is fine for items that have already been sealed and package such as condiments because they will be easily recognisable by cameras or have their barcodes read but the challenge occurs for items that look similar such as a shop's pre-made food that sell in offers such as meal deals. We will be taking the same approach as Amazon and their Go stores where they produce unique and camera-friendly barcodes.

Scalability is a massive challenge that we might face depending on how popular this cashier-less store will be. The cameras will be scanning customer's and we may need to implement more cameras to make sure that everything works efficiently. The use of all this data and these algorithms may present a storage problem so cloud-based storage will be the way to go.

Due to the nature of this store and because it's all IoT based with loads of nodes and data flowing, it means that the attack surface is exponential meaning that there are many areas that a malicious threat actor could attack. We will need to implement a security plan so comprehensive that it covers any potential attack on any potential IoT enabled device or node.

5.0 State-of-art and proposed technologies to be used

5.1 Custom App

The first state-of-art proposed technology that we will need is a custom app that will store customer's payment details and unique QR codes that will need to be scanned when entering the store. This QR code will be uniquely generated every time to combat the potential identity theft if these QR codes were static and not dynamic and were stolen or a customer's phone was stolen. This app will be featured on both app stores for both android and apple for usability purposes.

Figure 1: Amazon Go App. - (Bishop, 2018)

5.2 Cameras

The second proposed technology that we will implement is a large quantity of cameras all around the store on the roof as well as individual smaller cameras on each shelf. These cameras will be high quality and cover every part of the store, so we are able to stream the data being recovered to servers running our facial recognition software and AI. These cameras will also be able to recognise what items customers are picking up or putting down. These cameras are where the main privacy concerns come into play so it's important that we collect and process all the data without being in any breach of any regulation like the GDPR or the DPA. We will calibrate these cameras to specific zones in the store so that we are aware of where the footage is being captured.

Figure 2: Amazon Go Cameras. - (Bishop, 2018)

5.3 Camera Friendly Barcodes

The aforementioned cameras that detect what customers pick up or put down are able to do that easily enough with easily recognisable items and labels but the main issue with this is when a camera tries to recognise an item that looks extremely similar to other items such as a sandwich that is sealed in a plain white bag with only a name on it; this is where our camera friendly barcodes come in which allows us to seal these types of items in bags that features unique barcodes that can be easily read by any camera in the store.

Figure 3: Amazon Go Barcode. - (Amazon Go, Getting There with Better Retail! | Retailmatics Site, 2018)

5.4 Entry and Exit Sensors

When customers walk into the store, they will need to scan a unique barcode with a set of special sensors that will read the barcode and identify who they're by their account information which will be fed to the cameras. The customers will also need to exit via another sensor that will detect when a customer has left the store allowing this system to start the automation process for payment.

Figure 4: Amazon Go Entrance. - (Bishop, 2018)

5.5 Weight Sensors

We will implement small sensors on the individual shelfs that will be able to tell what item has been picked up or returned due to the shift in weight. We will also add small cameras which will help identify what customer has picked up an item using facial recognition. Both data sources will be able to identify what item has been either picked up or returned and by which customer.


Figure 5: Amazon Go Shelfs. - (Coldeway, 2018)

6.0 Overall security and privacy architecture

Customers must create an account before anything else. They will use an email address and password combination alongside MFA to ensure that the correct user is accessing the account. When customers attempt to login to their accounts and they're successful before they are authenticated another authentication mechanism is used. One of the three basic authentication factors 'something you have', in this case, the customer will receive a text message to a number they used when making the account and will be requested to put this number in before they're authenticated (What Is Two-Factor Authentication (2FA)?, 2020).

When a customer first enters the store, they will be met with our sensors. Customers will use their smart phones to scan a QR code that is uniquely generated every time. The unique QR code signature will also contain information related to a customer. This data is sent over encrypted HTTPs network traffic to be cross referenced with our backend database containing customer information, 128-bit symmetric encryption will be used to for the above encrypted network traffic to ensure that the data is secure. This database will manage by a database management system (DBMS) that will implement access control mechanisms that protect the confidentiality, integrity, and availability of our customer's data (Bertino, E. Ghinita, G. and Kamra, A. (2011). The database and the management software will be password protected adhering to a strict password policy ensuring that all passwords on this system are over 8 characters long, including both upper-case and lower-case characters as well as numbers and symbols. However, the big flaw in this authorisation method is that we have no way to authenticate whether the correct person is using the smartphone that is displaying the QR code.

Each node in the network will be running the LEACH protocol to ensuring energy efficient and constant availability. The IoT devices will be split into clusters and the nodes will work as cluster-head (CH). The energy consumption is calculated using this formula; $Etx\ (d) = kd\ 2\ [nj/bit]$, $where\ k = 1$ (Singh et al. 2017).

The cameras that are implemented around the store will be recording and streaming all footage to the backend on our server location over encrypted traffic. The cameras themselves will also be protected with complex passwords adhering to the previously mentioned policy to ensure that there are no malicious threats actors that have gained access and are able to watch the footage. As you can see from figure 6 which shows the results from only a single tag “server: boa WWW-Authenticate: Camera” showing that there are over 20,000 IoT camera-based devices that people are able to access hence the need for the complex password and encryption. This is also important as the footage recorded could also be classified as PII so to comply with GDPR compliance we need to ensure that we have implemented adequate protection and that no footage is kept for longer than necessary. To further compliance with the GDPR, in the worst-case scenarios where we have fallen victim to data breaches, we need to ensure that we have comprehensive policies in place to inform customers as well as the needed regulatory boards and have policies in place to remedy and contain the breach.

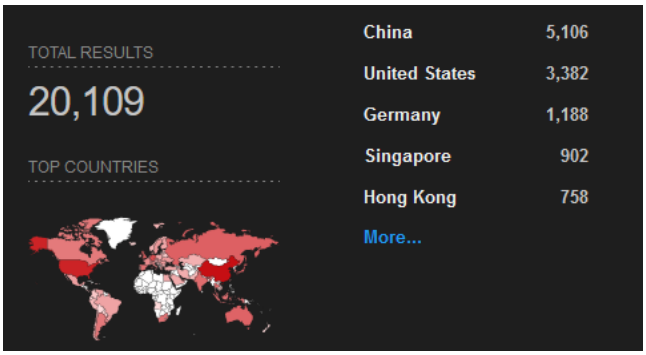

Figure 6: Shodan.io. - (Shodan Internet Exposure Dashboard, n.d.)

All this data is used to solely determine what customer has picked up and kept what item. When determined, the information is then stored in a customer’s virtual basket associated with their account and when a customer finish shopping and leaves the store the data is then processed from our server and then fed into the theoretical environment of our financial services which would be the location of the cashier-less store’s banking organisation. This environment provides the point-to-point encryption protocol that automatically encrypts customer’s payment methods and then decrypts at the bank.

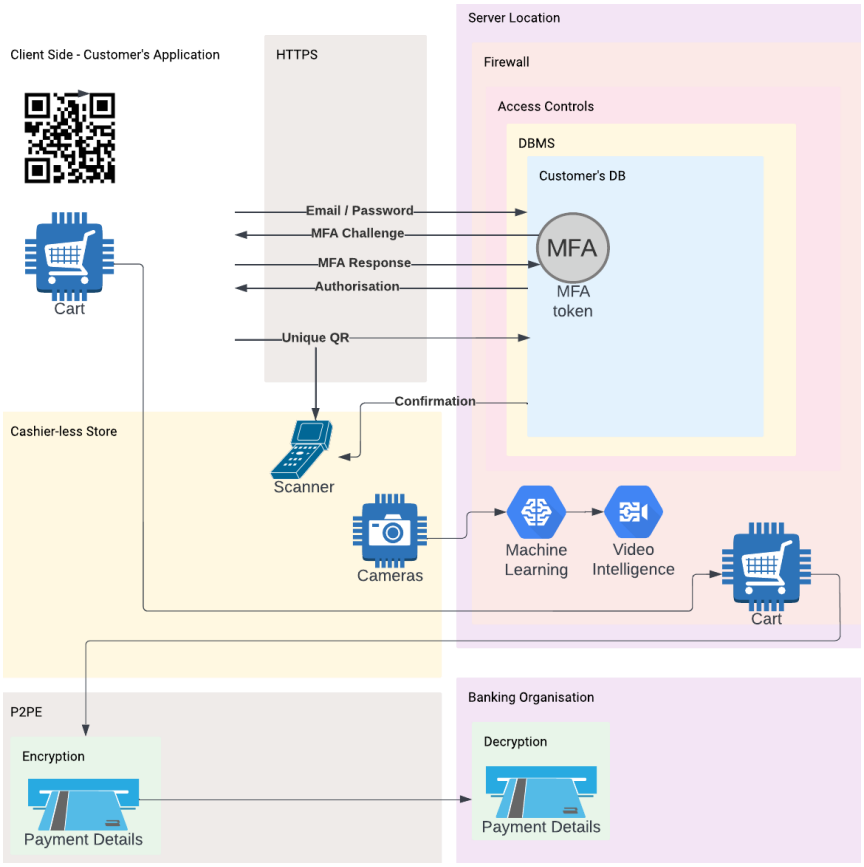

Figure 7: In depth Look - (Personal)

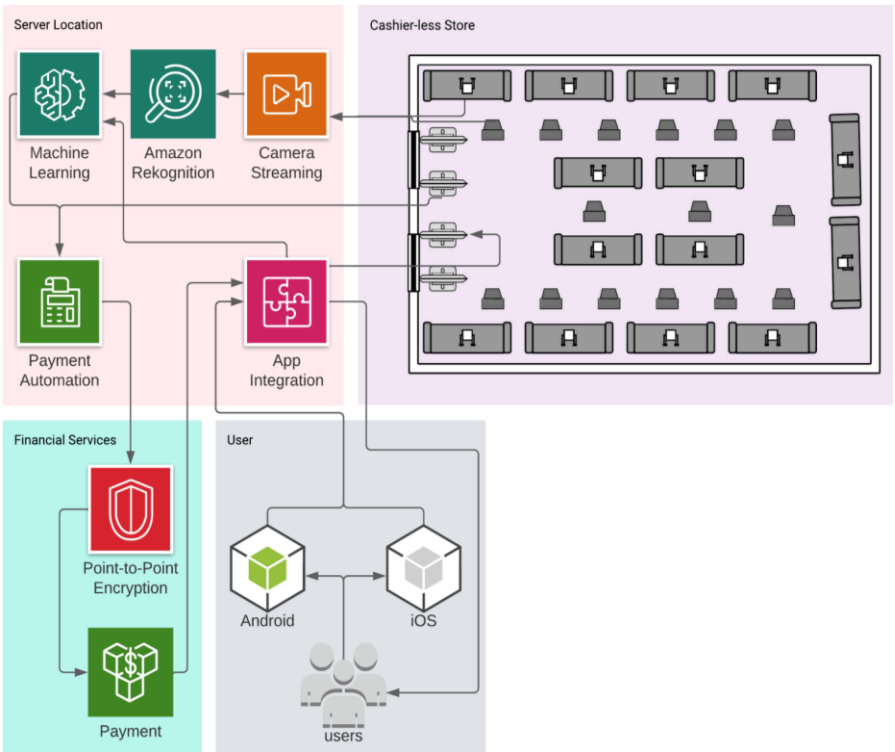

Figure 8: System Data Flow. - (Personal)

7.0 Conclusion

We have discussed the functional requirements for a cashier-less store to function as well as going into the necessary technology needing to be implemented to meet these requirements. We have considered the potential security and privacy risks and have offered recommendations to ensure that this store will comply with GDPR guidelines. We have also touched upon the security requirements needing to be added to the technology to ensure that it is safe from any malicious attacker. In addition, we have also considered the security measurements for privacy such as encryption and P2PE for payment information.

Bibliography

Fernando, J., 2021. *Point-to-Point Encryption (P2PE)*. [online] Investopedia. Available at: <https://www.investopedia.com/terms/p/pointtopoint-encryption-p2pe.asp> [Accessed 7 February 2022].

Exposure.shodan.io. n.d. *Shodan Internet Exposure Dashboard*. [online] Available at: <https://exposure.shodan.io/#/UK> [Accessed 3 February 2022].

GOV.UK. 2018. *Guide to the General Data Protection Regulation*. [online] Available at: <https://www.gov.uk/government/publications/guide-to-the-general-data-protection-regulation> [Accessed 3 February 2022].

Xle, J., 2018. *Amazon plans to open more cashierless grocery stores this year*. [online] Curbed. Available at: <https://archive.curbed.com/2018/1/22/16919268/amazon-go-first-store-seattle-open> [Accessed 7 February 2022].

Bishop, T., 2018. *Amazon Go is finally a go: Sensor-infused store opens to the public Monday, with no checkout lines*. [online] GeekWire. Available at: <https://www.geekwire.com/2018/check-no-checkout-amazon-go-automated-retail-store-will-finally-open-public-monday/> [Accessed 3 February 2022].

Retailmatics. 2018. *Amazon Go, Getting There with Better Retail! | Retailmatics Site*. [online] Available at: <https://retailmatics.com/amazon-getting-better-retail> [Accessed 3 February 2022].

Coldeway, D., 2018. *Inside Amazon's surveillance-powered, no-checkout convenience store*. [online] Techcrunch.com. Available at: <https://techcrunch.com/2018/01/21/inside-amazons-surveillance-powered-no-checkout-convenience-store/> [Accessed 3 February 2022].

2017. *In the picture: A data protection code of practice for surveillance cameras and personal information*. [ebook] Information Commissioner's Office, p.6. Available at: <https://ico.org.uk/media/1043340/surveillance-by-consent-cctv-code-update-2015-jonathan-bamford-20150127.pdf> [Accessed 15 February 2022].

Owasp.org. n.d. *OWASP Top 10 Privacy Risks | OWASP Foundation*. [online] Available at: <https://owasp.org/www-project-top-10-privacy-risks/> [Accessed 11 May 2022].

What Is Two-Factor Authentication (2FA)?. 2020. *What Is Two-Factor Authentication (2FA)?*. [online] Available at: <https://www.avg.com/en/signal/what-is-two-factor-authentication> [Accessed 11 May 2022].

Elisa Bertino, Gabriel Ghinita and Ashish Kamra (2011), "Access Control for Databases: Concepts and Systems", Foundations and Trends® in Databases: Vol. 3: No. 1–2, pp 1-148. http://dx.doi.org/10.1561/1900000014

IS&T Blog. 2020. *A Secure Password Is one of the Biggest Steps In Protecting Your Privacy*. [online] Available at: <https://www.is-t.net/blog/importance-of-secure-passwords/> [Accessed 11 May 2022].

S. K. Singh, P. Kumar and J. P. Singh, "A Survey on Successors of LEACH Protocol," in IEEE Access, vol. 5, pp. 4298-4328, 2017, doi: 10.1109/ACCESS.2017.2666082.