University of Glasgow | School of Computing Science

A secure client-server mobile chat application implementing an elliptic curve integrated encryption system (ECIES) and other security features.

Daniel Furnivall

School of Computing Science
Sir Alwyn Williams Building
University of Glasgow
G12 8QQ

A dissertation presented in part fulfilment of the requirements of the Degree of Master of Science at The University of Glasgow

1st April 2022

Abstract

abstract goes here

Education Use Consent

I hereby give my permission for this project to be shown to other University of Glasgow students and to be distributed in an electronic format. **Please note that you are under no obligation to sign this declaration, but doing so would help future students.**

Name: _____ Signature: _____

Acknowledgements

I would like to thank my supervisor, Mark McGill for consistently providing suggestions that made me question my own thought process and adjust my approach for the better. The constant support throughout the project is something I am incredibly grateful for.

I would also like to say thanks to Ewa Wanat, Tzvi Lipschitz, Thaïs Ramdani and Matt Weston for deeply helpful stylistic guidance, Bryce Campbell for top tier music recommendations, Conal Brosnan for Kotlin evangelism, Karim Al Tom for Lebanese sweets. Additionally, I'm extremely grateful to Emer Sweeney, Stuart Allan and Michael Callan for their much-needed assistance.

Lastly, I would like to thank my wife Rona and my furry friend Gordon for providing much needed moral support throughout the development and writeup process.

Contents

Chapter 1: Introduction

In an increasingly digital world, we are constantly producing data (and, of course, corresponding metadata). It's estimated that humanity created somewhere in the order of 2.5 exabytes of digital data per day in 2018[3] and the growth of digital data creation by individuals is constantly accelerating.[3]

As storage costs decrease over time[39], economic and political incentives have begun to develop for nation state actors and major organisations to develop profiling systems using large-scale data collection and mining. These "Big Data" profiling systems are already being used for targeted advertising[9], market segmentation[30], criminal investigations[40] and sentencing[35]. In the political sphere, these models have been used for increasingly effective traditional campaigning, [20] as well as (alleged) mass psychological manipulation[4] and disinformation[38] campaigns.

The question of whether the average individual is able to enjoy the fruits of such data collection is less clear. Do these large actors have the best interests of the subjects of their data in mind? If not, methods for obfuscating or hiding sensitive data become valuable considerations.

In a pre-digital world, an individual could avoid eavesdropping or data collection by malicious entities by simply speaking in a hushed voice, shredding documents, moving quietly or looking over their shoulder. In the current landscape, it is much more difficult to avoid surveillance without eschewing technology entirely. Our mobile phones are constantly communicating our triangulated locations, and even if our web traffic is encrypted by default, browser or ISP metadata still contains useful profiling information.

Secure messaging applications aim to allow individuals to communicate with other individuals around the globe while avoiding the potential for eavesdropping. In theory, this means the user can enjoy the benefits of a globally connected world while preserving their own privacy. However, in practice there are of course many implementation challenges to be considered.

1.1 Why are secure chat applications needed?

Secure messaging applications provide a means of communication between individuals or groups across a network of some kind. This can take the form of text, audio or video messaging, document or file sharing.

To begin to understand the users of secure chat applications, two important questions need to be answered:

1. Who desires secure messaging?
2. Who are they aiming to protect their data from?

There are many categories of potential users of such applications, and from wildly different settings. These can be mundane and innocuous, such as the organisation of a surprise party for a friend or family member.

However, another group of potential users are those who seek to hide criminal behaviour from law enforcement organisations. A high-profile example of this would be the EncroChat network of encrypted phones, predominantly used by organised crime, which was unveiled after a Europe-wide infiltration and investigation of the network by law enforcement groups (leading to several thousand arrests) [36].

Another group who may wish to evade police or law enforcement are political dissidents. In what has become known as the ”Million Dollar Dissident” [24] case, a dissident in the UAE, Ahmed Mansoor, was targeted by the NSO group (an Israeli cybersecurity firm which produces spyware for government use). He subsequently had his passport confiscated, he was beaten multiple times, his car was stolen and finally he was imprisoned by the UAE authorities - all within a week of posting anti-government posts online [29]. There are many parallels here with other groups who benefit from secure messaging - whistleblowers sharing information with journalists, and police informants who need to share data secretly with law enforcement.

In reality, there are plenty of reasons for *everyone* to use secure messaging such as protecting data in case of device theft, avoiding embarrassment, or limiting exposure to blackmail or government surveillance. Continuing the comparison to real-world communication - when speaking out loud, people don't tend to shout all the time, and tend to consciously limit and monitor who is listening to a conversation.

It should be clear that although there are many categories of potential secure messaging users, there are also many potential adversaries for these kinds of platforms, which means the development of such services is a complex undertaking. There are major tradeoffs which need to be made between usability and data security - for example, how can we preserve security of messages while also storing them on a mobile device?

1.2 Objectives

The primary development objective of this project was to develop an Android mobile application (and corresponding server) that uses complex security features including an end-to-end encryption solution that uses both asymmetric (Elliptic Curve Diffie-Hellman) and symmetric (AES) encryption approaches, pseudonymous identity and self-destructing messages.

During the development journey of this application, the intrinsic motivation was to come to a greater understanding of the complexities, assumptions and tradeoffs involved in creating these kinds of platforms.

Chapter 2: Requirements and Analysis

Unger et al's "Systematization of Knowledge" (SoK) 2015 paper on secure messaging captured in extensive detail some of the key considerations for developing secure messaging applications. They highlighted three points which comprise the major challenges in this field - trust establishment, conversation security and transport privacy.

Trust establishment describes the problem of ensuring that users are communicating with the party they intended. This is a difficult problem, as there are tradeoffs between security and usability that need to be made. One method which has seen extensive use is key-sharing. This ensures that the recipient of a message is able to verify that it was sent by the sender.

Conversation security in this context refers to the security protocol used to encrypt the data. Unger et al lament the fact that most secure messaging solutions use only static asymmetric encryption (i.e. long-term key persistence meaning users do not need to exchange keys regularly). This is another example of the usability/security tradeoffs that need to be made for wider adoption of secure messaging.

The final challenge is transport privacy, which concerns how messages are transmitted between users. This is a complex problem, as metadata such as the identity of the sender and recipient need to exist so that the server can route messages correctly, despite also being appealing attack vectors for malicious actors.

2.1 Existing applications in this field

There are many existing secure chat applications available for both iOS and Android, with wild variation in terms of encryption architecture, as well as other security and usability features. For brevity, this brief review focuses on the three largest mobile secure messaging providers in 2022 - Signal, Telegram and Whatsapp.

The table below highlights some of the features and attributes of each of these applications:

Application	Signal	Whatsapp	Telegram
Ownership	Signal Foundation (non-profit, USA)	Facebook Corporation (USA)	Durov Brothers (Russian)
End-to-end Encryption	Double Ratchet Algorithm[1], constantly cycling symmetric session keys and Elliptic Curve Diffie-Hellman (ECDH) keypairs - this approach is called the "Signal Protocol"	Proprietary implementation of the Signal Protocol (unverifiable). Does not apply to backed-up messages stored on Google servers.	Only available in "Secret Chats", and only on mobile devices - uses Telegram's own MTProto 2.0 protocol[18]. For group chats, messages are encrypted symmetrically and are theoretically readable by the server.
Open Source	Fully open source for both client and server (with the exception of a server-side anti-spam component)	Entire closed source	Open source client with a closed source server.
Self-destructing messages	Controlled by the sender (i.e. the messages will be deleted from the recipients device based on the sender's settings). Manageable via defaults or for individual conversations. Timespan variable.	Sender-controlled, can be managed individually or via defaults. Fixed timeframe of 7 days.	Sender-controlled, can be managed individually or via defaults. Timespan variable.
Group chats	Fully E2EE encrypted using "client-side fanout"[19], where each message is encrypted individually to all users in the group.	Same as Signal (unverifiable due to closed source)	Encrypted between client and server but not E2EE.
Pseudonymous Messaging	Not available (linked to phone number)	Not available (linked to phone number)	Not available (linked to phone number, although it is possible to hide phone number from contacts after creation.)
Metadata collection	Only stores phone number used to register, date of initial registration and date of last app use.	Historical message metadata, phone contacts, device metadata and activity, blocked numbers, read receipts, full name and more.	IP addresses, phone contacts, historical message metadata.

2.2 Issues in existing applications

As can be seen in the table, Signal appears to be the most secure of the available secure messaging applications. This comes from a combination of best-in-class encryption, open source code allowing for security audits, fine-grained control of privacy options and innovative handling of security such as their use of client-side fanout for group chats. Whatsapp and Telegram both closely guard some or all of their source code, making it impossible for independent researchers to verify that their security claims are accurate. Signal also seems to have limited financial or political incentive to store metadata about the behaviour of their users due to their non-profit ownership model, which contrasts with Whatsapp's ownership by an advertising company and Telegram's use of advertising via sponsored messages.

All three applications provide fairly strong capabilities for self-destructing messages. Importantly, all three allow the sender to define the parameters for deletion, although Whatsapp does not provide timeframe granularity compared to Signal and Telegram. This means that users can feel relatively secure in the knowledge that even if the recipient is compromised their own exposure is limited.

One of the major downsides of all three applications is the lack of pseudonymous messaging. This is an understandable omission due to implicit need for chat applications to incorporate some means of social network discovery (in all cases, phone numbers and contact lists are used for this function).

This means that the development of an application that provides comprehensive end-to-end encryption and self-destructing messages alongside pseudonymous messaging features is worth pursuing.

2.3 User Personas

Thaïs - a 37-year-old police informant deeply nested within a major international drug distribution network. She needs to communicate with her handlers within law enforcement securely without arousing the suspicion of her colleagues within the criminal organisation. One of the most important factors in her choice of messaging applications is self-destructing messages. This means that even if her device is compromised, there will be no evidence that she has been sharing information with police.

Ewa - a 28 year old human rights activist and political dissident in a large authoritarian state in South-East Asia. As an important agent for change within the country, she needs to alert the country's overseas diaspora, human rights organisations and international media to a new and bloody crackdown on freedom of expression that occurred. The ruling party of the country has implemented state of the art surveillance technology and deep packet inspection on all outgoing web traffic. Ewa needs a means of communicating securely with her intended recipients via an insecure channel.

Tzvi - a 45 year old political "fixer" working within the ruling party of a large African state. He has been tasked with giving a veneer of legitimacy to the process of allocating government contracts to political allies through a fraudulent tendering process. This requires him to have a secure communication channel

with the chosen contractor so they can ensure they are able to make the most competitive bid. Tzvi has his own political ambitions which could be tarnished by such dealings, making it vitally important for him to have plausible deniability. Using a pseudonymous identity for messaging is a key selling point for him.

2.4 Requirements

Requirements for this application were gathered after a detailed analysis of the existing products in this space. All the features in the application are present in some sense in some or all of the competitor applications analysed with the exception of pseudonymous messaging. Combining several key features (Self-destructing messages, a combination of symmetric and asymmetric encryption and pseudonymous messaging) formed the basis for the design.

Requirements were then outlined following the MoSCoW requirement prioritisation framework[15]. This method allows for prioritisation based on four priority levels:

1. “Must have” - the most basic requirements to have a functioning solution for the problem.
2. “Should have” - Requirements that the app should include, but does not absolutely need.
3. “Could have” - Requirements that could potentially be implemented.
4. “Would like to have” - Features which would be nice to include but are not strictly necessary or required.

2.4.1 MoSCoW table

Prioritisation	Requirement
Must have	User can connect to the server application via a client application.
Must have	User can send a message from one client to another client.
Must have	User can receive a message from another client.
Must have	User can create a pseudonymous username.
Must have	User identity can persist upon disconnect/reconnection events
Must have	User can see other connected users within a contacts page
Must have	Messages (sent and received) are stored locally on the client within a local database.
Must have	User can talk to more than one user independently in separate messaging sessions.
Must have	User messages reach the correct recipient only.
Should have	User can send encrypted messages which cannot be read by the server operator.
Should have	User is able to define how long to store messages in the local database.
Should have	User receives a push notification when a message is received.
Should have	User preferences (e.g. self-destruct preferences) should persist on application restart

Should have	Recipient public keys should be shared with sender as appropriate.
Should have	Private/Public keypairs should be stored securely to prevent unwanted access.
Should have	Generating keypairs should be fast while also secure.
Could have	Group messaging between multiple users.
Could have	Users can access the application via a web-based interface in addition to the mobile application.
Could have	Users can format their text using standard decorations (e.g. bold/italics/strikethrough)
Could have	Users can forward messages from one user to another user.
Would like to have	Message queueing to allow users to receive messages when they reconnect.
Would like to have	Sender-defined self-destructing messages.
Would like to have	E2EE on group chats using client side fan-out method.
Would like to have	Users can record audio messages and transmit to another user.

2.5 User stories

User stories are a means to map requirements to actionable feature development and separate workload into manageable chunks and a key tenet of Agile methodology. Indeed, a 2018 study found that approximately 90% of Agile practitioners utilised user stories in their requirements gathering process[7]. The user stories below represent the desired behaviours from the application and were used as a framework for developing the application in an agile framework.

- As a user, I want to be able to send messages from my client application to the server for handling.
- As a user, I want my message to be routed to the appropriate client device of my choosing.
- As a user, I want to create a pseudonymous username which is shown to other users.
- As a user, I want to view other connected users.
- As a user, I want messages on my device to disappear after a given time period
- As a user, I want to encrypt my messages so that the server cannot read the contents.
- As a user, I want to be able to decrypt messages that are transmitted to me from the server.
- As a user, I want my data (e.g. username) to persist after I close the application.

- As a user, I want to receive a push notification when a message is sent to my device if I am doing something else.
- As a user, I want to be able to click on a notification and be taken directly to the relevant chat window to reply to the sender.
- As a user, I want the contacts list to automatically update when a user connects or disconnects from the server.
- As the server operator, I do not want to be able to read the content of messages sent to me.
- As the server operator, I want to be able to update stored client public keys when they are changed.

The above stories formed the basis for the development process, the intricacies of which are discussed extensively in the following chapter.

Chapter 3: Design and Implementation

3.1 High Level Architecture

Some of the primary architectural considerations on this project are highlighted below:

1. Client-side encryption to allow messages to flow through the (untrusted) server.
2. Client-side self-destructing messages according to client-defined storage duration parameter.
3. Persistence of user identity and handling of disconnection/reconnection events.

The figure below represents a very high level view of the system architecture, which gives an overarching perspective of how the system fits together and some of the technological choices made. However, it does not give a comprehensive picture of the complexity of the overall system.

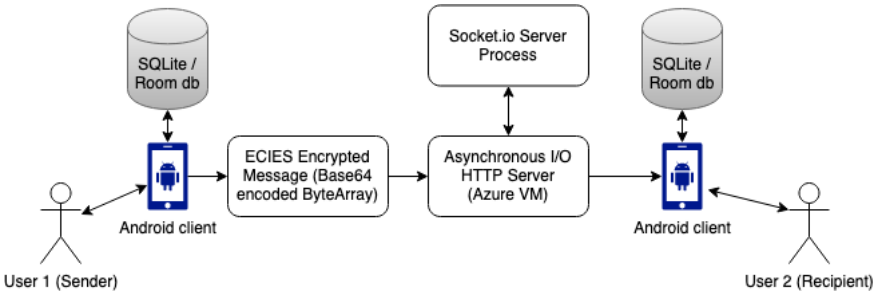

Figure 3.1: High level architecture of the system

3.2 WebSockets vs REST API

One of the major decisions required for this project was which transport protocol to use for propagating messages through the system. The two major options considered were a REST API via HTTP[27] and WebSockets[10].

REST (Representational State Transfer) is a set of design principles which assists in the development of web services which utilise a central server with one or many clients, based on the twin concepts of request and response. A client will send a request to a server and receive a response based on the content of the request. REST APIs are heavily used across many industries and mostly (though not exclusively) use the HTTP protocol to communicate between client and server. The primary weakness of REST is that it is not optimal

for circumstances where constant bi-directional communication is important between client and server. Indeed, the WebSocket Protocol Standards Track Document[10] states the problem as such:

```
Historically, creating web applications that need
bidirectional communication between a client and a server
(e.g. instant messaging and gaming applications) has required
an abuse of HTTP to poll the server for updates while
sending upstream notifications as distinct HTTP calls.
```

The WebSocket protocol is a newer system which excels in two-way communication between client and server systems. For chat applications like the subject of this project, WebSocket appears perfectly suited to the problem. Due to the comparative simplicity of bidirectional communication using WebSockets, the protocol uses significantly less energy[16] to maintain a client/server connection. This is especially relevant when considering that this is a mobile application which needs to preserve battery life. The one major downside of WebSockets is that it is a modern standard which is not necessarily fully supported across all devices or browsers yet.

During the literature review of this topic, it became apparent that there was a way to achieve the best of both worlds. Socket.io[31] is a cross-platform library which initially attempts to create a WebSocket connection and, if not possible, falls back to HTTP polling for devices which do not support it. It also provides helpful additional features like automatic reconnection. It was determined that both client and server implementations existed in the desired languages (Python/Kotlin) and subsequently included in the system design.

3.3 Encryption Implementation

As mentioned previously, the system was designed with the prevailing goal of having two clients who could communicate with each other through an insecure, untrusted environment (i.e. the server) without needing to worry about messages being vulnerable to man-in-the-middle attacks[23]. The most attractive means of achieving this is to use state of the art encryption.

An oft-repeated adage in secure software development is that a developer should ”Never roll your own crypto”[2]. This is due to the fact that encryption is almost always absolutely critical to the functioning of applications. Furthermore, it is very easy for a developer’s code to provide the illusion of working security features without fully understanding possible attack vectors. As such, all cryptography utilised in this project has an extensive track record of use in commercial and governmental settings.

Two of the most prevalent encryption approaches used in the modern world are asymmetric and symmetric encryption. Symmetric encryption is the simpler of the two - a secret key is used to alter the content of a piece of plain text in such a way that it’s unreadable to anyone except someone who also has the secret key and can use it to decrypt the message. The major flaw with symmetric encryption is the need to share the secret key with the recipient. If the recipient does not have access to the secret key, they cannot decrypt the message. Similarly,

if the secret key is intercepted during transport of the message, a bad actor can decrypt the message. A good example of the symmetric encryption approach is the Advanced Encryption Standard (AES), which is used in this project. The AES algorithm was designed by Vincent Rijmen and Joan Daemen in 1999 and defined as a standard by the United States Government in 2001.

Asymmetric encryption is a little more complex - a key pair consisting of a public and a private key, which are separate but mathematically linked is used to encrypt and decrypt data. To send an encrypted message to the intended recipient, the sender must use their own private key and the recipient's public key. The recipient can then use their own private key and the sender's public key to decrypt the message at the other side. The primary issue with asymmetric encryption is that it's significantly more complex to implement than symmetric alternatives, and key generation is slower.

Two of the most common types of asymmetric encryption are the Rivest Shamir Adleman (RSA) algorithm[32] and Elliptic Curve Cryptography (ECC)[22]. The RSA algorithm has been in use for 45 years, and utilises the factorisation of prime numbers to produce a unique keypair. ECC is a more modern algorithm which uses the structure of elliptic curves to generate keypairs. The primary advantage of ECC over RSA is the computational requirements to generate unique keys - as the key size required for secure communication increases (i.e. when malicious actors are able to use greater computational power), the key generation process becomes prohibitively expensive. In a 2015 study, an Indian team found that RSA keypair generation was 471 times slower than the equivalent ECC keypair process when keys reached a certain size[12].

The approach taken on this project was to use a combination of both asymmetric and symmetric approaches - this combination means that the comparative speed of symmetric approaches can be harnessed for converting the plaintext to ciphertext, and the additional security provided by asymmetric encryption, ensuring that only the intended recipient can open the message as the symmetric key is independently generated by both sender and receiver from a shared secret.The implementation used is known as an Elliptic Curve Integrated Encryption Scheme (ECIES)[26]. ECIES is a hybrid encryption system which was devised by Victor Shoup in 2001. It comprises of three key functions:

1. A key agreement function to generate "shared secrets" from a user's public key and another user's private key.
2. A key derivation function which is able to produce a key from an input of some kind.
3. A symmetric encryption algorithm which is used with the shared secret as a key to encrypt the plain text.

The advantage of ECIES is that it provides very strong encryption for secure communication within an untrusted/insecure channel while also having desirable properties for mobile devices such as much smaller keys than similar asymmetric/symmetric hybrid approaches (as shown by Martinez et al when comparing to RSA/AES hybrid systems[11])

The ECIES implementation used in this project uses the following algorithms and approaches:

- Elliptic Curve Diffie-Hellman[28] to generate a shared secret over an untrusted channel.
- The P-521 elliptic curve outlined by the National Institute of Standards and Technology (NIST)[5]
- The SHA-512 cryptographically secure hashing function, of which the first 32 bytes are used to generate the AES key for the key derivation function.
- AES-256 for the symmetric component, generated from the shared secret.

3.4 Self-destructing Messages

One of the features implemented in the application is self-destructing messages on a client device. This allows the user to define a duration to store received and sent messages on their device, after which the specific message is deleted. This feature was implemented by prompting the user on their first time starting the application to define a duration for message storage which is passed as a parameter to a daemon process (a background thread which runs at very low priority) which sends a parametrised SQL query to the database, deleting message objects based on their timestamp metadata. Discussion of how this duration is persisted after restarts can be found in the following section.

3.5 Persistence of users after disconnects

One of the most complex challenges faced during the development of the application was persisting clients after they disconnect and reconnect from the server. This was handled in the following manner:

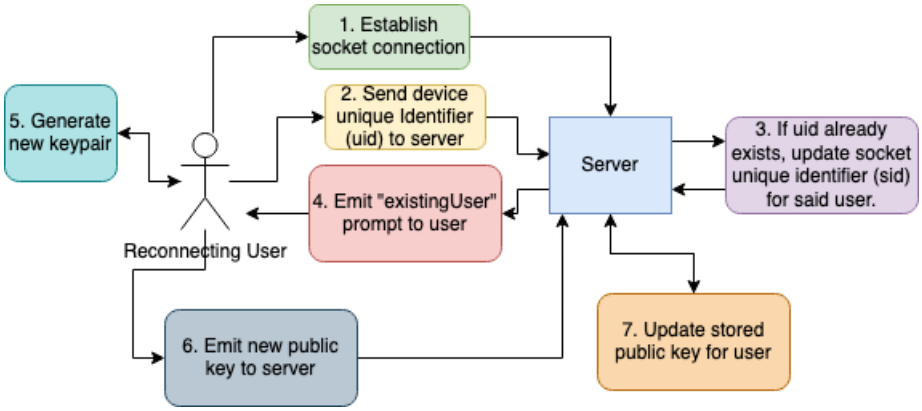

Figure 3.2: Reconnection flow for user persistence

The complexity of this process comes down to an inherent limitation of Android's built-in cryptography libraries. The system design was predicated on the assumption that it would be possible to securely store public and private keys within the Android KeyStore[6], a secure storage system within Android which is capable of storing public and private keys. However, the KeyStore does not

currently support Elliptic Curve keypairs, which meant that the design had to be altered to accomodate this. What this means in practice is that a new keypair is generated on reconnection - the private key is stored locally and the public key is emitted to the server which replaces the existing stored key.

The flexibility to easily generate new keypairs is a huge advantage of elliptic curve cryptography - the key generation process is so quick that it's imperceptible to the user - something that would not be possible with RSA encryption.

3.5.1 Storage of self-destruct duration

Although there is no secure key-value store that allowed for the storage of ECDH keypairs, there is a non-secure method for persistence of key-value data (not including the database) using "SharedPreferences" files. These are ideal for storing configuration properties which persist after closing the application. This means that the initial value prompted from the user to request self-destruct message duration is able to be persisted and a returning user will be able to maintain their configuration, while also maintaining separation of concerns and reducing excessive exposure to database operations.

3.6 Message Propagation

The figure below describes the entire journey of a message through the messaging system. The colour scheme for the various key stages is broadly represented by the following:

1. Green: Initial message composition, server querying and local storage.
2. Yellow: Asymmetric + symmetric encryption, encoding and send process
3. Red: Server routing for message
4. Purple: Receipt, decryption, display and storage of message.

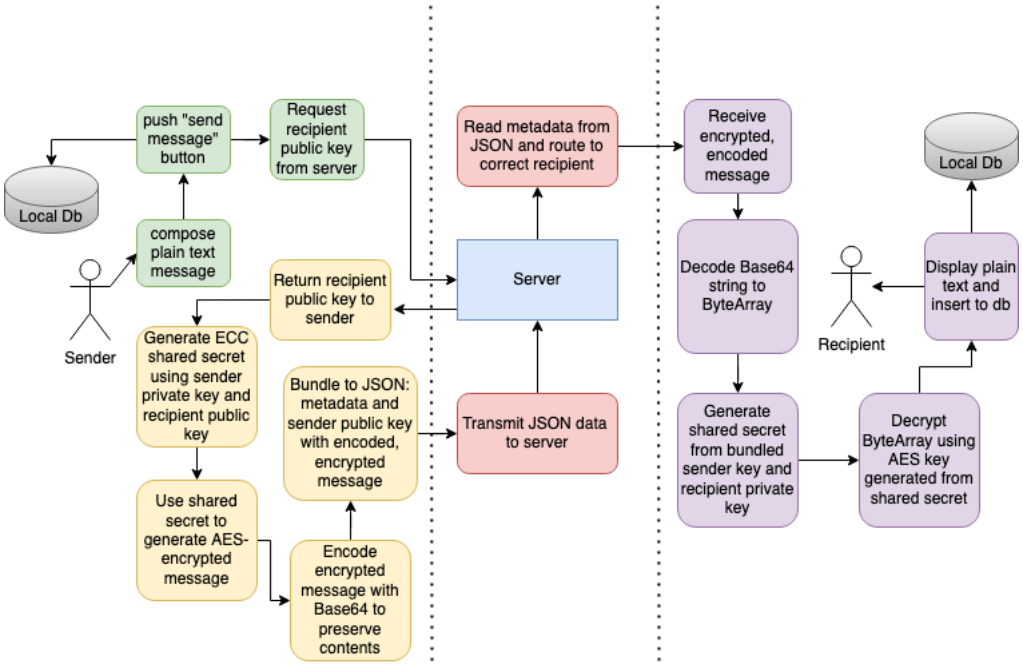

Figure 3.3: Flowchart of the journey of a message from one client to another

3.7 Design Patterns

During the development of this project, a concerted effort was made to follow established design patterns and best practice wherever possible. It is a testament to the design of the Android development environment and documentation that this was relatively trivial to achieve.

In addition, wherever possible, adherence to the “Don’t repeat yourself” (DRY)[17], “Keep it simple, stupid” (KISS), and five SOLID principles[25] was paramount.

The table below shows the design patterns which were deliberately utilised in the development of this project:

Pattern	Concise Definition	Implementation
Command Pattern[34]	Encapsulation of a request into an object	Socket.io event handler objects in both client and server
Observer Pattern[8]	When one object changes state, dependents update accordingly	LiveData queries updating views as events occur (ChatWindow and Contacts)
Adapter Pattern[14]	Allow an interface of one class to be used as another interface	ContactsAdapter for RecyclerView and MessageListAdapter for displaying chat messages.
Builder Pattern[33]	Simplifies object creation step-by-step	Push notification builder object allows for customisable notifications.
DAO Pattern[21]	Data Access Objects provide abstract interface for database, separating low level and high level operations	ChatMessageDAO allows us to specify the exact queries we can use on the database and prevents any other type of unplanned operations.
Singleton Pattern[37]	Restrict instantiation of class to exactly one instance	SocketHandler.kt in the client is an example of a class which should be (and is) a singleton.

It is also worth mentioning that the Model-View-Viewmodel architectural style was adopted throughout the application, as it is the recommended architectural pattern recommended by the Android core development team[13]. This architectural style is most clearly seen in all interactions between the database, the repository, the data access object and the ViewModel which displays data in the ChatWindow.kt activity.

3.8 User Interface and Visual Design

User interface was not a primary concern for the application, but during the requirements process, Balsamiq was used to design simple mockups of the application. These mockups do not reflect the final state of the application but they are included in the appendices for posterity.

Future development of the application will in part focus on improving the visual design.

3.9 Development tooling

There are two primary components to the overall system - one or more Android mobile client applications which communicate with a central, deployed server.

3.9.1 Client

The final client application was written entirely in the Kotlin language (v1.6.10) using the Android Studio IDE.

Development and debugging process

Traditional Android software development involves using the qEmu device emulator built into Android Studio to imitate the experience of a real user on their own device.

Debugging code during the HushChat development process was a complex endeavour, as the development machine had fairly limited RAM. As the purpose of the application was communication between two clients, it was important to be able to emulate two devices at the same time so messages could be exchanged. This was initially possible on the development machine, but became unfeasible over time as the complexity of the system grew and memory capacity became constrained.

The debugging environment eventually morphed into a single qEmu emulated device on the development machine combined with a physical android device (Google Pixel 4). This methodology became possible with the introduction of Android version 11 and Android Studio 'Bumblebee' (v2021.1), which allows the developer to run Android debugging tools (i.e the Android Debug Bridge, or ADB) via WiFi connection. This feature was extremely helpful in the development process, and it allowed for the introduction of further feature complexity that would not have been possible otherwise due to the technical limitations of the development machine.

Running a WebSocket-based application locally also presented a minor challenge, as qEmu virtual devices use the traditional localhost/127.0.0.1 address to represent their own internal loopback interface rather than that of the host development machine. This meant that to communicate with the locally hosted server, the special debugging address 10.0.2.2 was used. After the development of the server side of the application was complete, this problem was solved by deploying the server to a static IP address (via a Azure Virtual Machine).

Key Dependencies

- Socket.io - v2.0.0 - the developers of socket.io provide a native Java implementation. Due to Java's seamless interoperability with Kotlin, this did not cause any implementation problems with the version implemented.
- BouncyCastle - v1.67 - BouncyCastle is the cryptography API that performs a lot of the cryptographic operations within the application, including generation of elliptic curve keypairs. There is a built-in version of BouncyCastle within the Android SDK, but this does not support ECC, which meant that the client application needed to replace the inbuilt library with a more updated version. This can be seen within the codebase with the following calls within the MainActivity and ChatWindow classes:

  ```
  Security.removeProvider("BC")
  Security.addProvider(BouncyCastleProvider())
  ```

- Room ORM - v2.4.0 - Room provides an object-relational mapping (ORM) over the SQLite database built into the Android SDK. When developing the application, the main data model to consider on the client side was that of an individual chat message, including metadata (e.g. recipient) and

content. Working with Room allowed for a more streamlined development experience and reduction of boilerplate code, as it meant working directly with Kotlin (Java) entities instead of writing complex queries for inserting messages to our viewmodel. Traditional SQL queries were still used to capture relevant data to display to the user in chat windows.

3.9.2 Server

The server application was written in Python v3.9.1 in the PyCharm IDE and utilised several libraries which are highlighted below.

Key Dependencies

- python-socketio - v5.5.0 - a Python server implementation of socket.io, funded by the original socket.io developers.
- aiohttp - v3.8.1 - an asynchronous http server which (importantly) supports websockets. The socket.io process attaches itself to the aiohttp server which allows information to be transmitted and received from clients.

Deployment

To allow users to message other users, it was necessary to deploy the server application on a static IP address. To do this, a 1GB/1CPU Virtual Machine instance was used, running Ubuntu 20.04 in the West Europe region of the Microsoft Azure cloud platform.

3.9.3 Final report

This dissertation was written entirely in LaTeX using Neovim v0.5.1 with the VimTex plugin and the Skim PDF reader (which auto-refreshes at compile time).

3.9.4 Version Control

Git was the version control solution used throughout the project. Mild discomfort with the idea of storing the written report with a private company (Overleaf) while writing meant that it was worth taking the step of compiling the document locally. This had the fortunate benefit of meaning the entire project could be stored in a single git repository (as well as an opportunity to learn a lot of new things).

The host of choice for the upstream repository on this project was GitHub, and the entirety of the project can be found at the following link. It is important to note that the commenting approach may appear excessive (and does not follow the “Code tells you how, Comments tell you why” commenting practice) but this was done with the intention of assisting the reader to get through the code as quickly as possible.

https://github.com/furnivall/SEng_Final_Project

Chapter 4: Evaluation & Testing

4.1 User evaluations

4.1.1 Methodology

4.1.2 Demographics

4.1.3 Issues highlighted

Chapter 5: Conclusion

Appendix A: First appendix

A.1 Section of first appendix

Appendix B: Second appendix

Bibliography

[1] J. Alwen, S. Coretti, and Y. Dodis. The double ratchet: security notions, proofs, and modularization for the signal protocol. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 129–158. Springer, 2019.

[2] A. Apvrille and M. Pourzandi. Secure software development by example. *IEEE Security & Privacy*, 3(4):10–17, 2005.

[3] R. Baeza-Yates and U. M. Fayyad. The attention economy and the impact of artificial intelligence. In *Perspectives on Digital Humanism*, pages 123–134. Springer, Cham, 2022.

[4] H. Berghel. Malice domestic: the cambridge analytica dystopia. *Computer*, 51(5):84–89, 2018.

[5] M. Brown, D. Hankerson, J. López, and A. Menezes. Software implementation of the nist elliptic curves over prime fields. In *Cryptographers' Track at the RSA Conference*, pages 250–265. Springer, 2001.

[6] T. Cooijmans, J. de Ruiter, and E. Poll. Analysis of secure key storage solutions on android. In *Proceedings of the 4th ACM Workshop on Security and Privacy in Smartphones & Mobile Devices*, pages 11–20, 2014.

[7] F. Dalpiaz and S. Brinkkemper. Agile requirements engineering with user stories. In *2018 IEEE 26th International Requirements Engineering Conference (RE)*, pages 506–507. IEEE, 2018.

[8] A. Eales and A. Eales. The observer pattern revisited. *Educating, Innovating & Transforming: Educators in IT: Concise paper*, 2005.

[9] A. Farahat and M. C. Bailey. How effective is targeted advertising? In *Proceedings of the 21st international conference on World Wide Web*, pages 111–120, 2012.

[10] I. Fette and A. Melnikov. The websocket protocol, 2011.

[11] V. Gayoso Martínez, L. Hernandez Encinas, and A. Queiruga-Dios. Security and practical considerations when implementing the elliptic curve integrated encryption scheme. *Cryptologia*, 39:1–26, May 2015. DOI: 10.1080/01611194.2014.988363.

[12] M. Gobi, R. Sridevi, and R. Rahini. A comparative study on the performance and the security of rsa and ecc algorithm. In *Proceedings of Conference on Advanced Networking and Applications*, 2015.

[13] Guide to app architecture - android developers. URL: https://developer.android.com/jetpack/guide#recommended-app-arch.

[14] R. Harmes and D. Diaz. The adapter pattern. *Pro JavaScript Design Patterns*:149–158, 2008.

[15] D. Haughey. Moscow method. *Project Smart*:2011, 2011.

[16] V. Herwig, R. Fischer, and P. Braun. Assessment of rest and websocket in regards to their energy consumption for mobile applications. In *2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, volume 1, pages 342–347. IEEE, 2015.

[17] A. Hunt and D. Thomas. *The Pragmatic Programmer: From Journeyman to Master*. Addison-Wesley Longman Publishing Co., Inc., USA, 2000. ISBN: 020161622X.

[18] J. Jakobsen and C. Orlandi. On the cca (in) security of mtproto. In *Proceedings of the 6th Workshop on Security and Privacy in Smartphones and Mobile Devices*, pages 113–116, 2016.

[19] M. L. Jansen. A security analysis of the signal protocol's group messaging capabilities in comparison to direct messaging, 2020.

[20] D. Kreiss and S. C. McGregor. The "arbiters of what our voters see": facebook and google's struggle with policy, process, and enforcement around political advertising. *Political Communication*, 36(4):499–522, 2019.

[21] S. LONG, H. LIN, and X. CHEN. Data access object pattern. *Computer and Modernization, pp5*, 2004.

[22] J. Lopez and R. Dahab. An overview of elliptic curve cryptography, 2000.

[23] A. Mallik. Man-in-the-middle-attack: understanding in simple words. *Cyberspace: Jurnal Pendidikan Teknologi Informasi*, 2(2):109–134, 2019.

[24] B. Marczak, J. Scott-Railton, S. McKune, B. Abdul Razzak, and R. Deibert. HIDE AND SEEK: Tracking NSO Group's Pegasus Spyware to operations in 45 countries. Technical report, 2018.

[25] R. C. Martin. Design principles and design patterns. *Object Mentor*, 1(34):597, 2000.

[26] V. G. Martınez, L. H. Encinas, et al. A comparison of the standardized versions of ecies. In *2010 Sixth International Conference on Information Assurance and Security*, pages 1–4. IEEE, 2010.

[27] M. Masse. *REST API design rulebook: designing consistent RESTful web service interfaces*. " O'Reilly Media, Inc.", 2011.

[28] U. M. Maurer and S. Wolf. The diffie–hellman protocol. *Designs, Codes and Cryptography*, 19(2):147–171, 2000.

[29] M. Mazzetti, A. Goldman, R. Bergman, and N. Perlroth. A new age of warfare: how internet mercenaries do battle for authoritarian governments. *The New York Times*, 21:2019, 2019.

[30] K. Pantelis and L. Aija. Understanding the value of (big) data. In *2013 IEEE International Conference on Big Data*, pages 38–42. IEEE, 2013.

[31] R. Rai. *Socket. IO Real-time Web Application Development*. Packt Publishing Ltd, 2013.

[32] R. L. Rivest, A. Shamir, and L. M. Adleman. *A method for obtaining digital signatures and public key cryptosystems*. Routledge, 1982.

[33] V. Sarcar. Builder patterns. In *Java Design Patterns*, pages 89–95. Springer, 2016.

[34] V. Sarcar. Command pattern. In *Design Patterns in C#*, pages 315–335. Springer, 2020.

[35] R. Simmons. Big data and procedural justice: legitimizing algorithms in the criminal justice system. *Ohio St. J. Crim. L.*, 15:573, 2017.

[36] P. Sommer. Evidence from hacking: a few tiresome problems. *Forensic Science International: Digital Investigation*, 40:301333, 2022.

[37] K. Stencel and P. Wegrzynowicz. Implementation variants of the singleton design pattern. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*, pages 396–406. Springer, 2008.

[38] C. Stöcker. How facebook and google accidentally created a perfect ecosystem for targeted disinformation. In *Multidisciplinary International Symposium on Disinformation in Open Online Media*, pages 129–149. Springer, 2019.

[39] C. Walter. Kryder's law. *Scientific American*, 293(2):32–33, 2005.

[40] S. Zawoad and R. Hasan. Digital forensics in the age of big data: challenges, approaches, and opportunities. In *2015 IEEE 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems*, pages 1320–1325. IEEE, 2015.