

IMPROVISING IMAP PROTOCOL BY INTEGRATING SPAM DETECTION MODEL AND IMPLEMENTING SECURE COMMUNICATION USING RSA

Presented By

124157027: Inti Dhiraj,

Guide Name: Dr. R. Kavitha, CSE/SoC, SASTRA University

Date of Presentation: 14.12.2022

Points to be discussed

1. Aim & Objective
2. Problem Statement
3. Existing Solutions - Literature Survey
4. System Configuration : Software, Hardware, Network type, Protocols used
5. Network Architecture and Workflow
6. Work Done (Implementation code link)
7. Sample I/O (screen shot)
8. References

AIM & OBJECTIVES

The aim of the project is implementing the SMTP, IMAP protocols, in wired and wireless network to simulate the application on Client-Server environment(here Gmail server).

Objective 1: Understanding SMTP, IMAP, MIME protocols

Objective 2: Developing programs to make use of SMTP, IMAP protocols

Objective 3: Creating the spam filter model, RSA cryptosystem

Objective 4: Benefits to the E-mail users

Problem Statement: Nowadays many fraudulent mails pop on email clients inbox and he/she clicks on the malicious redirectable links and hence are betrayed or that causes a financial/moral/resources damage such as anonymous money transfers to another account, a virus getting injected into the system, etc.

Keywords: SMTP, IMAP, RSA, Machine Learning

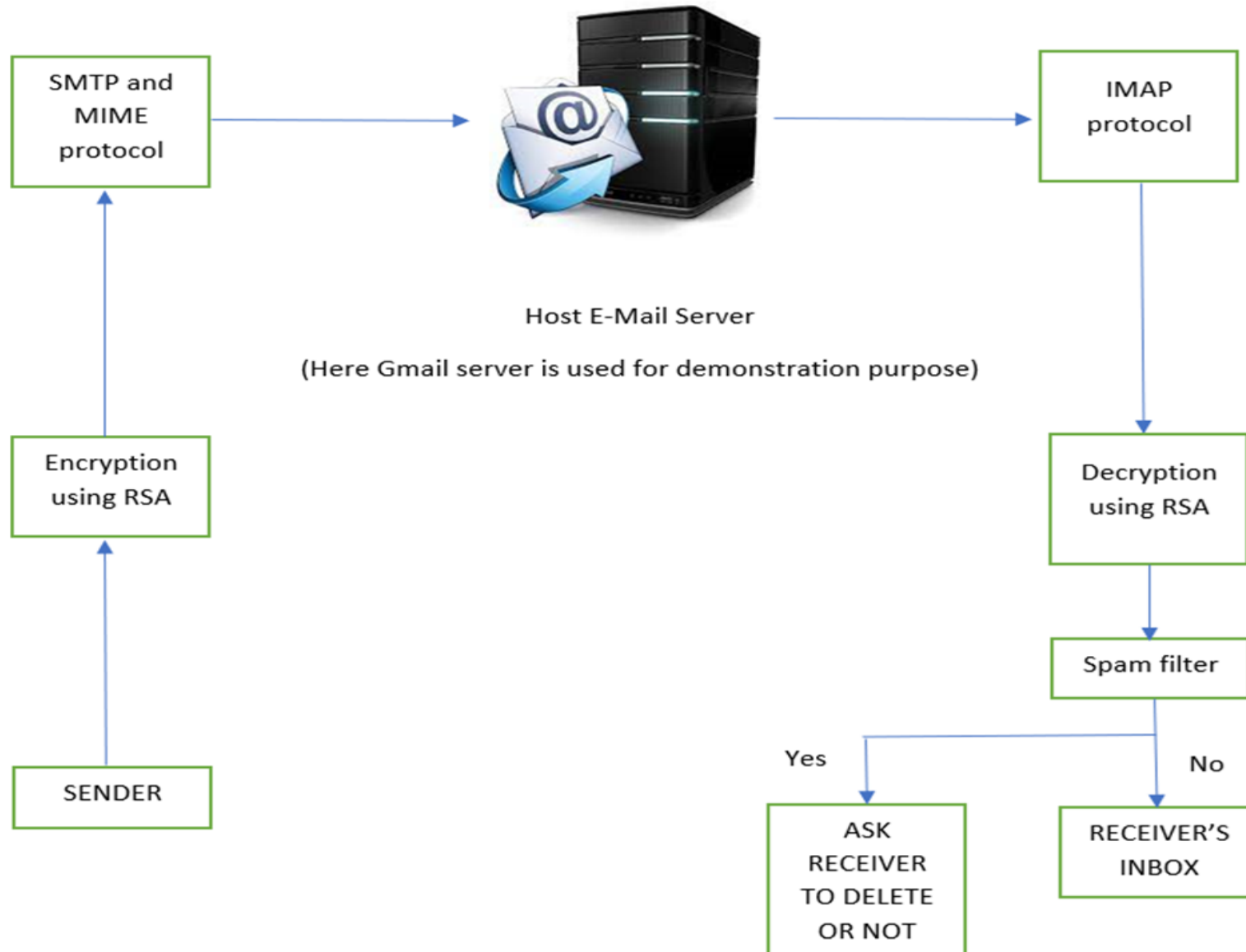
Solution: The key idea is to improvise the existing IMAP by introducing a spam filter, developed using Machine Learning, which will surely affect the current scenario and can decrease the number of fraudulent activities happening. Increases confidentiality of message transfer using RSA cryptosystem over third party servers.

Existing Solutions- LITERATURE SURVEY

| S.No | Title of the paper | Year | Comments on base and reference papers |
|------|--|------|---|
| 1 | The Development of the Open Machine-Learning Based Anti-Spam Open-MaLBAS | 2021 | Gives an idea about SMTP module, e-mail users security, Machine Learning approaches |
| 2 | A Comprehensive Survey for Intelligent Spam Email Detection | 2019 | Machine learning, phishing attack, spam detection, spam email, spam filtering. |

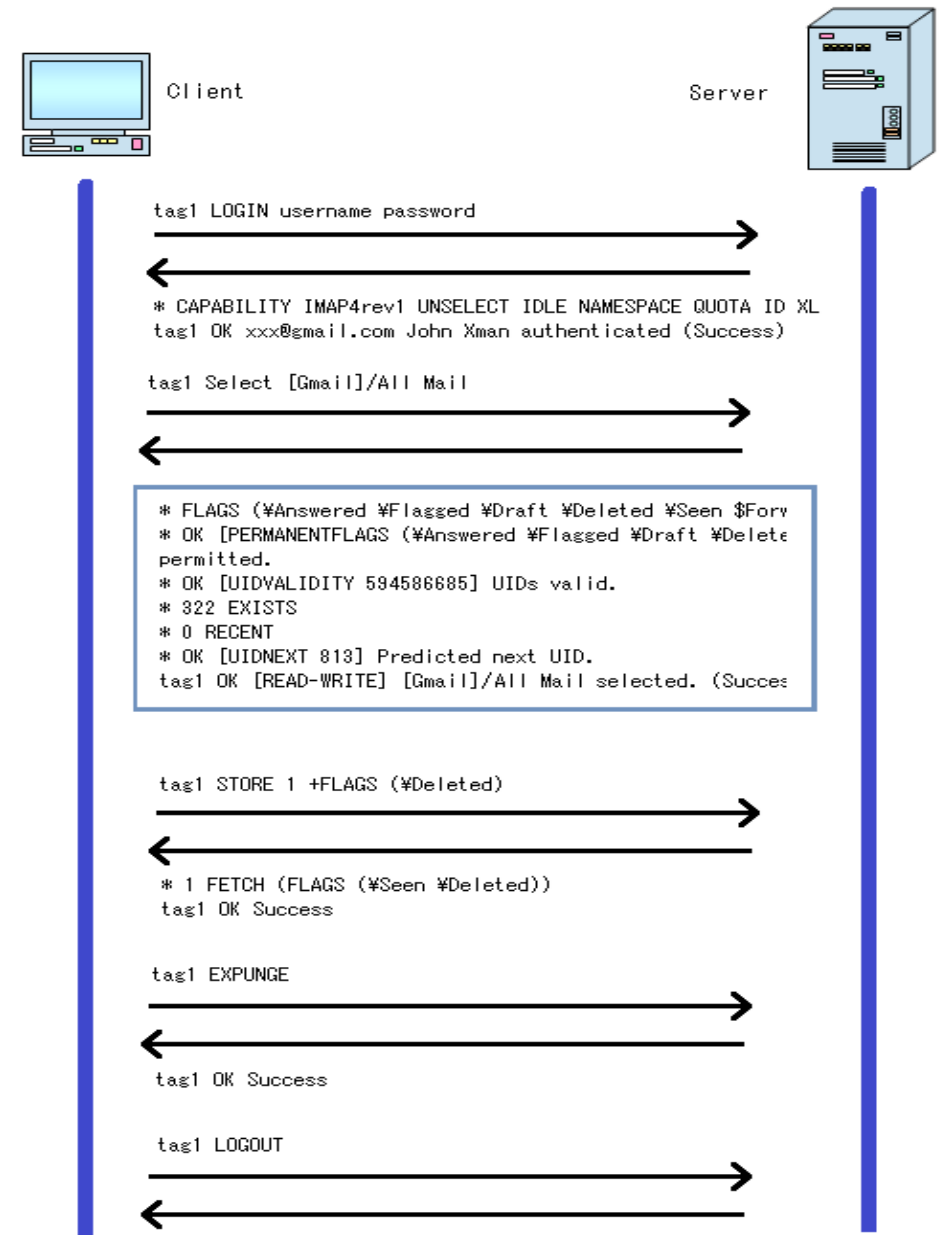
- Client side
 - Hardware used: i5 11th Gen processor
 - Software used: Python programming
 - Network: wireless
 - Protocol: SMTP/IMAP
- Server side
 - Hardware used: G-Mail Servers
 - Protocol: SMTP/IMAP

The Work Flow



Proposed Protocol Description

- **RFC number: 9051**
- **Features of the protocol :**
 - 1)IMAP only downloads a message to the client when the user clicks on it; attachments are not automatically downloaded.
 - 2)An IMAP server listens on port number 143, while IMAP over Secure Sockets Layer (SSL)/Transport Layer Security is assigned port number 993.
 - 3)A spam filter integrated with the protocol to detect spams.
- **Limitation of the protocol :**
 - 1)Mails won't work without an active internet connection.
 - 2)Accessing mails is little slower as compared to POP3, as all folders get synchronized every time there is a Send / Receive.



RSA CRYPTOGRAPHY:

RSA cryptosystem is one of the classical cryptography algorithms in which the keys are receiver generated key pairs, the public key is given to the authoritative sender and the private key is kept with the receiver. The assumption made while designing RSA system is it is difficult to factorize a large number which is in fact a product of two prime numbers. The public key consists of two numbers where one number is a multiplication of two large prime numbers. Private key is also derived from those two primes numbers. So, suppose the large number gets factorized, the private key will be unearthed. Therefore, encryption strength totally lies on the key size, solution to this can be increasing the size of the key, or the prime numbers contributing to form the large number. It is a receiver-initiated algorithm hence ensures confidentiality.

Step 1: Alice encrypts the message using Bob's public key in step one.

Step 2: Bob receives the encrypted message

Step 3: Bob decrypts the message using his private key

The Spam Filter

A program known as a spam filter is used to identify unsolicited, undesired, and virus-infected emails and stop them from reaching a user's mailbox. A spam filter, like other filtering systems, searches for certain standards on which to base its decisions.

Email spam filtering systems are used by Internet service providers (ISPs), enterprises, and free online email services to reduce the danger of spam distribution. For instance, one of the first and most basic spam filtering systems, like the one used by Microsoft's Hotmail, was configured to look for specific terms in message subject lines. When the filter detected one of the predefined terms, the email was not added to the user's inbox.

This approach is not particularly efficient and frequently leaves out completely acceptable messages, known as false positives, while allowing true spam messages to pass. Using suspicious word patterns or word frequency, more advanced systems, including Bayesian filters and other heuristic filters, can identify spam messages. They accomplish this by gaining insight into the user's preferences from the emails flagged as spam. When new emails are sent that are intended for the user's inbox, the spam program creates rules and applies them. Here we have developed an ML model using logistic regression model with an **accuracy of 96.70%**.

Work done so far

Dataset used:

https://drive.google.com/file/d/1uzbhec5TW_OjFr4UUZkoMm0rpyvYdhZw/view

Source code:

https://drive.google.com/drive/folders/1rDcz6oHNJOv6h_BbQebHd4DBGPgMRM8o

Output video:

https://drive.google.com/file/d/1kJJ3i1okX8NMIKIGFqjAujxBEFo1bAf3/view?sp=share_link

Sample I/O screenshots

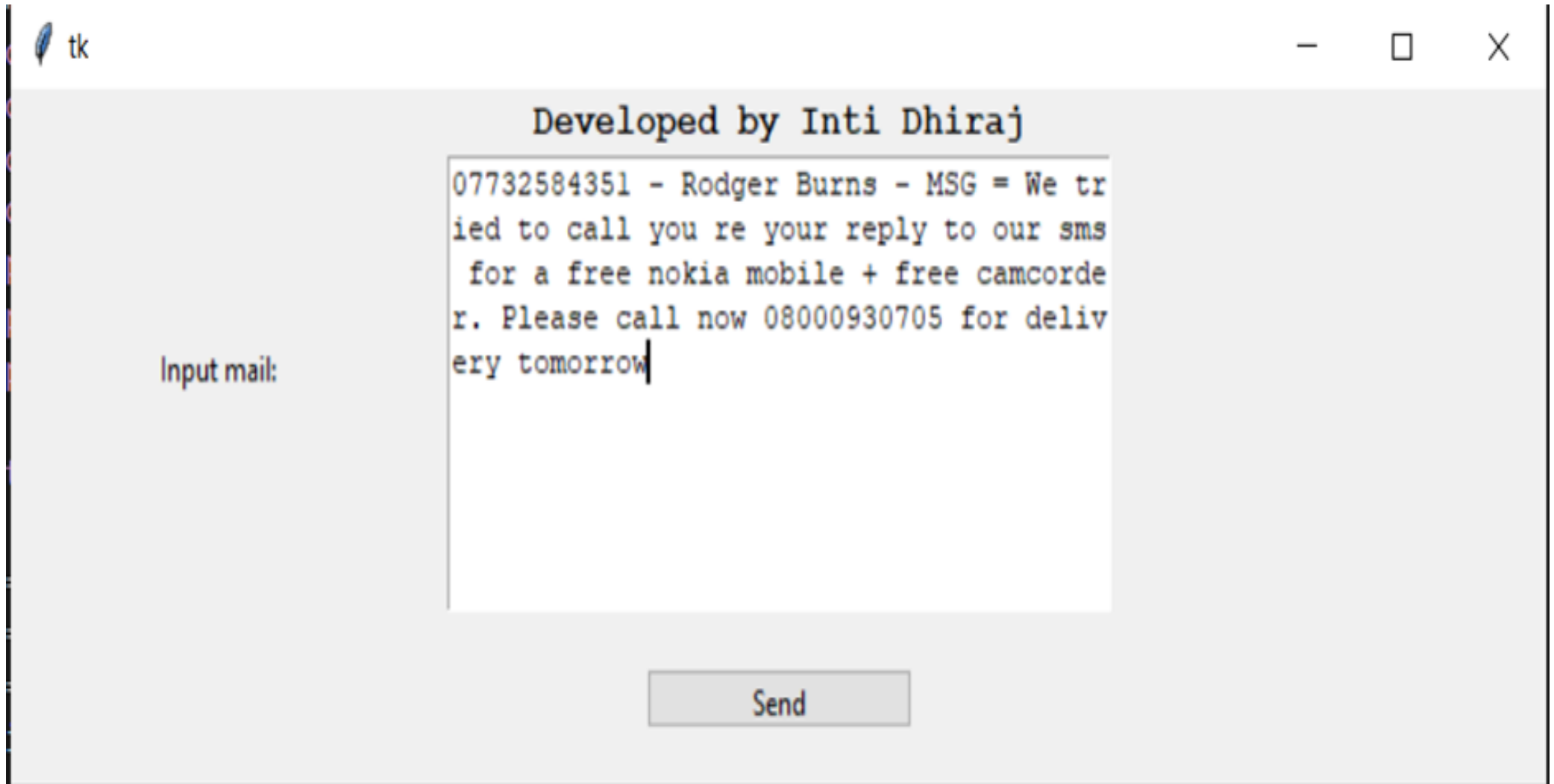
Input: Mail body

```
inp = ["WINNER!! As a valued network customer you have been selected to receivea £900 prize reward! To "  
      "claim call 09061701461. Claim code KL341. Valid 12 hours only."]
```

Output: Spam or Ham

```
C:\Python\python.exe "D:/SASTRA/Machine Learning/Spam Detection Model/predictor.py"  
[0]  
Spam mail
```

GUI to send mail: -



tk

Developed by Inti Dhiraj

Input mail:

07732584351 - Rodger Burns - MSG = We tried to call you re your reply to our sms for a free nokia mobile + free camcorde r. Please call now 08000930705 for deliv ery tomorrow

Send

Sending mail: -

```
PS D:\SASTRA\B Tech CSE materials\SEM 5\CN PROJECT 124157027\Src> python -u "d:\SASTRA\B Tech CSE materials\SEM 5\CN PROJECT 124157027\Src\smtp_impl.py"
send: 'mail FROM:<124157027@sastra.ac.in> size=927\r\n'
reply: b'250 2.1.0 OK c4-20020a17090a674400b002189ab866bfsm4558788pjm.5 - gsmtpl\r\n'
reply: retcode (250); Msg: b'2.1.0 OK c4-20020a17090a674400b002189ab866bfsm4558788pjm.5 - gsmtpl'
send: 'rcpt TO:<124157027@sastra.ac.in>\r\n'
reply: b'250 2.1.5 OK c4-20020a17090a674400b002189ab866bfsm4558788pjm.5 - gsmtpl\r\n'
reply: retcode (250); Msg: b'2.1.5 OK c4-20020a17090a674400b002189ab866bfsm4558788pjm.5 - gsmtpl'
send: 'data\r\n'
reply: b'354 Go ahead c4-20020a17090a674400b002189ab866bfsm4558788pjm.5 - gsmtpl\r\n'
reply: retcode (354); Msg: b'Go ahead c4-20020a17090a674400b002189ab866bfsm4558788pjm.5 - gsmtpl'
data: (354, b'Go ahead c4-20020a17090a674400b002189ab866bfsm4558788pjm.5 - gsmtpl')
send: b'Content-Type: multipart/mixed; boundary="=====2954395000233466684=="\r\nMIME-Version: 1.0\r\nSubject: Machine generated mail\r\n\r\n-----2954395000233466684==\r\nContent-Type: text/plain; charset="us-ascii"\r\nMIME-Version: 1.0\r\nContent-Transfer-Encoding: 7bit\r\n\r\n48 286 286 324 267 200 84 171 324 200 112 74 311 74 19 251 247 131 115 226 74 199 201 226 166 10
8 74 311 74 266 97 316 74 96 74 73 115 74 228 226 91 115 247 74 228 251 74 218 314 143 143 74 9 251 201 74 226 115 74 9 251 201 226
74 226 115 196 143 9 74 228 251 74 251 201 226 74 108 67 108 74 4 251 226 74 314 74 4 226 115 115 74 166 251 135 91 314 74 67 251 42
91 143 115 74 99 74 4 226 115 115 74 218 314 67 218 251 226 247 115 226 46 74 45 143 115 314 108 115 74 218 314 143 143 74 166 251
7 74 48 84 48 48 211 324 48 286 48 200 74 4 251 226 74 247 115 143 91 34 115 226 9 74 228 251 67 251 226 226 251 7\r\n-----
=====2954395000233466684==--\r\n.\r\n'
reply: b'250 2.0.0 OK 1669368883 c4-20020a17090a674400b002189ab866bfsm4558788pjm.5 - gsmtpl\r\n'
reply: retcode (250); Msg: b'2.0.0 OK 1669368883 c4-20020a17090a674400b002189ab866bfsm4558788pjm.5 - gsmtpl'
data: (250, b'2.0.0 OK 1669368883 c4-20020a17090a674400b002189ab866bfsm4558788pjm.5 - gsmtpl')
send: 'quit\r\n'
reply: b'221 2.0.0 closing connection c4-20020a17090a674400b002189ab866bfsm4558788pjm.5 - gsmtpl\r\n'
reply: retcode (221); Msg: b'2.0.0 closing connection c4-20020a17090a674400b002189ab866bfsm4558788pjm.5 - gsmtpl'
```

GUI on recipient side (Clicks yes):-

undefined

Developed by Inti Dhiraj

Mail received:

07732584351 - Rodger Burns
- MSG = We tried to call you
re your reply to our sms for
a free nokia mobile + free
camcorder. Please call now
08000930705 for delivery
tomorrow

Result: Spam

Delete the mail?(yes/no)

Yes No

Receiver clicks on yes and mail is deleted:-

```
PS D:\SASTRA\B Tech CSE materials\SEM 5\CN PROJECT 124157027\Src> python -u "d:\SASTRA\B Tech CSE materials\SEM 5\CN PROJECT 124157027\Src\imap_impl.py"
```

Subject: Machine generated mail

From: 124157027@sastra.ac.in

```
=====
```

07732584351 - Rodger Burns - MSG = We tried to call you re your reply to our sms for a free nokia mobile + free camcorder. Please call now 08000930705 for delivery tomorrow
It is a Spam mail

```
=====
```

Deleting Machine generated mail

IEEE papers: -

- A. Karim, “Comprehensive Survey for Intelligent Spam Email Detection”, December 4, 2019).
- Maryam Hina, Mohsan Ali, Abdul Rehman Javed, Gautam Srivastava, Thippa Reddy Gadekallu, Zunera Jalil, “Email Classification and Forensics Analysis using Machine Learning”, 2021.

Online resources: -

- [https://www.techtarget.com/whatis/definition/IMAP-Internet-Message-Access-Protocol#:~:text=The%20headers%20of%20all%20emails,Office%20Protocol%203%20\(POP3\).](https://www.techtarget.com/whatis/definition/IMAP-Internet-Message-Access-Protocol#:~:text=The%20headers%20of%20all%20emails,Office%20Protocol%203%20(POP3).)
- <https://www.youtube.com/watch?v=rxkGItX5gGE>
- [https://www.techtarget.com/whatis/definition/MIME-Multi-Purpose-Internet-Mail-Extensions#:~:text=MIME%20\(Multipurpose%20Internet%20Mail%20Extensions\)%20is%20an%20extension%20of%20the,and%20application%20programs%2C%20over%20email.](https://www.techtarget.com/whatis/definition/MIME-Multi-Purpose-Internet-Mail-Extensions#:~:text=MIME%20(Multipurpose%20Internet%20Mail%20Extensions)%20is%20an%20extension%20of%20the,and%20application%20programs%2C%20over%20email.)
- <https://realpython.com/python-send-email/>

Thank You