## MIT WORLD PEACE UNIVERSITY

Digital Forensics and Investigation Third Year B. Tech, Semester 5

# ANALYSING EMAIL HEADERS

LAB ASSIGNMENT 5

Prepared By

Krishnaraj Thadesar Cyber Security and Forensics Batch A1, PA 20

October 31, 2023

## Contents

| 1 | Aim   | 1  |
|---|---|--|
| 2 | Objectives  | 1  |
| 3 | Theory 3.1 Email Clients  | 1<br>1<br>3<br>4<br>4  |
| 4 | Platform  | 5  |
| 5 | 5.1.6 DKIM-Signature 5.1.7 From 5.1.8 To 5.1.9 Subject 5.1.10 Date 5.1.11 Message-ID 5.1.12 Content-Type 5.1.13 Return-Path 5.1.14 X-Sender-IP 5.1.15 X-MS-Exchange-Transport-EndToEndLatency 5.1.16 X-Microsoft-Antispam-Mailbox-Delivery 5.1.17 X-Message-Info and X-Message-Delivery | 5<br>8<br>9<br>9<br>10<br>10<br>10<br>11<br>11<br>11<br>11<br>12<br>12<br>12 |
| 6 | 6.1 Using VPN   | 15<br>15<br>16   |
| 7 |   | 17<br>19   |
|   |   | 20   |

## 1 Aim

To learn about the various types of email headers, and how to analyse them. To perform a live practical upon sent and received mails, while analysing the headers of the same.

## 2 Objectives

- 1. Understand the various types of email headers.
- 2. Learn how to analyse email headers.

## 3 Theory

#### 3.1 Email Clients

Email clients, also known as email programs or email software, are applications or platforms designed for managing, sending, and receiving email messages. They are essential tools for communication in both personal and professional settings. Here are key points related to email clients:

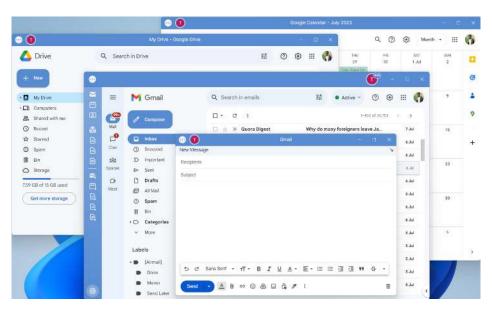
- Purpose of Email Clients: Email clients are designed to provide a user-friendly interface for managing email communication. They allow users to read, compose, send, and organize emails.
- 2. **Types of Email Clients:** There are various types of email clients, including desktop clients (e.g., Microsoft Outlook, Mozilla Thunderbird), web-based clients (e.g., Gmail, Outlook.com), and mobile clients (e.g., Apple Mail, Gmail app).
- 3. **Features of Email Clients:** Email clients offer a wide range of features, including the ability to access multiple email accounts, organize emails into folders, set up filters and rules, and manage attachments.

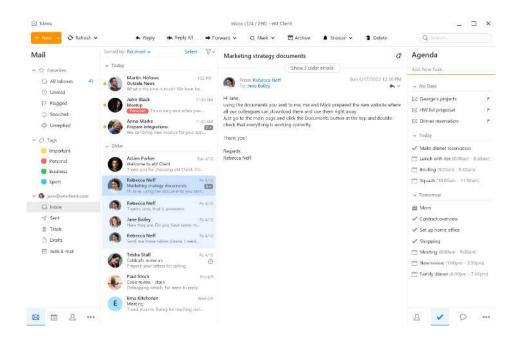


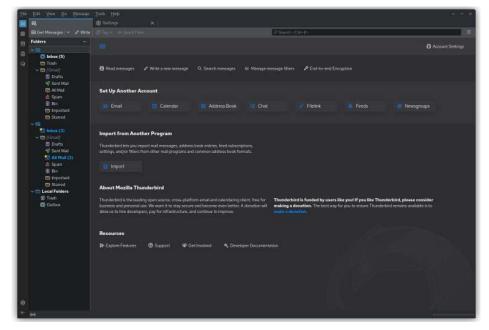
4. **Integration with Protocols:** - Email clients integrate with email protocols such as IMAP (Internet Message Access Protocol) and POP3 (Post Office Protocol) to retrieve and store emails on local devices or remote servers.

- 5. **User Interface:** The user interface of email clients varies, but it typically includes an inbox, folders for organization, a compose window, and options for formatting emails.
- Cross-Platform Compatibility: Many email clients are available on multiple platforms, making it easy for users to access their emails on different devices such as computers, smartphones, and tablets.
- 7. **Security and Encryption:** Email clients often support encryption and security features to protect sensitive information, including SSL/TLS for secure connections and S/MIME for email message encryption.
- 8. **Customization:** Users can often customize the appearance and behavior of email clients through themes, plugins, and settings to suit their preferences and needs.
- 9. **Productivity Tools:** Some email clients offer productivity tools like calendars, task management, and contact organization, making them comprehensive communication hubs.
- 10. **Email Standards:** Email clients adhere to email standards such as RFC 5322 for formatting and displaying emails, ensuring compatibility and consistent rendering.

Email clients play a pivotal role in modern communication, providing users with a centralized platform for managing their electronic correspondence.







### 3.2 Email Headers

Email headers are crucial components of every email message, providing essential information about the email's origin, routing, and content. They are often hidden from the average email user but are invaluable for analysis and troubleshooting. Here are the key points related to email headers:

- 1. **Purpose of Email Headers:** Email headers serve the purpose of facilitating the smooth transmission of email messages from the sender to the recipient.
- 2. **Structure of Email Headers:** Email headers are structured as a set of key-value pairs. Each pair is on a separate line, and the header section is separated from the email body by a blank line.

- 3. **Common Header Fields:** Email headers consist of various fields, including "From," "To," "Subject," "Date," and "Message-ID." These fields provide information about the sender, recipient, subject, timestamp, and a unique identifier for the email.
- 4. **Received Headers:** One of the most critical sections of an email header is the "Received" field. It provides a trail of servers through which the email passed, helping trace the email's path and verify its authenticity.
- 5. Message Routing: Email headers contain information about how the message traveled from the sender's email client to the recipient's email server. This routing information is useful for diagnosing email delivery issues.

## 3.3 Types of Email Headers

There are several types of email headers, each serving a specific purpose:

- 1. **MIME Headers:** MIME (Multipurpose Internet Mail Extensions) headers are used to specify the type and structure of the email's content, allowing for the inclusion of multimedia elements like images and attachments.
- 2. **Authentication Headers:** Authentication headers, such as SPF (Sender Policy Framework) and DKIM (DomainKeys Identified Mail), are used to verify the authenticity of the sender and prevent email spoofing.
- 3. **User-Agent Headers:** User-Agent headers provide information about the email client or software used by the sender. This helps in identifying the source of the message.
- 4. **Custom Headers:** Email headers can also include custom fields created by the sender or email servers. These fields are specific to certain applications or organizations.

#### 3.4 Email Header Analysis

Email header analysis is a crucial skill in various fields, including cybersecurity, digital forensics, and troubleshooting email delivery issues. Here are key aspects of email header analysis:

- 1. **IP Address Tracking:** Analyzing email headers allows one to trace the path of the email through IP addresses, helping identify potential issues or suspicious activity.
- 2. **Email Authentication:** Authentication headers like SPF and DKIM can be scrutinized to verify the sender's authenticity, ensuring the email is not a forgery.
- 3. **Server Timestamps:** Email headers include timestamps from servers, which can be compared to identify delays or bottlenecks in email delivery.
- 4. **Identifying Email Spoofing:** By examining routing and authentication information, one can detect instances of email spoofing or phishing attempts.
- 5. **Legal and Compliance:** In legal matters, email headers can serve as evidence, proving the origin and path of an email message.

When analyzing email headers, it's essential to refer to email standards and protocols, such as RFC 5322, for a comprehensive understanding of the header fields and their meanings.

## 4 Platform

Operating System: Arch Linux x86-64

IDEs or Text Editors Used: Visual Studio Code

Compilers or Interpreters: Python 3.12

## 5 Analysis

### 5.1 Downloading Email Headers

```
[1]: username = "krishnaraj.kpt@outlook.com"
  password = "BBQtJTSs8uQh57aw"
  imap_server = "outlook.office365.com"
```

```
[5]: import imaplib
import email
from email.header import decode_header
from email.parser import BytesParser
import json
```

#### Creating Imap Object

```
[6]: # create an IMAP4 class with SSL
imap = imaplib.IMAP4_SSL(imap_server)
# authenticate
imap.login(username, password)
status, messages = imap.select("INBOX")

# number of top emails to fetch
N = 15

# total number of emails
messages = int(messages[0])
print("The Total Number of Messages in your account are: ", messages)
```

The Total Number of Messages in your account are: 1017

#### **Downloading Headers**

```
email_parser = BytesParser()
        email_header = email_parser.parsebytes(data[0][1])
         # create a dictionary to store the email header
         email = {
             "subject": email_header["Subject"],
             "headers": [],
        }
         # add the email headers to the dictionary
        for header in email_header.items():
             email["headers"].append({
                 "header": header[0],
                 "value": header[1],
             })
         # add the email to the list of emails
        emails.append(email)
# close the IMAP connection
imap.close()
# write the emails to a JSON file
with open("emails.json", "w") as f:
    json.dump(emails, f, indent=4)
Converting the json file to a python dictionary
```

```
[23]: import json
with open("emails.json", "r") as f:
    emails = json.load(f)
```

Looking at Header keys of one of the emails

```
[29]: header_names = [i['header'] for i in emails[1]["headers"]]
```

Focusing on 8 Test Emails

```
[24]: emails = emails[2:10]
```

Looking at their Subjects

```
[26]: [i['subject'] for i in emails]
```

'Mail sent from phone',

```
"Mother's phone Gmail app",
       'Sending from browser with vpn connection on']
     Analyse each header and find out what it means
[59]: print("Total number of headers in provided email: ", len(header_names))
     Total number of headers in provided email: 77
[45]: # print header names with their index
      for index, header in enumerate(header_names):
          print(index, header)
     0 MIME-Version
     1 Received
     2 ARC-Seal
     3 ARC-Message-Signature
     4 ARC-Authentication-Results
     5 Received
     6 Received
     7 Authentication-Results
     8 Received-SPF
     9 Received
     10 X-IncomingTopHeaderMarker
     11 ARC-Seal
     12 ARC-Message-Signature
     13 ARC-Authentication-Results
     14 DKIM-Signature
     15 Received
     16 Received
     17 From
     18 To
     19 Subject
     20 Thread-Topic
     21 Thread-Index
     22 Date
     23 Message-ID
     24 Accept-Language
     25 Content-Language
     26 X-MS-Has-Attach
     27 X-MS-TNEF-Correlator
     28 msip_labels
     29 x-tmn
     30 x-ms-traffictypediagnostic
     31 X-MS-Office365-Filtering-Correlation-Id
     32 X-Microsoft-Antispam-Untrusted
     33 X-Microsoft-Antispam-Message-Info-Original
     34 X-MS-Exchange-AntiSpam-MessageData-Original-ChunkCount
```

- 35 X-MS-Exchange-AntiSpam-MessageData-Original-0
- 36 Content-Type
- 37 X-MS-Exchange-Transport-CrossTenantHeadersStamped
- 38 X-IncomingHeaderCount
- 39 Return-Path
- 40 X-MS-Exchange-Organization-ExpirationStartTime
- 41 X-MS-Exchange-Organization-ExpirationStartTimeReason
- 42 X-MS-Exchange-Organization-ExpirationInterval
- 43 X-MS-Exchange-Organization-ExpirationIntervalReason
- 44 X-MS-Exchange-Organization-Network-Message-Id
- 45 X-EOPAttributedMessage
- 46 X-EOPTenantAttributedMessage
- 47 X-MS-Exchange-Organization-MessageDirectionality
- ${\tt 48\ X-MS-Exchange-Transport-CrossTenantHeadersStripped}$
- 49 X-MS-Exchange-Transport-CrossTenantHeadersPromoted
- 50 X-MS-PublicTrafficType
- 51 X-MS-Exchange-Organization-AuthSource
- 52 X-MS-Exchange-Organization-AuthAs
- 53 X-MS-UserLastLogonTime
- 54 X-MS-Office365-Filtering-Correlation-Id-Prvs
- 55 X-MS-Exchange-EOPDirect
- 56 X-Sender-IP
- 57 X-SID-PRA
- 58 X-SID-Result
- 59 X-MS-Exchange-Organization-PCL
- 60 X-MS-Exchange-Organization-SCL
- 61 X-Microsoft-Antispam
- 62 X-MS-Exchange-CrossTenant-OriginalArrivalTime
- 63 X-MS-Exchange-CrossTenant-Network-Message-Id
- 64 X-MS-Exchange-CrossTenant-Id
- 65 X-MS-Exchange-CrossTenant-RMS-PersistedConsumerOrg
- 66 X-MS-Exchange-CrossTenant-rms-persistedconsumerorg
- 67 X-MS-Exchange-CrossTenant-AuthSource
- 68 X-MS-Exchange-CrossTenant-AuthAs
- 69 X-MS-Exchange-CrossTenant-FromEntityHeader
- 70 X-MS-Exchange-Transport-CrossTenantHeadersStamped
- 71 X-MS-Exchange-Transport-EndToEndLatency
- 72 X-MS-Exchange-Processed-By-BccFoldering
- 73 X-Microsoft-Antispam-Mailbox-Delivery
- 74 X-Message-Info
- 75 X-Message-Delivery
- 76 X-Microsoft-Antispam-Message-Info

#### 5.1.1 MIME-Version

MIME-Version indicates the email's message format. It's important in investigations to understand how the message is structured and if it includes multimedia or attachments.

```
[38]: emails[1]['headers'][0]['header'], emails[1]['headers'][0]['value']

[38]: ('MIME-Version', '1.0')
```

#### 5.1.2 Received

Received headers track the path of the email through various servers. This is crucial for tracing the email's journey, identifying potential anomalies, or investigating its source.

```
[40]: emails[1]['headers'][1]['header'], emails[1]['headers'][1]['value']
```

Received from SJOPR17MB4837.namprd17.prod.outlook.com (2603:10b6:a03:37a::10) by DS7PR17MB6730.namprd17.prod.outlook.com with HTTPS; Sun, 29 Oct 2023 17:32:29 +0000

#### 5.1.3 ARC-Seal

ARC (Authenticated Received Chain) headers help verify the authenticity of email forwarding. ARC-Seal ensures the integrity of email headers, reducing the risk of spoofing.

```
[41]: emails[1]['headers'][2]['header'], emails[1]['headers'][2]['value']
```

[41]: ('ARC-Seal',

'i=2; a=rsa-sha256; s=arcselector9901; d=microsoft.com; cv=pass;\r\n b=UERZv4B1 U4weLVUXdoohvbjGYpYf4pb9pFtOwayMK+mtwTbjMhWrqskYHiEhqgH1rrxhnvYgrK7YkvSkKXiEypdA OakOf4KcLaHNb/KEBpCVvQoKhVUX2zWzFMxVLsIRkMgoltrKRs0JShcFwrbt6XCvxCZUTbsGQs/hFpaN OsYlFys1Qu41etiVDrmS8ZYpq4ZnHuXxdBzxW6A8Aql06f5sr4CF2fSeAIjAFu5JB5/tTHlu9wFIYa49 rmvL4i2S8QLH18IdHsvPpz0oNrK0BVzf8bFat2iF7qnIX2J1lkXo21nWrGEioqJjPP6uusrYJBc5R+Sa tB5i3rmSYpqn0w==')

#### 5.1.4 ARC-Message-Signature

ARC-Message-Signature is part of ARC headers and provides cryptographic assurance of email header integrity, aiding in detecting email tampering.

r9Wu/Vx5h/kN9LqqjvwNmvzTLuRE1oPLaIch8T2dBnIf6leWeKENffAjFK+Kba43a68mbAnbxlvsg=='

#### 5.1.5 ARC-Authentication-Results

These headers indicate the email's authentication status. Investigators can use this to assess the email's legitimacy and potential for phishing.

#### 5.1.6 DKIM-Signature

DKIM (DomainKeys Identified Mail) verifies that the email content hasn't been altered in transit. It's crucial for email integrity checks and source verification.

#### 5.1.7 From

The "From" header shows the sender's email address. It's essential for identifying the sender, although it can be spoofed.

```
[47]: emails[1]['headers'][17]['header'], emails[1]['headers'][17]['value']

[47]: ('From', 'Krishnaraj Thadesar <Krishnaraj.kpt@outlook.com>')
```

#### 5.1.8 To

The "To" header reveals the email's recipient, which is significant for understanding the email's target and potential threat actors.

```
[48]: emails[1]['headers'][18]['header'], emails[1]['headers'][18]['value']

[48]: ('To', 'Krishnaraj Thadesar <krishnaraj.kpt@outlook.com>')
```

#### 5.1.9 Subject

The subject line provides insight into the email's content, which is crucial for assessing the email's purpose and relevance to an investigation.

```
[49]: emails[1]['headers'][19]['header'], emails[1]['headers'][19]['value']

[49]: ('Subject', 'Self mail from web outlook client')
```

#### **5.1.10** Date

The date header shows when the email was sent. It's valuable for establishing timelines and correlations with other events.

```
[50]: emails[1]['headers'][22]['header'], emails[1]['headers'][22]['value']

[50]: ('Date', 'Sun, 29 Oct 2023 17:32:25 +0000')
```

#### 5.1.11 Message-ID

The Message-ID is unique to each email and can be used for tracking and associating related messages in an investigation.

#### 5.1.12 Content-Type

Content-Type specifies the format of the email content. It helps investigators interpret the email's structure and potential for malicious attachments.

#### 5.1.13 Return-Path

Return-Path indicates where undeliverable emails should be sent. It can assist in identifying email redirection or bouncing patterns.

```
[54]: emails[1]['headers'][39]['header'], emails[1]['headers'][39]['value']

[54]: ('Return-Path', 'krishnaraj.kpt@outlook.com')
```

#### 5.1.14 X-Sender-IP

This header contains the IP address of the sender, which is essential for tracking the origin of the email and potential geolocation.

```
[55]: emails[1]['headers'][56]['header'], emails[1]['headers'][56]['value']

[55]: ('X-Sender-IP', '40.92.20.10')
```

#### 5.1.15 X-MS-Exchange-Transport-EndToEndLatency

End-to-end latency is crucial for assessing the email's delivery speed, which might reveal anomalies or delays in transit.

```
[56]: emails[1]['headers'][71]['header'], emails[1]['headers'][71]['value']

[56]: ('X-MS-Exchange-Transport-EndToEndLatency', '00:00:02.4306131')
```

#### 5.1.16 X-Microsoft-Antispam-Mailbox-Delivery

This header provides information about the email's delivery and its classification as spam or not, aiding in filtering and threat analysis.

#### 5.1.17 X-Message-Info and X-Message-Delivery

These headers contain miscellaneous information about the email's handling, which can be valuable for tracking and understanding the email's journey.

```
[58]: emails[1]['headers'][75]['header'], emails[1]['headers'][75]['value']

[58]: ('X-Message-Delivery', 'Vj0xLjE7dXM9MDtsPTA7YT0x00Q9MTtHRD0x01NDTD0tMQ==')
```

```
Defining a function to get information about an ip address for sender ip analysis
```

```
def get_ip_information(ip_address):
    def get_ip_location():
        # Make a GET request to ipinfo.io with the IP address
        url = f"https://ipinfo.io/{ip_address}/json"
        response = requests.get(url)
```

```
if response.status_code == 200:
        data = response.json()
        return data
    else:
        return None
location_info = get_ip_location()
# print(location_info)
if location_info:
    # Print the location information
    print(f"IP Address: {location_info['ip']}")
    print(f"Hostname: {location_info['hostname']}")
    print(f"City: {location_info['city']}")
    # print(f"Region: {location_info['region']}")
    print(f"Country: {location_info['country']}")
    print(f"Location: {location_info['loc']}")
    # print(f"Organization: {location_info['org']}")
    print(f"Timezone: {location_info['timezone']}")
else:
    print("Unable to retrieve location information for the IP address.")
```

#### Analysing Sender IP Address for all emails

```
[66]: for email in emails:
          for header in email['headers']:
              if header['header'] == 'X-Sender-IP':
                  print("Subject of Email: ", email['subject'])
                  print("IP Address Information:")
                  get_ip_information(header['value'])
                  print()
     Subject of Email: Sending from phone gmail app
     IP Address Information:
     IP Address: 209.85.218.48
     Hostname: mail-ej1-f48.google.com
     City: Oudeschip
     Country: NL
     Location: 53.4300,6.8264
     Timezone: Europe/Amsterdam
     Subject of Email: Self mail from web outlook client
     IP Address Information:
     IP Address: 40.92.20.10
     Hostname: mail-bn8nam110lkn2010.outbound.protection.outlook.com
     City: Boydton
     Country: US
     Location: 36.6676,-78.3875
```

Timezone: America/New\_York

Subject of Email: Self test mail from gmail web client

IP Address Information:
IP Address: 209.85.167.178

Hostname: mail-oi1-f178.google.com

City: Tulsa Country: US

Location: 36.1540,-95.9928 Timezone: America/Chicago

Subject of Email: self test mail from mit gmail client

IP Address Information:
IP Address: 209.85.219.47

Hostname: mail-qv1-f47.google.com

City: Raleigh Country: US

Location: 35.7721,-78.6386 Timezone: America/New\_York

Subject of Email: From movile outlook app

IP Address Information: IP Address: 40.92.242.29

Hostname: mail-ps2kor01olkn2029.outbound.protection.outlook.com

City: Busan Country: KR

Location: 35.1017,129.0300

Timezone: Asia/Seoul

Subject of Email: Mail sent from phone

IP Address Information:
IP Address: 209.85.160.175

Hostname: mail-qt1-f175.google.com

City: Charlotte Country: US

Location: 35.2271,-80.8431 Timezone: America/New\_York

Subject of Email: Mother's phone Gmail app

IP Address Information:
IP Address: 209.85.160.171

Hostname: mail-qt1-f171.google.com

City: Charlotte Country: US

Location: 35.2271,-80.8431 Timezone: America/New\_York

Subject of Email: Sending from browser with vpn connection on

IP Address Information:
IP Address: 209.85.208.52

Hostname: mail-ed1-f52.google.com

City: Oudeschip Country: NL

Location: 53.4300,6.8264 Timezone: Europe/Amsterdam

## 6 Analysing Effect of Network Connection on IP Headers

## 6.1 Using VPN

Emails were sent by "parthzarekar3@outlook.com" to "krishnaraj.kpt@outlook.com" with connection to a VPN.

The IP Information was taken from the headers.

```
{
    "header": "X-Sender-IP",
    "value": "209.85.208.52"
},
```

From This the IP Address was taken and searched on iplookup.com

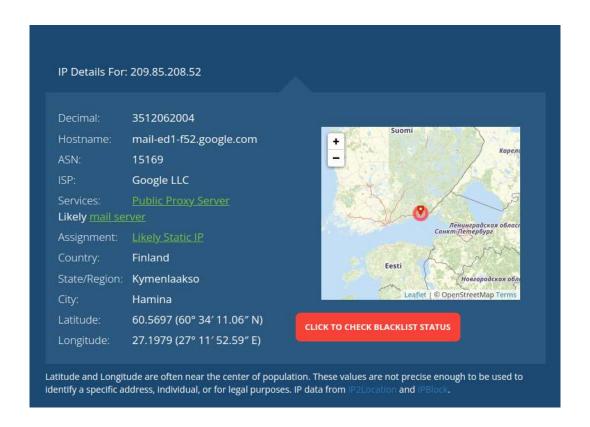
The IP Information was:

IP Address Information: IP Address: 209.85.208.52

Hostname: mail-ed1-f52.google.com

City: Oudeschip Country: NL

Location: 53.4300,6.8264 Timezone: Europe/Amsterdam



## 6.2 Without using VPN, on Wifi

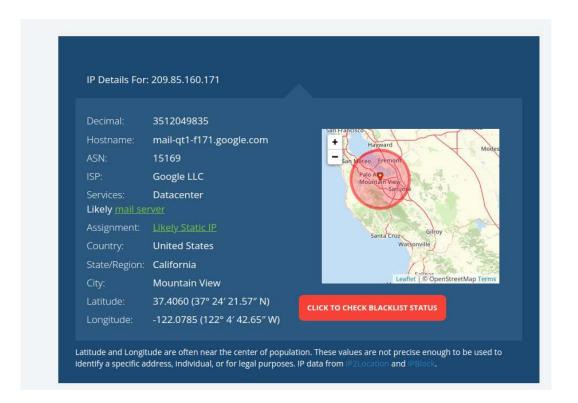
```
{
    "header": "X-Sender-IP",
    "value": "209.85.160.171"
},
```

From This the IP Address was taken and searched on iplookup.com

#### The IP Information was:

```
Subject of Email: Mother's phone Gmail app
IP Address Information:
IP Address: 209.85.160.171
Hostname: mail-qt1-f171.google.com
City: Charlotte
Country: US
Location: 35.2271,-80.8431
```

Timezone: America/New\\_York



## 6.3 Without using VPN, on Mobile Data

```
{
    "header": "X-Sender-IP",
    "value": "209.85.208.41"
},
```

From This the IP Address was taken and searched on iplookup.com

### The IP Information was:

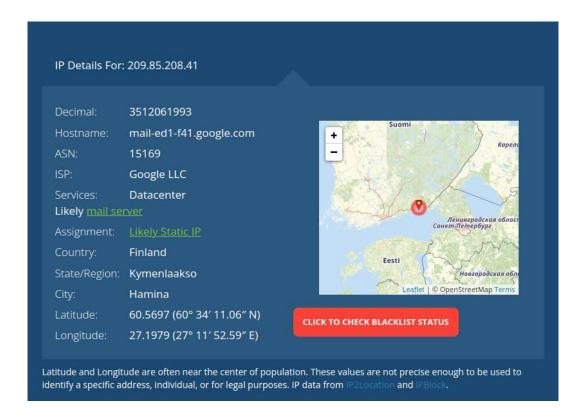
```
Subject of Email: Sending from browser using mobile hotspot
```

IP Address Information: IP Address: 209.85.208.41

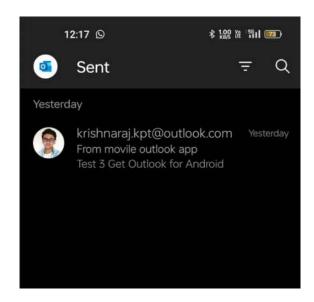
Hostname: mail-ed1-f41.google.com

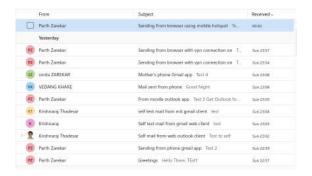
City: Oudeschip Country: NL

Location: 53.4300,6.8264 Timezone: Europe/Amsterdam









## 7 Conclusion

We have explored how to analyze email headers and understand their significance:

- 1. Email headers contain vital information about the email's origin, routing, and content.
- 2. Analyzing email headers is crucial for diagnosing email delivery issues, ensuring authenticity, and enhancing cybersecurity.
- 3. Key header fields include "From," "To," "Received," and "Message-ID."

We have also learned about email routing:

- 1. Regardless of the user's connection method (VPN, Mobile Data, or WiFi), emails are routed through servers.
- 2. Standard email protocol involves servers sending emails from one server to another.

Additionally, it was discovered that Outlook has security vulnerabilities:

- 1. Outlook's security, especially regarding password protection, has been found lacking.
- 2. Even after enabling 2 Factor Authentication (2FA), a simple Python script using the imaplib package can extract emails with just the password.

## References

- [1] Internet Message Format Internet Message Format
- [2] Email Headers
  "Email Headers: What They Are and How to Analyze Them" by Email on Acid
- [3] How to Read Email Headers
  MX Toolbox