CSEC 730 = Advanced Computer Forensics

Email Forensics - Report

Extracting hidden data from email headers

- Ashwin Ramesh, Hans Johnson, Ryan Whittier

Overview

The tool looks at raw emails with headers and verifies whether headers are valid (registered with IANA) pulling out any headers that are not. The tool also looks at headers to find out if they are formatted properly. Some further checks it does are whether or not some email defences are valid such as SPF and DKIM.

Introduction

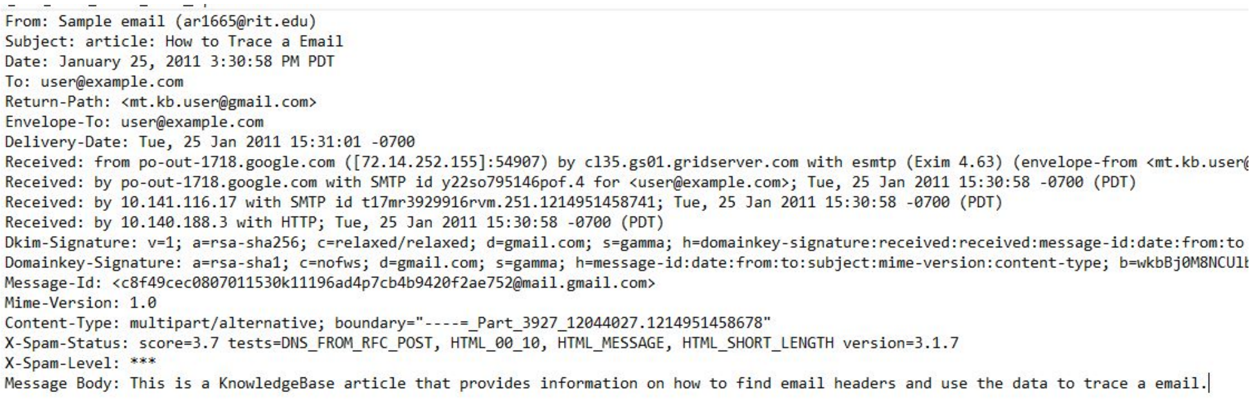
The header is a portion of the email message that remains invisible to most people. However, it contains a significant amount of metadata which can be useful for forensic analysis.

Breaking down headers

Figure 1 is a sample email header which we can breakdown and analyse. It should be noted that when reading an email header every line can be forged, therefore only the Received lines generated by the user's service or company can be completely trusted.

- FROM : Displays who the message is from, however this header can be easily forged and is the least reliable
- SUBJECT : Created by the sender placed as a topic of the email body.
- DATE : The date and time when the email was composed.
- TO : Shows to whom the email was addressed, but typically may not contain the recipient's address
- RETURN-PATH : The email address for return mail, also known as "Reply-To:".

- ENVELOPE-TO : This header shows that this email was delivered to the mailbox of a subscriber whose email address is user@example.com



```
From: Sample email (ar1665@rit.edu)
Subject: article: How to Trace a Email
Date: January 25, 2011 3:30:58 PM PDT
To: user@example.com
Return-Path: <mt.kb.user@gmail.com>
Envelope-To: user@example.com
Delivery-Date: Tue, 25 Jan 2011 15:31:01 -0700
Received: from po-out-1718.google.com ([72.14.252.155]:54907) by cl35.gs01.gridserver.com with esmtp (Exim 4.63) (envelope-from <mt.kb.user@
Received: by po-out-1718.google.com with SMTP id y22so795146pof.4 for <user@example.com>; Tue, 25 Jan 2011 15:30:58 -0700 (PDT)
Received: by 10.141.116.17 with SMTP id t17mr3929916rvm.251.1214951458741; Tue, 25 Jan 2011 15:30:58 -0700 (PDT)
Received: by 10.140.188.3 with HTTP; Tue, 25 Jan 2011 15:30:58 -0700 (PDT)
Dkim-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed; d=gmail.com; s=gamma; h=domainkey-signature:received:received:message-id:date:from:to
Domainkey-Signature: a=rsa-sha1; c=nofws; d=gmail.com; s=gamma; h=message-id:date:from:to:subject:mime-version:content-type; b=wkbBj0M8NCUl
Message-Id: <c8f49cec0807011530k11196ad4p7cb4b9420f2ae752@mail.gmail.com>
Mime-Version: 1.0
Content-Type: multipart/alternative; boundary="----=_Part_3927_12044027.1214951458678"
X-Spam-Status: score=3.7 tests=DNS_FROM_RFC_POST, HTML_00_10, HTML_MESSAGE, HTML_SHORT_LENGTH version=3.1.7
X-Spam-Level: ***
Message Body: This is a KnowledgeBase article that provides information on how to find email headers and use the data to trace a email.
```

Fig 1. Sample email headers generated for a simple email message[5].

- DELIVERY-DATE : This shows the date and time at which the email was received by the user's service or email client.
- RECEIVED : Considered the most important part of the email header and is usually the most reliable. It generates a list of all the servers through which the messaged hopped across to reach the recipient. This header is generally best read bottom to top, the first line being the recipient's system or mail server, and the last line being the origin of the email.
- DKIM SIGNATURE : An email authentication method designed to detect email spoofing.
- MESSAGE-ID : A unique string assigned by the mail system when the mail is initially created. These can be easily forged.
- MIME-VERSION : Multipurpose Internet Mail Extensions (MIME) is an internet standard that extends the format of emails.
- CONTENT-TYPE : Details on the format of the message such as html or plaintext
- X-SPAM-STATUS : Shows a spam score created by the email service or mail client.
- X-SPAM-LEVEL : Shows a spam score created by the email service or mail client.
- MESSAGE BODY : Content of the email, written by the sender.

SPF, DKIM and DMARC

SPF (Sender policy Framework) is a DNS text entry which lists servers that that are considered safe for forwarding mail for a specific domain. This list is authoritative for the domain, since only the administrators are allowed to add/change the main domain zone.

DKIM (DomainKeys Identified Mail) alternatively is a method to verify that the messages' body/content are trustworthy, i.e was not changed from the moment the email left the initial mail server. This is achieved by implementing a standard public/private key signing scheme. The administrators of the domain add a DNS entry with the public DKIM key which is used by the recipients to verify that the message is correct.

DMARC (Domain-based Message Authentication, Reporting and Conformance) in addition with SPF and DKIM is used to state a clear policy which should be followed, it also sets an address which can be used to send reports about the mail message statistics gained by recipients against a specific domain.

SPF

- Upon receipt, the HELO message and the sender address are fetched by the receiving mail server.
- The mail server runs a TXT DNS query against the claimed domain SPF entry
- The SPF data is then used to verify the sender's server.
- If the check fails a rejection message is given to the sender's server.

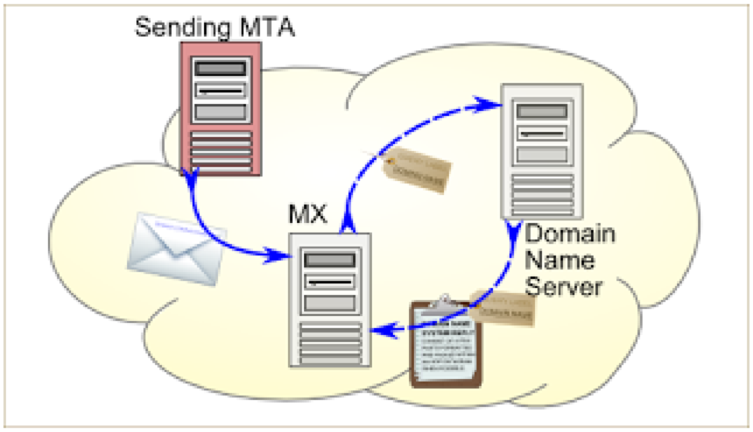


Fig 2. SPF authenticating the sender's IP address [2]

DKIM

- A new header (DKIM-SIGNATURE) is added to the email by using the private key of the domain on the email content.
- The email content cannot be modified, otherwise the DKIM header will not match.
- Upon receipt the receiving server will make a TXT DNS query to retrieve the key used in the DKIM-Signature field.
- The DKIM header check result can now be used to check if the content has been spoofed.

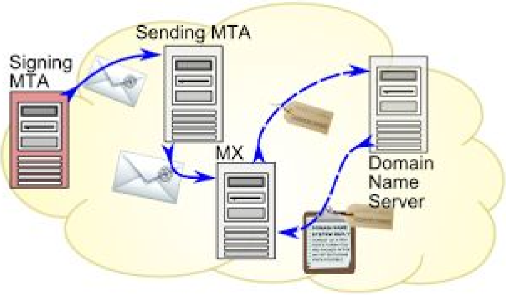


Fig 3. DKIM authentication of parts of the message content[2].

DMARC

- Upon receiving the email the receiving mail server check if there is an existing DMARC policy published in the domain used by the SPF/DKIM checks.
- If one or both of the SPF and DKIM checks succeeds and is in accordance with the policy set by DMARC, then the check is successful, otherwise it failed.
- If the check fails, an action based on the DMARC policy is taken.

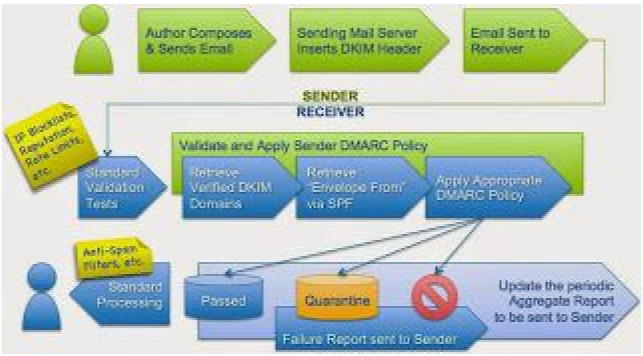

Author Composes & Sends Email
Sending Mail Server Inserts DKIM Header
Email Sent to Receiver
SENDER
RECEIVER
IP Blocklists, Reputation, Rate Limits, etc.
Standard Validation Tests
Validate and Apply Sender DMARC Policy
Retrieve Verified DKIM Domains
Retrieve "Envelope From" via SPF
Apply Appropriate DMARC Policy
Anti-Spam Filters, etc.
Standard Processing
Passed
Quarantine
Update the periodic Aggregate Report to be sent to Sender
Failure Report sent to Sender

Fig 4. A sample DMARC policy framework[1].

Related Tools

All other tools only parse the headers to make them more human readable. For example this tool “https://mxtoolbox.com/EmailHeaders.aspx” only reads the header and creates a table of the header/value pairs and the hops of email servers and tells whether the IPs were blacklisted or not. Whereas our tool will parse email headers and find non-compliant headers for further investigation.

Our Tool

The significance of this tool in the forensics field is that it can help you identify anomalies in email headers that may have been overlooked or not seen at all. These headers can be indicative of attacks or abuse.

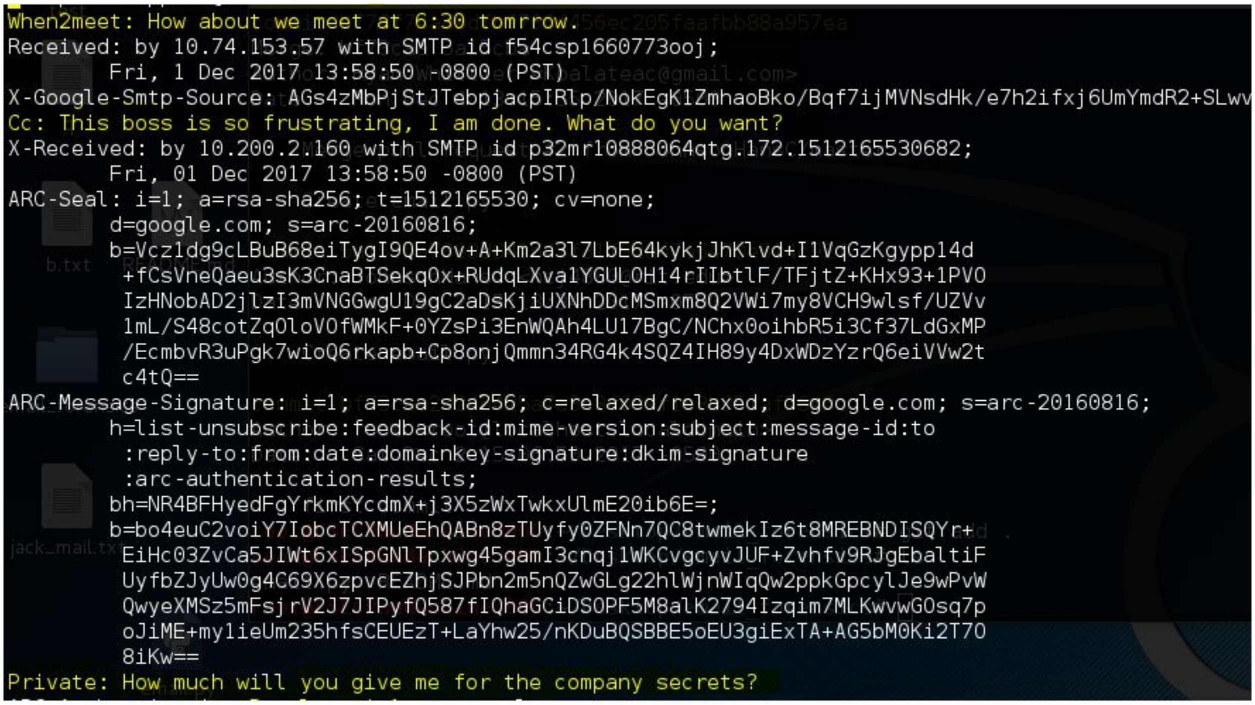
How to use this tool:

1. Obtain the original message file. In gmail select “Show Original” from the drop down in the top right of the email and download.
2. Run the python file and pass in the message file.
3. Observe output.

The necessary files can be downloaded from : https://github.com/KoalaTea/email_forensics

The CASE

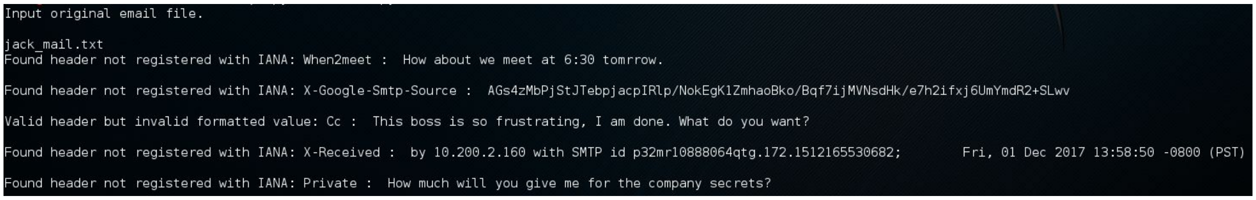
Jack from ACME is frustrated with his boss making him work weekends and not recommending him for a promotion. Jack since has decided to plot against his employers. He recently figured out that he can edit invisible email headers, and he can use that to contact his friend working at Belkin.



```
When2meet: How about we meet at 6:30 tomrrow.
Received: by 10.74.153.57 with SMTP id f54csp1660773ooj;
        Fri, 1 Dec 2017 13:58:50 -0800 (PST)
X-Google-Smtp-Source: AGs4zMbPjStJTebpjacpIRlp/NokEgK1ZmhaoBko/Bqf7ijMVNsdHk/e7h2ifxj6UmYmdR2+SLwv
Cc: This boss is so frustrating, I am done. What do you want?
X-Received: by 10.200.2.160 with SMTP id p32mr10888064qtg.172.1512165530682;
        Fri, 01 Dec 2017 13:58:50 -0800 (PST)
ARC-Seal: i=1; a=rsa-sha256; t=1512165530; cv=none;
        d=google.com; s=arc-20160816;
        b=Vcz1dg9cLBuB68eiTygI9QE4ov+A+Km2a3l7LbE64kykjJhKlvd+I1VqGzKgypp14d
         +fCsVneQaeu3sK3CnaBTSekqOx+RUdqLXva1YGULOH14rIIbtlF/TFjtZ+KHx93+1PVO
         IzHNobAD2jlzI3mVNGGwgU19gC2aDsKjiUXNhDDcMSmxm8Q2VWi7my8VCH9wlsf/UZVv
         1mL/S48cotZqOloVOfWMkF+0YZsPi3EnWQAh4LU17BgC/NChx0oihbR5i3Cf37LdGxMP
         /EcmbvR3uPgk7wioQ6rkapb+Cp8onjQmmn34RG4k4SQZ4IH89y4DxWDzYzrQ6eiVVw2t
         c4tQ==
ARC-Message-Signature: i=1; a=rsa-sha256; c=relaxed/relaxed; d=google.com; s=arc-20160816;
        h=list-unsubscribe:feedback-id:mime-version:subject:message-id:to
         :reply-to:from:date:domainkey-signature:dkim-signature
         :arc-authentication-results;
        bh=NR4BFHyedFgYrkmKYcdmX+j3X5zWxTwkxUlmE20ib6E=;
        b=bo4euC2voiY7IobcTCXMUeEhQABn8zTUyfy0ZFNn7QC8twmekIz6t8MREBNDISQYr+
         EiHc03ZvCa5JIWt6xISpGNlTpxwg45gamI3cnqj1WKCvgcyvJUF+Zvhfv9RJgEbaltiF
         UyfbZJyUw0g4C69X6zpvcEZhjSJPbn2m5nQZwGLg22hlWjnWIqQw2ppkGpcylJe9wPvW
         QwyeXMSz5mFsjrV2J7JIPyfQ587fIQhaGCiDSOPF5M8alK2794Izqim7MLKwvwG0sq7p
         oJiME+my1ieUm235hfsCEUEzT+LaYhw25/nKDuBQSBBE5oEU3giExTA+AG5bM0Ki2T70
         8iKw==
Private: How much will you give me for the company secrets?
```

Fig 5 Sample of Jack's outgoing email including all headers.

However, ACME screens all their outgoing emails. Incidentally Jack's email was flagged because some of his email headers were not compliant. Upon investigation it was found that Jack was sending secret information by hiding it in his email headers and severe action was taken.



```
Input original email file.

jack_mail.txt
Found header not registered with IANA: When2meet :  How about we meet at 6:30 tomrrow.

Found header not registered with IANA: X-Google-Smtp-Source :  AGs4zMbPjStJTebpjacpIRlp/NokEgK1ZmhaoBko/Bqf7ijMVNsdHk/e7h2ifxj6UmYmdR2+SLwv

Valid header but invalid formatted value: Cc :  This boss is so frustrating, I am done. What do you want?

Found header not registered with IANA: X-Received :  by 10.200.2.160 with SMTP id p32mr10888064qtg.172.1512165530682;        Fri, 01 Dec 2017 13:58:50 -0800 (PST)

Found header not registered with IANA: Private :  How much will you give me for the company secrets?
```

Fig 6 Flags generated by our tool post reaching the email domain server.

Limitations

The limitations of this tool are:

- It does not look at the email content
- It is not a command line tool with flags for different options
- It does not validate all of the registered header values format
    - For example, it does not check that the value of header Subject is formatted correctly
- It cannot look at multiple emails and correlate data
- Does not support a pst file from outlook

References

1. https://dmarc.org/overview/
2. https://en.wikipedia.org/wiki/Email_authentication
3. https://www.endpoint.com/blog/2014/04/15/spf-dkim-and-dmarc-brief-explanation
4. https://en.wikipedia.org/wiki/DomainKeys_Identified_Mail
5. https://mediatemple.net/community/products/dv/204643950/understanding-an-email-header
6. http://www.makeuseof.com/tag/what-can-you-learn-from-an-email-header-metadata/