You Have Mail: Email deception and how to detect it.[1]

Tim Watson

How vulnerable email protocols can be abused and how to catch those who do it.

Email started life as a novelty and has risen to become a necessity. But the speed, flexibility and low costs of email communication have been turned into a weapon. From spam to spear phishing, your inbox can place you one click away from disaster. In fact, you don't even need to click to be in danger. How can you tell the good from the bad, the genuine from the fake? How is a deceptive email constructed and how can it be spotted? Let's find out.

As with any form of defence, knowledge is power. The main weakness exploited by those who send malicious emails is the weakness of ignorance. The fact that the vast majority of users do not have a clue how emails work, how they are constructed and how they get from source to destination is both a credit to the design of the email system, which provides a simple and reliable communication method with no need for the user to understand the mechanisms used, and an opportunity for those who do understand the system to perform nefarious, electronic sleight of hand to deceive the trusting masses of email users who embrace its magic.

To understand the dangers and the ways to reduce them, we need to peek behind the curtains and discover the secrets of the processes and protocols that make up the modern email system. By understanding how emails work, we will be able to spot the weak points and to discover the trail of clues left by those who seek to abuse the system for their own advantage. We will start by following the typical journey of an email from composition to the point at which it is read at its destination.

In simple terms, an email is composed in a mail client such as Mozilla Thunderbird or Outlook Express, sent to a mail server (e.g. Sendmail), which then forwards it through other mail servers until it reaches the destination mail server. To be precise, if the sender and receiver use the same mail server then there will only be one mail server involved and if the email is sent to diverse recipients then there will be several destination mail servers. After the email has arrived, the recipient can use a mail client to download and read the email. If you explore the various standards and documentation relating to email you will discover that there are further components defined, such as mail submission agents , mail delivery agents and mail access agents. You'll also see that clients are often called mail user agents (MUAs) and that mail servers are called mail transport agents (MTAs) (see Figure 1).

For the purposes of this article, we need to explore the format of emails, the client and server programs that process them and the protocols used to transport them. There is also another area that provides an attacker with a wealth of opportunities and that is HTML, commonly found within emails and often used to mislead and compromise victims but, since the topic is vast and not specific to emails, it will not be covered here. The interested reader is directed to the many resources on the Web to do with Web-based attacks, drive-by downloads, cross-site scripting etc. I have to admit that there is a certain, delicious irony in directing readers to HTML pages to discover more about HTML attacks.

---

1 *This article originally appeared in Digital Forensics Magazine and is reproduced with the kind permission of the publishers.*

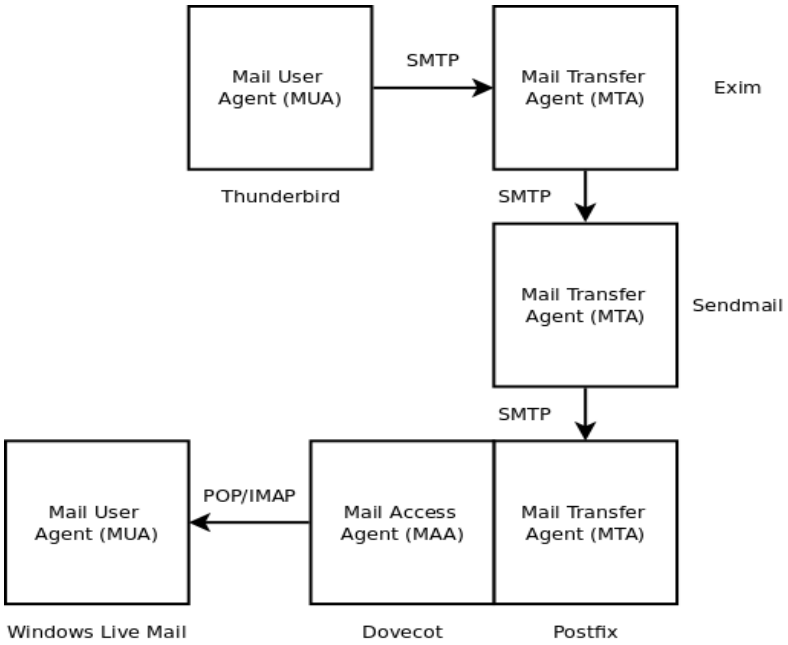

Figure 1. Delivering an email

As well as looking at how attackers can exploit emails to deceive victims, we are also interested in how to detect their deception and how to determine the identity of the attacker. Again, the limitations of space prevent us from covering a number of useful avenues of investigation. These include the various attribution techniques that rely on the details contained in the network packets associated with sending and receiving emails and the evidence contained in the machines running mail servers. Our investigation will be based solely on the information available from an email retrieved by a mail client.

Email Message Format

An email message is contained in an envelope. The envelope is defined in the RFC 5321 document that describes the Simple Mail Transfer Protocol (SMTP) and, just like a standard mail envelope, it tells the mail system where to deliver it. We'll look more closely at the envelope later but for now we will concentrate on the message itself.

A typical email, as viewed by a user, is shown in Figure 2. The mail client shows which mail folder is being viewed, a list of email subject lines, usually in date order, and a preview pane that displays the contents of the currently selected email. However, this is often only a selected part of the email. The actual email source can be viewed (using CTRL+U or choosing 'view message source' in a menu) and doing so will reveal the full email as received by the mail client. RFC 5322 and RFC 2045 together provide an authoritative description of the format of an email message.

If we look at the simplified source of the email from Figure 2 (see Appendix 1) we can see that it is made up of different logical sections. The overall message is divided into two: first come the email headers and then, after a blank line, is the message body. The body itself is typically divided into parts: a plaintext version, an HTML version and any attachments. You can see that one of the headers defines the string of characters "b1_5fc6d29ab..." that will be used in this email as a boundary to separate the different parts of the body. Each part is preceded by a blank line and then a line that starts with two dashes followed by the boundary string. The mail client understands this message format and uses the headers to show who sent the email, its subject and when it was sent, and it chooses whether to display

the plaintext or HTML version of the contents, depending on the preference of the user. It may be that the user gets no choice and only gets to see the HTML version.

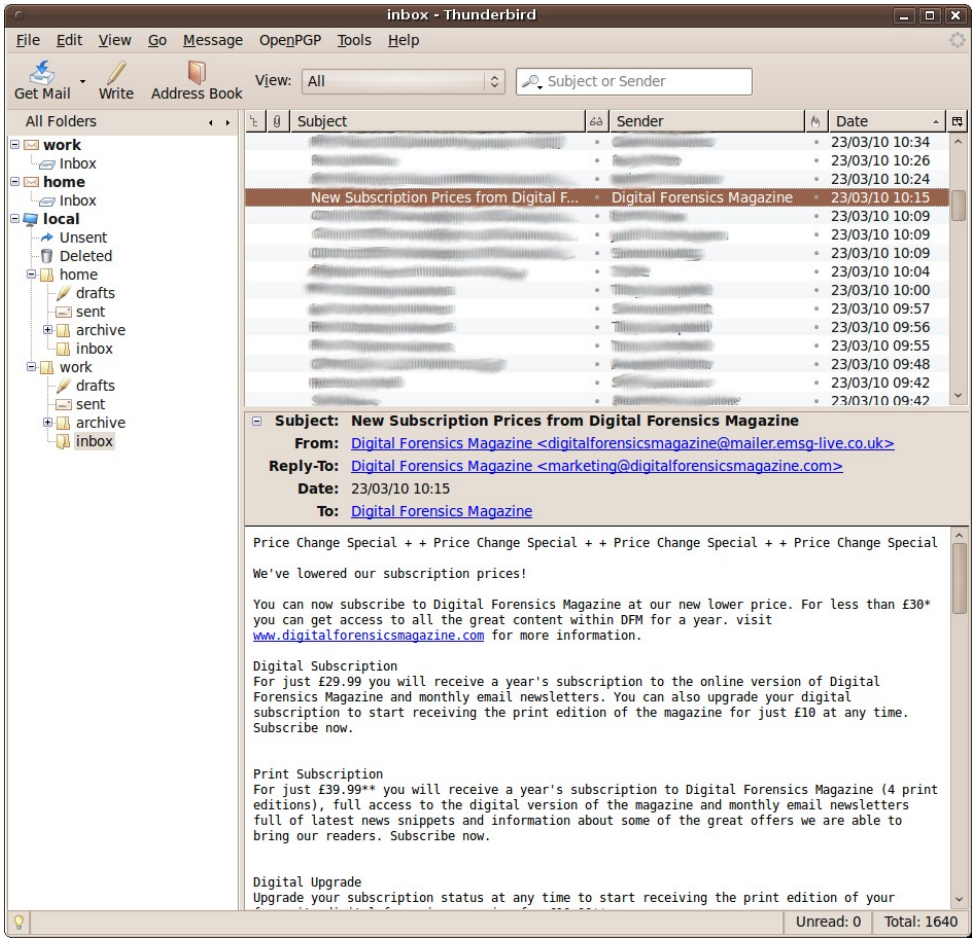

Figure 2. A typical email client

This is where we get our first surprise. All of this information is completely under the control of the sender. While the details of the sender and recipient are 'on' the envelope, the envelope is removed by the mail server and it is only the contents of the message that are sent to the user receiving the email. The message headers that state who the email is from, who it's to, when it was sent and the entire message body can all be made up by the sender and do not have to relate to the information on the envelope. I can construct an envelope to your email address today but when you receive the email I can make it appear that it was sent to anyone I like, from anyone I like, at any date I like and with any contents I like. We'll see how to do this shortly. For now, it is enough to worry that the weakness in the global email system just revealed means that you can never trust another email unless you view the source. Oh, and you'd also better worry about someone sending a forged email to your boss, or your partner, that appears to come from you.

While the subject of HTML-based attacks is beyond the scope of this article, it is worth noting that the email shown in Appendix 1 contains a common, hidden extra. If you look closely you'll see that the plaintext section ends with the words, “Add us to your contact list to make sure you can receive future emails safely”, whereas the HTML version has an extra bit of code after this text, as follows:

This code tells your email client to load an image from the “clicks.emsg-live.co.uk” website. It is used to track details about the email when it is read. Even if you don't click on the email, if it is displayed in your preview pane as an HTML message it will retrieve this image. When it does so, the website you have just accessed can log details such as the date and time of access, the IP address of the source of the request and the HTTP request headers that show details about your computer's software. With a different image included in each email, this allows the sender to monitor each recipient every time they view the email. Unless, like me, you don't open the HTML version of emails. In the hands of a malicious sender, this ability to make the receiver automatically access an arbitrary webserver and download an image of the attacker's choosing is obviously very dangerous. Of course, if you are communicating with a suspected criminal by email, the same technique can be used to help trace them.

We will return to the message content when we explore the message headers but, for now, we need to understand how to construct and how to send a forged email.

Simple Mail Transfer Protocol

When your email client sends an email it does so by communicating with a mail server using the Simple Mail Transfer Protocol (SMTP). The details of this protocol can be found in RFC 5321. Although it is recommended that mail user agents don't talk to mail transfer agents directly, but rather that they use a mail submission agent as described in RFC 4409, both MTAs and MSAs use the same SMTP protocol and it is still normal for mail clients to talk directly to MTAs.

Your mail client typically connects to the mail server using TCP port 25 and receives an identification message from the server. The client then says hello and the server responds with a list of services available. The client will then typically send the envelope details saying where to send the email and who it's from, and then the email data is transferred from client to server. This data includes the message headers and the message content and it is all treated as undifferentiated data by the mail server. Your mail client will include several headers in the email to show which mail client software you are using, your sender's email address, the date and time etc.

However, we don't have to use a conventional mail client. If we use a low-level network tool such as netcat, we can directly control the information passed to the server. The command-line output in Appendix 2 is a dialogue between my “attacker's machine” and a mail server and Figure 3 shows the email as it appears to the receiver after it has been sent. I have highlighted in bold the parts of this dialogue that were typed in by me, the rest is produced by the server.

You'll notice that there was no authentication needed. As long as I'm accessing the mail server from the same domain it will happily accept commands from me. You'll also notice that it was perfectly comfortable accepting an obviously forged “MAIL FROM:” command and that it didn't attempt to check for consistency between the envelope and the message headers. This isn't some badly configured home-user mail server, this is a real mail server in a large organisation and is typical of most current mail servers in use today.

So how do we hope to spot when we are being deceived by email? The answer is in the headers.

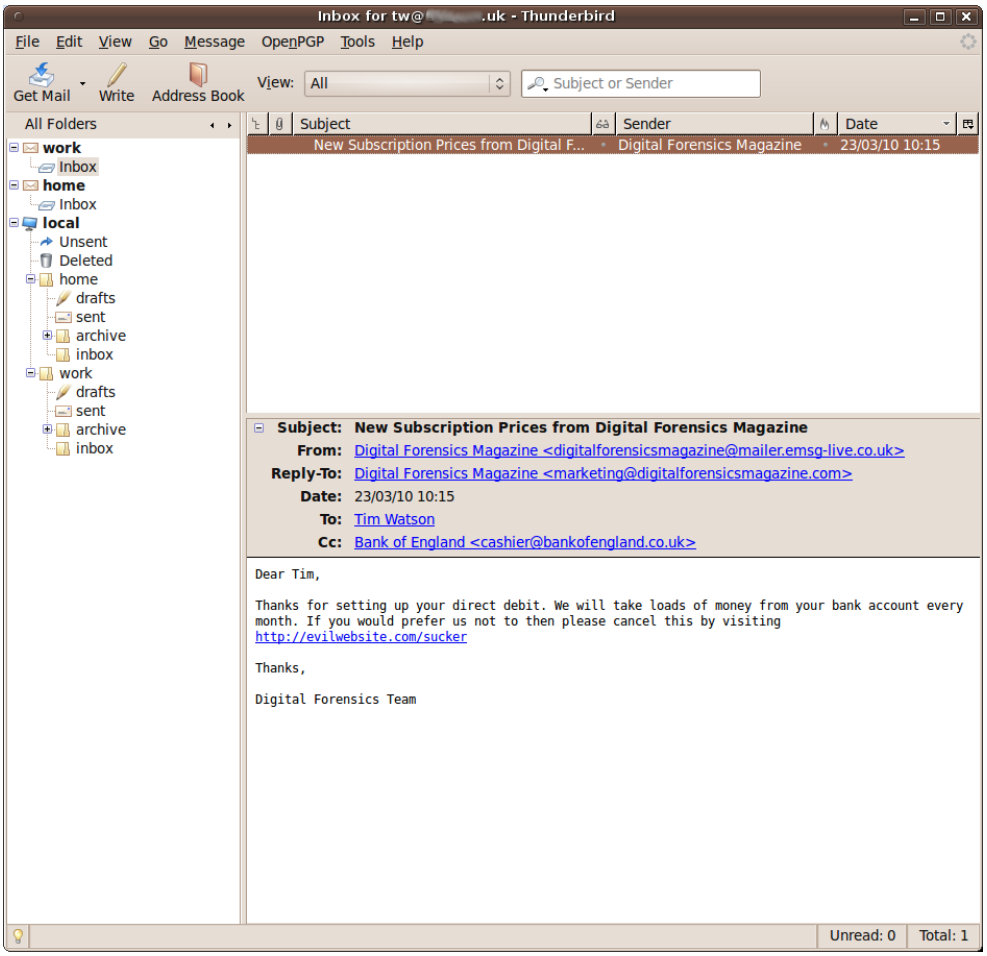

Figure 3. A malicious email as seen by the receiver

Email Headers

If we look at the message source of our forged email and focus on the headers we can see that something isn't right:

From - Mon Apr 05 20:42:30 2010
X-Mozilla-Status: 0001
X-Mozilla-Status2: 00000000
Return-Path: <digitalforensicsmagazine@mailer.emsg-live.co.uk>
Received: from mailer.emsg-live.co.uk (me.mydomain.co.uk [146.XXX.XX.XXX])
by mail.mydomain.co.uk (8.13.6/8.13.6) with ESMTP id o35JZFxs008825
for <tw@victim.co.uk>; Mon, 5 Apr 2010 20:35:32 +0100 (BST)
From: "Digital Forensics Magazine" <digitalforensicsmagazine@mailer.emsg-live.co.uk>
To: Tim Watson <tw@victim.co.uk>
Cc: Bank of England <cashier@bankofengland.co.uk>
Date: Tue, 23 Mar 2010 10:15:10 +0000
Sender: digitalforensicsmagazine@mailer.emsg-live.co.uk
Reply-to: Digital Forensics Magazine <marketing@digitalforensicsmagazine.com>
Subject: New Subscription Prices from Digital Forensics Magazine
Message-Id: <20100323101540.527637DCB1C@mailer.emsg-live.co.uk>

The headers from “From: "Digital Forensics Magazine"...” downwards are the ones supplied by the attacker. The three headers at the top are added by the mail client and show when the email was received and what its status is (“X-Mozilla-Status: 0001” means it has been read).

The most interesting header and the one that gives the game away is the “Received:” header. It says it is from “mailer.emsg-live.co.uk” but then, in brackets, it shows that it actually came from “me.mydomain.co.uk” and it helpfully gives the IP address of the machine it was sent from in square brackets (this IP address has been obscured in this article to protect the guilty). We can also see, from the same header, that the time it was actually sent was not the time that was specified in the “Date:” header lower down.

Each mail server that handles the email will add its own “Received:” header above the previous ones and it is this trail of headers that can be used to trace an email back to its source. However, one word of warning. A sophisticated attacker can manipulate this trail by inserting false “Received:” headers to make it look as though the email came from somewhere it didn't. As RFC 5321 states:

“SMTP mail is inherently insecure in that it is feasible for even fairly casual users to negotiate directly with receiving and relaying SMTP servers and create messages that will trick a naive recipient into believing that they came from somewhere else. Constructing such a message so that the "spoofed" behavior cannot be detected by an expert is somewhat more difficult, but not sufficiently so as to be a deterrent to someone who is determined and knowledgeable.”

So it would appear that detecting and attributing malicious emails is a hopeless task. Luckily for us, those who aim to deceive us are normally not that clever. Here are the headers from the last three malicious emails that I have received. See if you can spot the errors and see if you can work out where they came from:

```
Return-path: <secure-alert@alliance-leicester.co.uk>
Envelope-to: tw@victim.co.uk
Delivery-date: Wed, 31 Mar 2010 19:44:56 +0100
Received: from [121.10.121.80] (helo=Hostmail)
            by inmx06.plus.net with esmtp (PlusNet MXCore v2.00) id 1Nx2uN-0005DO-Fp
            for tw@victim.co.uk; Wed, 31 Mar 2010 19:44:56 +0100
Received: from User [61.137.93.80] by Hostmail with ESMTP
  (SMTPD-8.21) id A32628218; Wed, 31 Mar 2010 22:58:46 +0800
From: "Alliance & Leicester Security Alert"<secure-alert@alliance-leicester.co.uk>
Date: Wed, 31 Mar 2010 16:59:39 +0200
MIME-Version: 1.0
Content-Type: text/html;
        charset="Windows-1251"
Content-Transfer-Encoding: 7bit
X-Priority: 3
X-MSMail-Priority: Normal
X-Mailer: Microsoft Outlook Express 6.00.2600.0000
X-MimeOLE: Produced By Microsoft MimeOLE V6.00.2600.0000
Message-Id: <201003312259158.SM03620@User>
To:
Subject: Alliance & Leicester: Notification of Irregular Account Activity On Your Account
```

```
Return-path: <apache@server.hardtec.srv.br>
Envelope-to: tw@victim.co.uk
Delivery-date: Thu, 01 Apr 2010 14:15:37 +0100
Received: from [187.0.211.213] (helo=server.hardtec.srv.br)
            by pih-inmx09.plus.net with esmtp (PlusNet MXCore v2.00) id 1NxKFE-0006Wc-4I
            for tw@victim.co.uk; Thu, 01 Apr 2010 14:15:36 +0100
Received: from server.hardtec.srv.br (unknown [127.0.0.1])
          by server.hardtec.srv.br (Postfix) with ESMTP id 2E3E6D93054
          for <tw@victim.co.uk>; Thu,  1 Apr 2010 13:11:05 +0000 (UTC)
Received: by server.hardtec.srv.br (Postfix, from userid 48)
          id 0454D710AFF9; Thu,  1 Apr 2010 13:08:11 +0000 (UTC)
To: tw@victim.co.uk
From: Halifax Bank <info@halifax.co.uk>
Reply-To:
MIME-Version: 1.0
Content-Type: text/html
Content-Transfer-Encoding: 8bit
Message-Id: <20100401130811.0454D710AFF9@server.hardtec.srv.br>
Date: Thu,  1 Apr 2010 10:08:11 -0300 (BRT)
Subject: Halifax Online Team Account Notification
```

```
Return-path: <xzznrg@yahoo.com>
Envelope-to: tw@victim.co.uk
Delivery-date: Mon, 05 Apr 2010 07:39:24 +0100
Received: from [95.168.183.140] (helo=srv.multimedyahosting.com)
            by inmx04.plus.net with esmtp (PlusNet MXCore v2.00) id 1Nyfy0-0004Vc-A7
            for tw@victim.co.uk; Mon, 05 Apr 2010 07:39:24 +0100
Received: (qmail 18101 invoked from network); 5 Apr 2010 09:40:47 +0300
Received: from unknown (HELO User) (190.254.17.41)
  by softdnserror with SMTP; 5 Apr 2010 09:40:47 +0300
Reply-To: xzznrg@yahoo.com
From: PayPal<xzznrg@yahoo.com>
Date: Mon, 5 Apr 2010 01:39:15 -0500
MIME-Version: 1.0
Content-Type: multipart/mixed;
          boundary="----=_NextPart_000_009E_01C2A9A6.023CCCD0"
X-Priority: 3
X-MSMail-Priority: Normal
X-Mailer: Microsoft Outlook Express 6.00.2600.0000
X-MimeOLE: Produced By Microsoft MimeOLE V6.00.2600.0000
Message-ID: <E1Nyfy0-0004Vc-A7@inmx04.plus.net>
To:
Subject: Notification of Limited Account Access
```

Additional Reading

https://tools.ietf.org/html/rfc5321
https://tools.ietf.org/html/rfc5322

Appendix 1

…

From: "Digital Forensics Magazine" <digitalforensicsmagazine@mailer.emsg-live.co.uk>
To: "Digital Forensics Magazine" <tw@mydomain.co.uk>
Date: Tue, 23 Mar 2010 10:15:09 +0000
Sender: digitalforensicsmagazine@mailer.emsg-live.co.uk
Reply-to: Digital Forensics Magazine <marketing@digitalforensicsmagazine.com>
Subject: New Subscription Prices from Digital Forensics Magazine
MIME-Version: 1.0
Content-Type: multipart/alternative;
boundary="b1_5fc6d29ab134767240d462b85f431cfa"
Message-Id: <20100323101540.527637DCB1B@.com>

--b1_5fc6d29ab134767240d462b85f431cfa
Content-Type: text/plain; charset = "utf-8"
Content-Transfer-Encoding: 8bit

Price Change Special ++ Price Change Special …

Please visit the following URL in your web browser to unsubscribe -
http://clicks.emsg-live.co.uk/profile/S-10768@7354432@1

Don't forget to forward this email to people who you think will find Digital Forensics Magazine of interest
Add us to your contact list to make sure you can receive future emails safely

--b1_5fc6d29ab134767240d462b85f431cfa
Content-Type: text/html; charset = "utf-8"
Content-Transfer-Encoding: 8bit

<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">

...

<h2 class="style1" align="left" style="font-family:Arial, Arial, Helvetica, sans-serif;
font-weight: normal; font-size: 16px; margin: 0px; padding: 0px;">
Price Change Special ++ Price Change Special …
</h2>

...

Please click here to unsubscribe.

Don't forget to forward this email to people who you think will find Digital Forensics Magazine of interest
Add us to your contact list to make sure you can receive future emails safely

...

</body>
</html>

--b1_5fc6d29ab134767240d462b85f431cfa--

Appendix 2

$ **nc mail.mydomain.co.uk 25**

220 mail.mydomain.co.uk ESMTP Sendmail 8.13.6/8.13.6; Mon, 5 Apr 2010 20:35:15 +0100 (BST)

EHLO mailer.emsg-live.co.uk

250- mail.mydomain.co.uk Hello me.mydomain.co.uk [146.XXX.XX.XXX], pleased to meet you

250-ENHANCEDSTATUSCODES

250-PIPELINING

250-EXPN

250-VERB

250-8BITMIME

250-SIZE 16000000

250-DSN

250-ETRN

250-DELIVERBY

250 HELP

MAIL FROM:<digitalforensicsmagazine@mailer.emsg-live.co.uk>

250 2.1.0 <digitalforensicsmagazine@mailer.emsg-live.co.uk>... Sender ok

RCPT TO:<tw@victim.co.uk>

250 2.1.5 <tw@victim.co.uk>... Recipient ok

DATA

354 Enter mail, end with "." on a line by itself

From: "Digital Forensics Magazine" <digitalforensicsmagazine@mailer.emsg-live.co.uk>

To: Tim Watson <tw@victim.co.uk>

Cc: Bank of England <cashier@bankofengland.co.uk>

Date: Tue, 23 Mar 2010 10:15:10 +0000

Sender: digitalforensicsmagazine@mailer.emsg-live.co.uk

Reply-to: Digital Forensics Magazine <marketing@digitalforensicsmagazine.com>

Subject: New Subscription Prices from Digital Forensics Magazine

Message-Id: <20100323101540.527637DCB1C@mailer.emsg-live.co.uk>

Dear Tim,

Thanks for setting up your direct debit. We will take loads of money from your bank account every month. If you would prefer us not to then please cancel this by visiting http://evilwebsite.com/sucker

Thanks,

Digital Forensics Team

.

250 2.0.0 o35JZFxs008825 Message accepted for delivery

QUIT

221 2.0.0 mail.mydomain.co.uk closing connection