

SMTP Information gathering

Lluís Mora, Neutralbit
llmora@neutralbit.com

Black Hat Europe
Amsterdam, NL // March 2007

- E-mail is present in nearly every organization
- We all understand how it works
 - How envelope and headers work
 - How it can be spoofed
 - How it can be read in transit
 - What a message looks like
 - What to say and what to keep to ourselves
- But what does a message tell about its sender?



- What makes SMTP messages so interesting?
- Control information is embedded in the message
 - Some headers are mandatory, others can be stripped
 - All of them usually end up stored in the mailbox
- Mailing list archives
 - Public logs of our communications
 - Stored over the years
 - The ultimate SMTP information gatherer source!

- Received headers: an advanced “record route”
 - Probably the most well-known information gathering aspect of SMTP
 - Mandatory, per RFC2821: each node adds its header, no one touches the headers
 - Used to prevent mail loops and debug delivery
 - Strip with caution

- Each relay adds
 - IP address of sending gateway
 - FQDN of receiving server
 - Transfer protocol
 - MTA server software
 - Timestamp, including time zone

```
Received: from relay.example.com (201.20.51.192)
        by neutralbit.com (Postfix) with ESMTP id 35B83500EC
        for <llmora@neutralbit.com>; Mon, 15 May 2006 20:26:52 +0000
(UTC)
```

- Not a traceroute...
 - SMTP path, not at the IP level
- ... but has its own advantages
 - Allows us to peek behind NAT and firewalls
 - Point-to-point relaying
 - It is initiated by the victim, part of the communication
- Not rocket science
 - Everybody knows about them, but are we conscious of what they tell about us?

- Corporate IP subnetting
 - Received header addresses are not translated
 - Internal IP addressing scheme
 - Type of connection to the internet

```
Received: from smtp.example.com (6.Net-45-12-192.dynami cIP.example.net  
[192.12.45.6])
```

```
by mail.example.org (Postfix) with ESMTP id 0AB0E147B1
```

```
Received: from smtp.example.com (smtp.example.com [172.18.5.21])
```

```
by mx1.example.com (8.11.6/8.11.6) with ESMTP id i82sokwis;
```

```
Received: from vailo (172.16.1.100)
```

```
by smtp.example.com (Postfix) with ESMTP id i82shwk;
```

- Corporate Internet access policies
 - Centralized Internet access?
 - Each location has a public connection?

```
Received: from mx1.uk.example.com ([195.166.192.8])  
        by vger.kernel.org
```

```
From: John Doe <j.doe@uk.example.com>
```

```
Received: from smtp.de.example.com ([32.1.120.11])  
        by vger.kernel.org
```

```
From: Pam Plinas <pplinas@de.example.com>
```


- Server fingerprinting
 - Software and versions
 - Location based on time zones

```
Received: from mx2.example.mil [192.18.1.12]
        by gatekeeper with POP3 (fetchmail-6.3.0)
        for <j.doe@example.com> (single-drop); Mon, 02 Jan 2006 14:43:41 -0800
```

(PST)

```
Received: from mx1.example.mil ([192.168.1.2])
        by mx2.example.mil with Microsoft SMTPSVC(6.0.3790.211);
        Tue, 3 Jan 2006 07:44:01 +0900
```

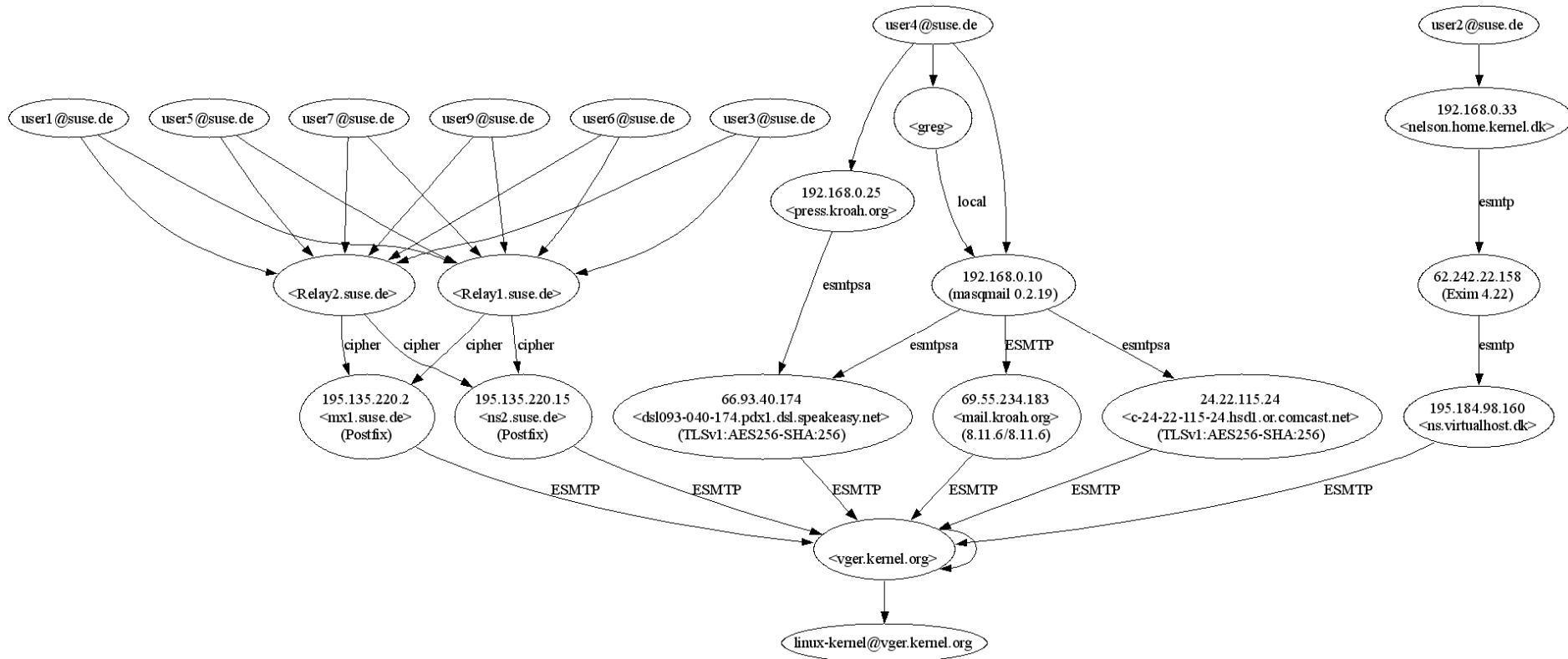
- Relay link information
 - SMTP Link encryption

```
Received: from lappy (192.168.1.4) by pub.example.net (qmail) with ESMTPLink
  ID MG0007DA (SSL/TLS, 3DES, CBC mode, keysize 192 bits) ; 8 Sep 2006 16:40:03
+0200
```

```
Received: from [24.26.7.196] (ilm.example.com [24.26.7.196])
  (using TLSv1 with cipher DHE-RSA-AES256-SHA (256/256 bits))
  (No client certificate requested)
```

- Graphic representation of SMTP paths
 - Definitively flashier than staring at logs
 - Parsing of “Received” headers is challenging
 - Absorb more information at once
 - One image...
- A few examples
 - Data extracted from Linux kernel mailing list
 - Around 3 months in early 2006

SMTP Network mapping (VIII)

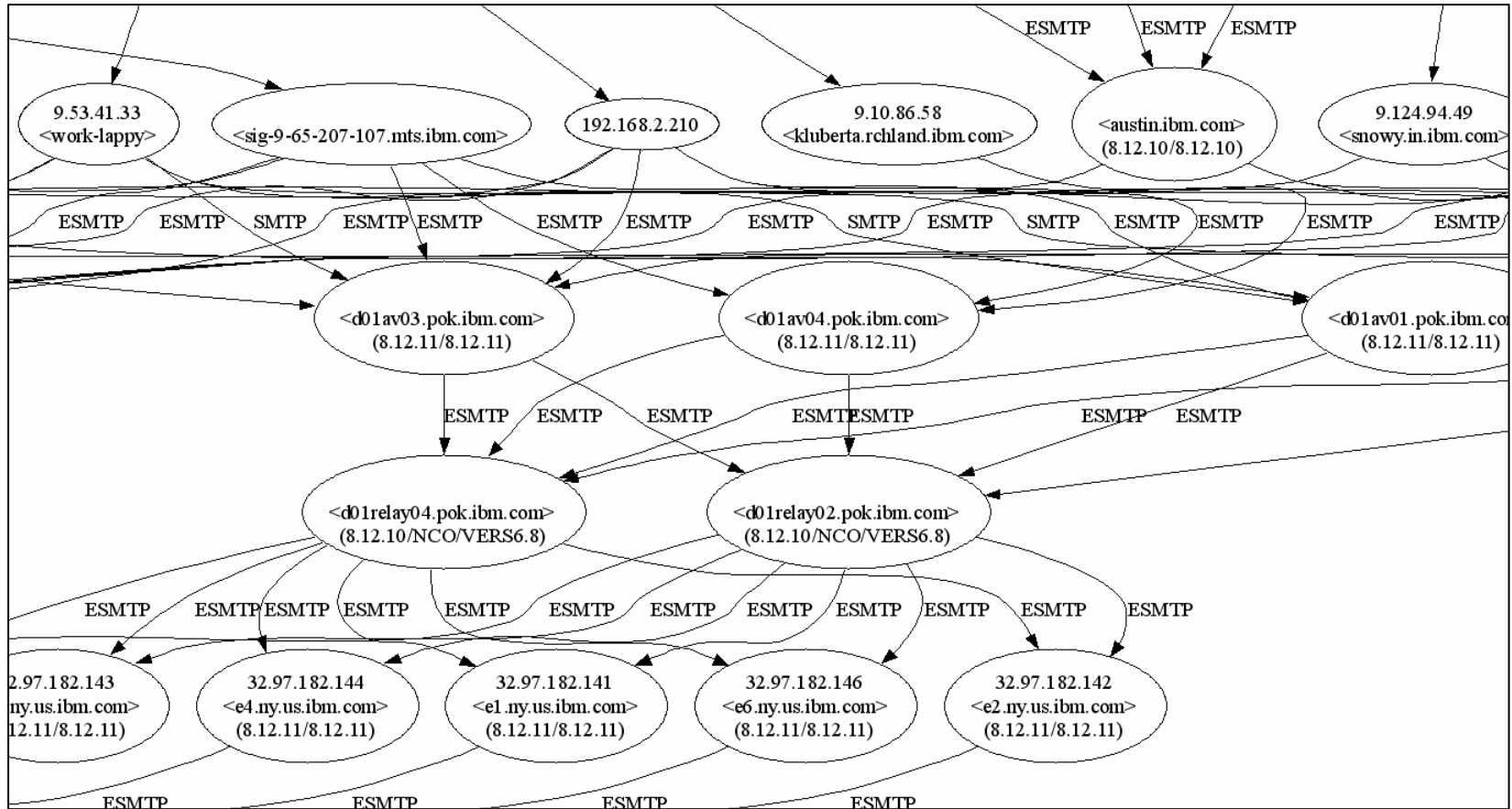


spot the telecommuters ...



... target selection?

SMTP Network mapping (IX)



where is wally?

- Based on a different set of headers
 - User-Agent
 - X-Mailer
 - X-MIME-OLE
- Excellent level of details
 - Down to the patch level
- Not used for anything else

X-Mailer: Microsoft Office Outlook, Build 11.0.5510

User-Agent: Thunderbird 1.5.0.7 (Windows/20060909)

X-Mailer: ColdFusion MX Application Server

X-MimeOLE: Produced By Microsoft MimeOLE V6.00.2900.2962

X-Mailer: Evolution 2.2.3 (2.2.3-4.fc4)

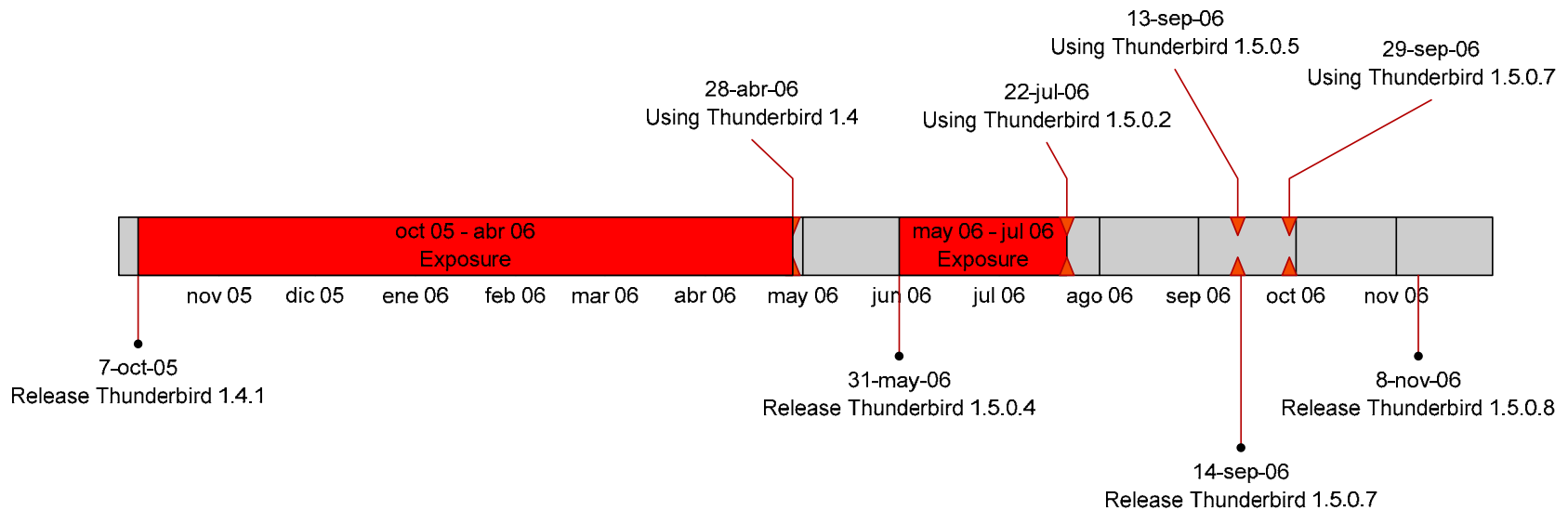
X-Mailer: iPlanet Messenger Express 5.2 Patch 2 (built Jul 14 2004)

X-Mailer: Lotus Notes Release 5.0.6a January 17, 2001

User-Agent: SquirrelMail/1.4.3a

User-Agent: Wanderlust/2.12.0 (Your Wildest Dreams) SEMI/1.14.6 (Maruoka)
FLIM/1.14.7 APEL/10.6 MULE XEmacs/21.5 (beta21)
(corn) (+CVS-20050720) (i386-suse-linux)

- Long term analysis
 - If we get access to a long stretch of messages
 - Plot client mailers over time...
 - ... then add mailer release dates

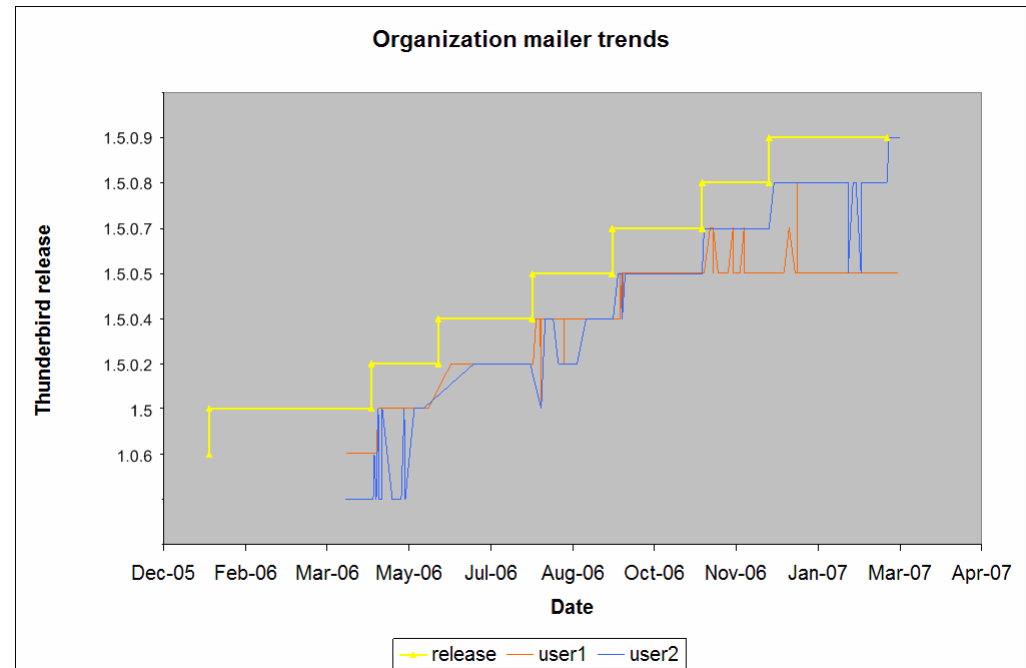


- Organization trend analysis

- With enough e-mails, we can find out details about the organization policies

- Patching policies
- Application usage
- Security gaps
- Policy exceptions

...maybe not just for SMTP servers?



- Other interesting facts can be guessed
 - Same e-mail address + alternating mailers + multiple IP addresses → multiple locations (home / work?)
 - Same e-mail address + same mailer + multiple IP addresses → take the laptop home
 - Various e-mail domains + same mailer + same IP address → non-corporate mail at work
 - Changing “Date” time zones → user on the go?

• Indirect sources of information

– Implementation differences

- Ordering of headers
- Quoted replies

– Custom X-Headers

- X-Originating-IP, etc.
- Antivirus / Antispam

```
Subject: Re: [RELEASE 4] Testing patch #49192
Date: Tue, 21 Feb 2006 10:21:14 +0100
X-Originating-IP: 10.2.1.122
X-Virus-Scanned: by amavisd-new-20030616-p10
(Debian)
X-Spam-Checker-Version: SpamAssassin 3.0.2 (2004-
11-1
X-Spam-Status: No, score=-1.4 required=2.0
```

– Message contents

- User data
- Encoding data

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML><HEAD>
<META http-equiv=Content-Type content="text/html; charset=iso-8859-1">
<META content="MSHTML 6.00.6000.16414" name=GENERATOR></HEAD>
<BODY>
<DIV><SPAN class=044560813-09032007><FONT face=Verdana size=2>Dear user,
</FONT></SPAN></DIV>
<IMG height=92 src="cid:044560813@09032007-1B02" width=191
```

- Indirect sources of information
 - Encoded data in unsuspecting headers

```
Message-ID: <Pine.LNX.4.21.0611280421440.26304-100000@example.org>
```

```
Message-ID: <1103.203.41.53.196.1128283359.squirrel@mail.example.com>
```

```
Message-ID: <11363603.1154544476739.JavaMail.root@as.example.net>
```

```
Content-Type: multipart/mixed; boundary=Apple-Mail-1-944594902
```

- Strip unneeded information at border gateways whenever possible
- Find out what has already leaked and fix it
- Analysis relies on client provided data, handle with care

Thank you!

Lluís Mora

llmora@neutralbit.com

neutralbit*securityinnovation*