Network Working Group
Request for Comments: 4407
Category: Experimental

J. Lyon
Microsoft Corp.

April 2006

Purported Responsible Address in E-Mail Messages

Status of This Memo

Copyright Notice

IESG Note

The following documents (RFC 4405, RFC 4406, RFC 4407, and RFC 4408) are published simultaneously as Experimental RFCs, although there is no general technical consensus and efforts to reconcile the two approaches have failed. As such, these documents have not received full IETF review and are published "AS-IS" to document the different approaches as they were considered in the MARID working group.

The IESG takes no position about which approach is to be preferred and cautions the reader that there are serious open issues for each approach and concerns about using them in tandem. The IESG believes that documenting the different approaches does less harm than not documenting them.

Note that the Sender ID experiment may use DNS records that may have been created for the current SPF experiment or earlier versions in this set of experiments. Depending on the content of the record, this may mean that Sender-ID heuristics would be applied incorrectly to a message. Depending on the actions associated by the recipient with those heuristics, the message may not be delivered or may be discarded on receipt.

Participants relying on Sender ID experiment DNS records are warned that they may lose valid messages in this set of circumstances. Participants publishing SPF experiment DNS records should consider the advice given in section 3.4 of RFC 4406 and may wish to publish both v=spf1 and spf2.0 records to avoid the conflict.

Participants in the Sender-ID experiment need to be aware that the way Resent-* header fields are used will result in failure to receive legitimate email when interacting with standards-compliant systems (specifically automatic forwarders which comply with the standards by not adding Resent-* headers, and systems which comply with RFC 822 but have not yet implemented RFC 2822 Resent-* semantics). It would be inappropriate to advance Sender-ID on the standards track without resolving this interoperability problem.

The community is invited to observe the success or failure of the two approaches during the two years following publication, in order that a community consensus can be reached in the future.

Abstract

This document defines an algorithm by which, given an e-mail message, one can extract the identity of the party that appears to have most proximately caused that message to be delivered. This identity is called the Purported Responsible Address (PRA).

Table of Contents

1. Introduction

Most e-mail flows relatively directly from a sender to a recipient, with a small number of Mail Transfer Agents (MTAs) in between. Some messages, however, are resent by forwarding agents, mailing list servers, and other such software. These messages effectively result in two or more mail transactions: one from the sender to the forwarding agent, and another from the agent to the destination.

In some cases, messages travel through more than one of these agents. This can occur, for example, when one mailing list is subscribed to another, or when the address subscribed to a mailing list is a forwarding service.

Further complicating the situation, in some cases the party that introduces a message is not the author of the message. For example, many news web sites have a "Mail this article" function that the

public can use to e-mail a copy of the article to a friend. In this case, the mail is "from" the person who pressed the button, but is physically sent by the operator of the web site.

This document defines a new identity associated with an e-mail message, called the Purported Responsible Address (PRA), which is determined by inspecting the header of the message. The PRA is designed to be the entity that (according to the header) most recently caused the message to be delivered.

Note that the results of this algorithm are only as truthful as the headers contained in the message; if a message contains fraudulent or incorrect headers, this algorithm will yield an incorrect result. For this reason, the result of the algorithm is called the "Purported Responsible Address" -- "purported" because it tells you what a message claims about where it came from, but not necessarily where it actually came from.

This document does not prescribe any particular uses for the Purported Responsible Address. However, [RFC4406] describes a method of determining whether a particular MTA is authorized to send mail on behalf of the domain contained in the PRA.

1.1. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

2. Determining the Purported Responsible Address

The PRA of a message is determined by the following algorithm:

1. Select the first non-empty Resent-Sender header in the message. If no such header is found, continue with step 2. If it is preceded by a non-empty Resent-From header and one or more Received or Return-Path headers occur after said Resent-From header and before the Resent-Sender header, continue with step 2. Otherwise, proceed to step 5.

2. Select the first non-empty Resent-From header in the message. If a Resent-From header is found, proceed to step 5. Otherwise, continue with step 3.

3. Select all the non-empty Sender headers in the message. If there are no such headers, continue with step 4. If there is exactly one such header, proceed to step 5. If there is more than one such header, proceed to step 6.

4. Select all the non-empty From headers in the message. If there is exactly one such header, continue with step 5. Otherwise, proceed to step 6.

5. A previous step has selected a single header from the message. If that header is malformed (e.g., it appears to contain multiple mailboxes, or the single mailbox is hopelessly malformed, or the single mailbox does not contain a domain name), continue with step 6. Otherwise, return that single mailbox as the Purported Responsible Address.

6. The message is ill-formed, and it is impossible to determine a Purported Responsible Address.

For the purposes of this algorithm, a header field is "non-empty" if and only if it contains any non-whitespace characters. Header fields that are otherwise relevant but contain only whitespace are ignored and treated as if they were not present.

Note that steps 1 and 2 above extract the Resent-Sender or Resent-From header from the first resent block (as defined by section 3.6.6 of [RFC2822]) if any. Steps 3 and 4 above extract the Sender or From header if there are no resent blocks.

Note that what constitutes a hopelessly malformed header or a hopelessly malformed mailbox in step 5 above is a matter for local policy. Such local policy will never cause two implementations to return different PRAs. However, it may cause one implementation to return a PRA where another implementation does not. This will occur only when dealing with a message containing headers of questionable legality.

Although the algorithm specifies how messages that are not in strict conformance with the provisions of RFC 2822 should be treated for the purposes of determining the PRA, this should not be taken as requiring or recommending that any systems accept such messages when they otherwise would not have done so. However, if a liberal implementation accepts such messages and desires to know their PRAs, it MUST use the algorithm specified here.

Where messages conform to RFC 822 rather than RFC 2822, it is possible for the algorithm to give unexpected results. An RFC822 message should not normally contain more than one set of resent headers; however, the placement of those headers is not specified, nor are they required to be contiguous. It is therefore possible that the Resent-From header will be selected even though a Resent-Sender header is present. Such cases are expected to be rare or non-existent in practice.

3. Security Considerations

The PRA, as described by this document, is extracted from message headers that have historically not been verified. Thus, anyone using the PRA for any purpose MUST be aware that the headers from which it is derived might be fraudulent, malicious, malformed, and/or incorrect. [RFC4406] describes one mechanism for validating the PRA.

A message's PRA will often be extracted from a header field that is not normally displayed by existing mail user agent software. If the PRA is used as part of a mechanism to authenticate the message's origin, the message SHOULD NOT be displayed with an indication of its authenticity (positive or negative) without the PRA header field also being displayed.

4. Acknowledgements

The PRA concept was first published in [CallerID]. It has been refined using valuable suggestions from members of the MARID working group.

5. References

5.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2822] Resnick, P., "Internet Message Format", RFC 2822, April 2001.

5.2. Informative References

[CallerID] Microsoft Corporation, Caller ID for E-Mail Technical Specification, http://www.microsoft.com/mscorp/safety/technologies/senderid/resources.mspx

[RFC4406] Lyon, J. and M. Wong, "Sender ID: Authenticating E-Mail", RFC 4406, April 2006.

Author's Address

Jim Lyon
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052
USA

EMail: jimlyon@microsoft.com

Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).