Email Author Header Field
draft-crocker-email-author-00

Abstract

   Internet mail defines the From: field to indicate the author of the
   message's content and the Sender: field to indicate who initially
   handled the message. The Sender: field is optional, if it has the
   same information as the From: field. That is, when the Sender: field
   is absent, the From: field has conflated semantics, as both a
   handling identifier and a content creator identifier. This was not a
   problem, until development of stringent protections on use of the
   From: field. It has prompted Mediators, such as mailing lists, to
   modify the From: field, to circumvent mail rejection caused by those
   protections.

   This affects end-to-end behavior of email, between the author and the
   final recipients, because mail from the same author is not treated
   the same, depending on what path it followed. In effect, the From:
   field has become dominated by its role as a handling identifier. The
   current specification augments the current use of the From: field, by
   specifying the Author: field, which identifies the original author of
   the message and is not subject to modification by Mediators.

Copyright Notice

Table of Contents

1. Introduction

   Internet mail conducts asynchronous communication from an author to
   one or more recipients, and is used for ongoing dialogue amongst
   them. Email has a long history of serving a wide range of human uses
   and styles, within that simple framework, and the mechanisms for
   making email robust and safe serve that sole purpose.

   Internet mail defines the From: field to indicate the author of the
   message's content and the Sender: field to indicate who initially
   handled the message. [Mail-Fmt] The Sender: field is optional, if it
   has the same information as the From: field. That is, when the
   Sender: field is absent, the From: field has conflated semantics, as
   both a handling identifier and a content creator identifier. These
   fields were initially defined in [RFC733] and making the redundant
   Sender: field optional was a small, obvious optimization, in the days
   of slower communications, expensive storage and less powerful
   computers.

The dual semantics was not a problem, until development of stringent
protections on use of the From: field. It has prompted Mediators,
such as mailing lists, to modify the From: field, to circumvent mail
rejection caused by those protections. This affects end-to-end
usability of email, between the author and the final recipients,
because mail received from the same author is treated differently by
the recipient's software, depending on what path the message
followed.

By way of example, mail from "Example User <user@example.com>" which
is sent directly to a recipient, will show the user's display name
correctly and can correctly analyze and aggregate mail from that
user, based on their email address. However if the user sends
through a mailing list, and the mailing list conducts a common form
of From: modification, needed to bypass enforcement of stringent
authentication policies, then the received message might have a From:
field along the lines of "Example User via Example List
<listname@list.example.com>". The change inserts an operational
address, for the Mediator, into the From: field, and distorts the
field's display-name, as a means of recording the modification. The
result is that the recipient's software will see the message as being
from an entirely different author and will handle it separately.
(Mediators might create a Reply-To: field, with the original From:
field email address, but this does nothing to aid other processing
done by the recipient's MUA based on what it believes is the author's
address or original display-name.)

In effect, the From: field has become dominated by its role as a
handling identifier. The current specification augments the current
use of the From: field, by specifying the Author: field, which
identifies the original author of the message and is not subject to
modification by Mediators.

While it might be cleanest to move towards more reliable use of the
Sender: field and then to target it as the focus of authentication
concerns, enhancement of standards works best with incremental
additions, rather than efforts at replacement. To that end, this
specification provides a means of supplying author information that
is not subject to modification by processes seeking to enforce
stringent authentication.

Terminology and architectural details in this document are
incorporated from [Mail-Arch].

Discussion of this draft is directed to the ietf-822@ietf.org mailing
list.

2. Author Header Field

A new message header field is defined: Author:. It has the same
syntax as From:. [Mail-Fmt] As with the original and primary intent
for the From: header field, the Author: header field is to contain
the email address and can contain the displayable human name of the
author for the message content.

The [ABNF] for the field's syntax is:

                        author = "Author:" mailbox-list CRLF

This header field can be added as part of the original message create
process, or it can be added later, to preserve the original author
information from the From: field.

3. Discussion

The Author: header field, here, is intended for creation during
message generation or during mediation. It is intended for use by
recipient MUAs, as they typically use the From: field. In that
regard, it would be reasonable for an MUA that would normally
organize or display information based on the From: field to give the
Author: header field priority.

The X-Original-From: header field is referenced in [rfc5703] but is
not actually defined. Further, it is registered with IANA, but the
registry cites RFC5703 as the controlling source for the entry.
Lastly, the field is solely intended for use by Mediators, to
preserve information from a modified From:

Obviously any security-related processing of a message needs to
distinguish From: from Author: and treat their information
accordingly.

4. Security Considerations

Any header field containing identification information is a source of
security and privacy concerns, especially one pertaining to content
authorship. Generally, the handling of the Author: header field
needs to receive scrutiny and care comparable to that given to the
From: header field.

Given the semantics of this field, it is easy to believe that use of
this field will create a new attack vector for tricking end-users.
However, for all of the real and serious demonstration of users'
being tricked by deceptive or false content in a message, there is no
evidence that problematic content in a field providing information

   about message's author directly contributes to differential and
   problematic behavior by the end user.

5. IANA Considerations

5.1. Registration of the Author header field

   Header field name: Author

   Applicable protocol: mail

   Status: provisional

   Author/Change controller: *** ??? ***

   Specification document(s): *** This document ***

6. References

6.1. Normative References

   [ABNF]   Dave, D., Ed. and P. Paul, "Augmented BNF for Syntax
              Specifications: ABNF", RFC 5234, January 2008.

   [Mail-Arch]
              Crocker, D., "Internet Mail Architecture", RFC 5598, July
              2009.

   [Mail-Fmt]
              Resnick, P., Ed., "Internet Message Format", RFC 5322,
              October 2008.

6.2. Informative References

   [rfc5703] Hansen, T. and C. Daboo, "Sieve Email Filtering: MIME Part
              Tests, Iteration, Extraction, Replacement, and Enclosure",
              RFC 5703, October 2009.

   [RFC733]   Crocker, D., Vittal, J., Pogran, K., and D. Henderson,
              "Standard for the Format of ARPA Network Text Messages",
              RFC 733, November 1977.

Author's Address

   Dave Crocker
   Brandenburg InternetWorking

   Email: dcrocker@bbiw.net