Network Working Group
Internet Draft
<draft-shafranovich-feedback-report-02.txt>
Intended status: Informational
Expires: November 2007

Y. Shafranovich
SolidMatrix Technologies, Inc.
J. Levine
P. Hoffman
Domain Assurance Council
May 2007

An Extensible Format for Email Feedback Reports

draft-shafranovich-feedback-report-02.txt

Status of this Memo

Copyright Notice

Abstract

This document defines an extensible format and MIME type that may be used by network operators to report feedback about received email to other parties. This format is intended as a machine readable replacement for various existing report formats currently used in Internet email.

Table of Contents

1. Introduction

As the spam problem continues to expand and potential solutions evolve, network operators are increasingly exchanging abuse reports among themselves and other parties. However, different operators have defined their own formats, and thus the receivers of these reports are forced to write custom software to interpret each. In addition, many operators use various other report formats to provide non-abuse-related feedback about processed email. This memo seeks to define a standard extensible format and the "message/feedback-report" MIME type for these reports in accordance with RFC 2048 [RFC2048]. This format and content type is intended to be used within the scope of the framework of the "multipart/report" content type defined in RFC 3462 [RFC3462]. While there have been previous work in this area([STRADS_BCP] and [ASRG_ABUSE]), none of them have yet been sucessful. It is hoped that this document will have a better fate.

This format is intended primarily as an Abuse Reporting Format (ARF) for reporting email abuse but also includes support for direct feedback via end user mail clients, reports of some types of virus activity, and some similar issues.

This document only defines the format and content type to be used for these reports. Determination of where these reports should be sent, how trust among report generators and report recipients is established, and reports related to more than one message are outside the scope of this document. It is assumed that best practices will evolve over time, and will be codified in future documents.

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 RFC 2119 [RFC2119].

2. Intent

The reports defined in this document are intended for several purposes:

- To inform ISPs about email abuse originating from or related to their networks
- To inform email service providers or other primarily outbound senders that there may be issues regarding their mail. These issues include (but are not limited to) reports that the mail may be considered to be "spam" by a recipient of the message.
- To inform email service provides about opt-out requests
- To advise providers that certify or otherwise make assertions about mail of recipient disagreement with the assertions.

Please note that while the parent "multipart/report" content type defined in RFC 3462 is used for all kinds of administrative messages, this format is intended specifically for communications among providers regarding email abuse and related issues, and SHOULD NOT be used for other reports.

3. Requirements

The following requirements are necessary for feedback reports (the actual standard is defined in the next sections) :

- They must be both human and machine readable
- A copy of the original email message (body and headers) or the message headers must be enclosed in order to allow the receiver to properly handle the report.
- The machine readable section must provide ability for the report generators to share metadata with receivers,
- The format must be extensible.

4. Format of Email Feedback Reports

To satisfy the requirements, an email feedback report is defined as a MIME message with a top level MIME content type of "multipart/report" (as defined in RFC 3462). The following apply:

a. The "report-type" parameter of "multipart/report" type is set to "feedback-report"

b. The first MIME part of the message contains a human readable description of the report and MUST be included.

c. The second MIME part of the message is a machine-readable section with the content type of "message/feedback-report" (defined later on in this document) and MUST be included. This section is intended to convey metadata about the report in question that may not be readily available from the included email message itself.

d. The third MIME part of the message contains either a full copy of the original message with a MIME content type of "message/rfc822" (as defined in RFC 2046 [RFC2046]) OR a copy of the headers from the original message with MIME content type of "text/rfc822-headers" (as defined in RFC 3462). This part MUST be included (unlike RFC 3462). While some operators may choose to modify or redact this portion for privacy or legal reasons, it is RECOMMENDED that the entire original email message be included without any modification.

e. Each feedback report MUST be related to only a SINGLE email message. Summary and aggregate formats are outside the scope of this specification.

f. The subject line of the feedback report SHOULD be the same as the included email message and MAY include only the standard forwarding prefix used by MUAs such as "FW:". (Many smaller operators using MUAs for abuse handling rely on the subject lines for processing.)

5. Format of 'message/feedback-report' Content Type

This content type provides a machine-readable section intended to let the report generator convey metadata to the report receiver. The intent of this section is to convey information which may not be obvious or may not be easily extracted from the original email message or headers.

The body of this content type consists of multiple "fields" formatted according to the ABNF of RFC 822 [RFC0822] header "fields". This section defines the initial set of fields provided by this specification. Additional fields may be registered according to the procedure described later on in this document. Altough these fields have a syntax similar to those of mail message headers, they are semantically distinct; hence they SHOULD NOT be repeated in the header area of the message containing the report. Note that these fields represent information that the receiver is asserting about the report in question, but are not necessarily verifiable. Report receivers MUST NOT assume that these assertions are always accurate.

5.1 Required Fields

The following header fields are REQUIRED and MUST only appear once:

- "Feedback-Type:" - contains the type of feedback report (as defined in the corresponding IANA registry). This is intended to let report generators distinguish among different types of reports.
- "User-Agent:" - indicates the name and version of the software program that generated the report. The format of this field MUST follow section 14.43 of RFC 2616 [RFC2616]. This field is for documentation only; there is no registry of user agent names or versions, and report receivers SHOULD NOT expect user agent names to belong to a known set.
- "Version" - indicates the version of specification that the report generator is using to generate the report. The version number in this specification is set to "0.1".

5.2 Optional Fields Appearing Once

The following header fields are OPTIONAL and MUST NOT appear more than once:

- "Original-Mail-From:" - copy of the email address used in the MAIL FROM portion of the original SMTP transaction. The format of this field is defined in section 4.1.1.2 of RFC 2821 [RFC2821].
- "Original-Rcpt-To:" - copy of the email address used in the RCPT TO portion of the original SMTP transaction. The format of this field is defined in section 4.1.1.3 of RFC 2821.
- "Received-Date:" - indicates the date the original message was received by recipient system's MTA. This field MUST BE formatted as per section 3.3 of RFC 2822 [RFC2822].
- "Source-IP:" - contains an IPv4 or IPv6 address of the MTA from which the original message was received. IPv4 addresses are to be formatted in dot-decimal notation as currently used by the community. IPv6 addresses MUST BE formatted as per section 2.2 of RFC 2373 [RFC2373].

5.3 Optional Fields Appearing Multiple Times

The following set of header fields are OPTIONAL and MAY appear more than once:

- "Authentication-Results:" - indicates the result of an authentication check run by the report generator. The format of this field is is defined in draft-kucherawy-sender-auth-header [AUTH-HEADER]. Report receivers should note that this field only indicates an assertion made by the report generator.

- "Reported-Domain:" - indicates a domain name that the report generator believes to be relevant to the report. Domain format is defined in section 2.3.1 of RFC 1035 [RFC1035].
- "Reported-URI:" - indicates a URI that the report generator believes to be relevant to the report. URI format is defined in RFC 2396 [RFC2396].
- "Removal-Recipient:" - indicates the email address to be removed from the mailing list (MUST only be used with "opt-out" and "opt-out-list" types). The format of this field is defined in section 3.4.1 of RFC 2822.

6. MIME Type Registration of message/feedback-report

This section provides the media type registration application (as per RFC 2048 [RFC2048].

To: ietf-types@iana.org

Subject: Registration of MIME media types message/feedback-report

MIME media type name: message

MIME subtype name: feedback-report

Required parameters: none

Optional parameters: none

Encoding considerations:

"7bit" encoding is sufficient and MUST be used to maintain readability when viewed by non-MIME mail readers.

Security considerations:

See the "Security Considerations" of this document.

Interoperability considerations: implementors MUST ignore any fields they do not support

Published specification: this document

Applications which use this media type: Abuse helpdesk software for ISPs, mail service bureaus, mail certifiers, and similar organizations

Additional information:

Magic number(s): none

File extension(s): none

Macintosh File Type Code(s): none

Person and email address to contact for further information:

Yakov Shafranovich <ietf@shaftek.org>

Intended usage: COMMON

Author/Change controller: IESG

7. Extensibility

Like many other formats and protocols, this format may need to be extended over time to fit the ever changing landscape of the Internet. Therefore, extensibility is provided via two IANA registries: one for feedback types and a second for header fields. The feedback type registry is to be used in conjunction with the "Feedback-Type" field above. The header name registry is intended for registration of new metadata fields to be used in the machine readable portion (part 2) of this format. Please note that version numbers do not change with new field registrations unless a new specification of this format is published. Also note that all new field registrations can only registered as OPTIONAL fields. Any new required fields REQUIRE a new version of this specification to be published.

In order to encourage extensibility and interoperability of this format, implementors MUST ignore any fields they do not support.

8. IANA Considerations

IANA is requested to register MIME type "message/feedback-report" using the application provided in this document and setup two registries: one for header field names and a second for "Feedback-Type" values. This section contains the templates used for registration of new entries in these registries and initial values. New registrations to these two registries MUST have approval by a Designated Expert in accordance with the Expert Review guidelines as described in RFC 2434 [RFC2434] (the expert should be appointed by the Area Directors of the Applications Area). Any new field registered is considered OPTIONAL unless a new version of this specification is published.

For the header name registry, the following MUST be provided in order to register a new header field name:

1. Name of the field being registered
2. Short description of the field
3. Whether the field can appear more than once
4. Which "Feedback-Type" types does this field apply to (or "any")
5. The RFC number (or Internet draft name) in which this header is registered

If the header field being registered requires its own IANA registry, than the appropriate registry MUST be properly defined.

For the feedback type registry, the following MUST be provided in order to register a new header field name:

1. Name of the feedback type being registered
2. Short description
3. The RFC number (or Internet draft name) in which this feedback type is registered

8.1 Initial Values for the Header Names Registry

The data below are populated from this document. The RFC number used for registration of these values is this document.

Field Name: Authentication-Results
Description: results of authentication check
Multiple Appearances: Yes
Related "Feedback-Type": any

Field Name: Feedback-Type
Description: type of feedback report
Multiple Appearances: No
Related "Feedback-Type": N/A

Field Name: Original-Mail-From
Description: email address used in the MAIL FROM portion of the original SMTP transaction
Multiple Appearances: No
Related "Feedback-Type": any

Field Name: Original-Rcpt-To
Description: copy of the email address used in the RCPT TO portion of the original SMTP transaction
Multiple Appearances: No

Related "Feedback-Type": any

Field Name: Received-Date
Description: date the original message was received
Multiple Appearances: No
Related "Feedback-Type": any

Field Name: Reported-Domain
Description: relevant domain name
Multiple Appearances: Yes
Related "Feedback-Type": any

Field Name: Reported-URI
Description: relevant URI
Multiple Appearances: Yes
Related "Feedback-Type": any

Field Name: Removal-Recipient
Description: email address to be removed from the mailing list
Multiple Appearances: Yes
Related "Feedback-Type": opt-out, opt-out-list

Field Name: Source-IP
Description: IPv4 or IPv6 address from which the original message was received
Multiple Appearances: No
Related "Feedback-Type": any

Field Name: User-Agent
Description: name and version of the program used
Multiple Appearances: No
Related "Feedback-Type": any

Field Name: Version
Description: version of specification used
Multiple Appearances: No
Related "Feedback-Type": any

8.2 Initial values for the "Feedback-Type" registry

The initial names and descriptions are provided below. The RFC number used for registration of these values is this document.

- abuse - spam or some other kind of email abuse
- fraud - indicates some kind of fraud or phishing activity.
- miscategorized - indicates that the content categorization applied in connection with a certification or reputation system was incorrect
- not-spam - indicates that a message that was tagged or categorized as spam (such as by an ISP) is not spam
- opt-out - a request to opt out from ALL mailing lists from this provider.
- opt-out-list - a request to opt out from THIS mailing list ONLY.
- virus - report of a virus found in the originating message
- other - any other feedback that doesn't fit into other types.

9. Security Considerations

All of the Security Considerations from RFC 3462 are inherited here.

This specification describes a report format. This document does not say what a recipient of such a report must, should, or even may do with any report in the format described here.

10. Acknowledgments

The authors would like to thank many of the members of the email community who provided helpful comments and suggestions for this document including many of the participants in ASRG, IETF and MAAWG activities, and all of the members of the abuse-feedback-report public mailing list.

11. References

11.1 Normative References

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.

[RFC2046] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types", RFC 2046, November 1996.

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

[RFC2373] Hinden, R.M. and S.E. Deering, "IP Version 6 Addressing Architecture", RFC 2373, July 1998.

[RFC2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999.

[RFC2821] Klensin, J., "Simple Mail Transfer Protocol", RFC 2821, April 2001.

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

[RFC3462] Vaudreuil, G., "The Multipart/Report Content Type for the Reporting of Mail System Administrative Messages", RFC 3462, January 2003.

11.2 Informative References

[ASRG_ABUSE] Anti-Spam Research Group (ASRG) of the Internet Research Task Force (IRTF), "Abuse Reporting Standards Subgroup of the ASRG", May 2005, <http://asrg.sp.am/subgroups/abuse_reports.shtml>.

[AUTH-HEADER] Kucherawy, M, "Message Header for Indicating Sender Authentication Status", Internet-Draft draft-kucherawy-sender-auth-header-04 (work in progress).

[RFC0822] Crocker, D.H., "Standard for the format of ARPA Internet text messages", STD 11, RFC 822, August 1982.

[RFC2048] Freed, N., Klensin, J., and J. Postel, "Multipurpose Internet Mail Extensions (MIME) Part Four: Registration Procedures", BCP 13, RFC 2048, November 1996.

[RFC2396] Berners-Lee, T., Fielding, R.T., and L. Masinter, "Uniform Resource Identifiers (URI): Generic Syntax", RFC 2396, August 1998.

[RFC2434] Narten, T. and H.T. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.

[STRADS_BCP] Crissman, G., "Proposed Spam Reporting BCP Document", May 2005, <http://www.tmisnet.com/~strads/spam/bcp.html>.

Authors' Addresses

Yakov Shafranovich
SolidMatrix Technologies, Inc.
EMail: ietf@shaftek.org
URI: http://www.shaftek.org

John Levine
Domain Assurance Council
EMail: john.levine@domain-assurance.org

Paul Hoffman
Domain Assurance Council
EMail: paul.hoffman@domain-assurance.org

A. Appendix A - Sample Feedback Reports

A.1 Simple Report for Email Abuse without Optional Headers

```
From: <abusedesk@example.com>
Date: Thu, 8 Mar 2005 17:40:36 EDT
Subject: FW: Earn money
To: <abuse@example.net>
MIME-Version: 1.0
Content-Type: multipart/report; report-type=feedback-report;
     boundary="part1_13d.2e68ed54_boundary"

--part1_13d.2e68ed54_boundary
Content-Type: text/plain; charset="US-ASCII"
Content-Transfer-Encoding: 7bit

This is an email abuse report for an email message received from IP
10.67.41.167 on Thu, 8 Mar 2005 14:00:00 EDT. For more information
about this format please see http://www.mipassoc.org/arf/.

--part1_13d.2e68ed54_boundary
Content-Type: message/feedback-report

Feedback-Type: abuse
User-Agent: SomeGenerator/1.0
Version: 0.1

--part1_13d.2e68ed54_boundary
Content-Type: message/rfc822
Content-Disposition: inline

From: <somespammer@example.net>
Received: from mailserver.example.net
     (mailserver.example.net [10.67.41.167])
     by example.com with ESMTP id M63d4137594e46;
     Thu, 08 Mar 2005 14:00:00 -0400
To: <Undisclosed Recipients>
Subject: Earn money
MIME-Version: 1.0
Content-type: text/plain
Message-ID: 8787KJKJ3K4J3K4J3K4J3.mail@example.net
Date: Thu, 02 Sep 2004 12:31:03 -0500

Spam Spam Spam
Spam Spam Spam
Spam Spam Spam
Spam Spam Spam
--part1_13d.2e68ed54_boundary--
```

A.2 Opt-Out Report without Message Body

```
From: <abusedesk@example.com>
Date: Thu, 8 Mar 2005 17:40:36 EDT
Subject: FW: Earn money
To: <abuse@example.net>
MIME-Version: 1.0
Content-Type: multipart/report; report-type=feedback-report;
     boundary="part1_13d.2e68ed54_boundary"

--part1_13d.2e68ed54_boundary
Content-Type: text/plain; charset="US-ASCII"
Content-Transfer-Encoding: 7bit

This is an opt-out report for an email message received from IP
10.67.41.167 on Thu, 8 Mar 2005 14:00:00 EDT. For more information
about this format please see http://www.mipassoc.org/arf/.

--part1_13d.2e68ed54_boundary
Content-Type: message/feedback-report

Feedback-Type: opt-out
User-Agent: SomeGenerator/1.0
Version: 0.1
Removal-Recipient: user@example.com

--part1_13d.2e68ed54_boundary
Content-Type: message/rfc822-headers
Content-Disposition: inline

From: <somespammer@example.net>
Received: from mailserver.example.net
     (mailserver.example.net [10.67.41.167])
     by example.com with ESMTP id M63d4137594e46;
     Thu, 08 Mar 2005 14:00:00 -0400
To: <Undisclosed Recipients>
Subject: Earn money
MIME-Version: 1.0
Content-type: text/plain
Message-ID: 8787KJKJ3K4J3K4J3K4J3.mail@example.net
Date: Thu, 02 Sep 2004 12:31:03 -0500
--part1_13d.2e68ed54_boundary--
```

A.3 Full Report for Email Abuse with All Headers

```
From: <abusedesk@example.com>
Date: Thu, 8 Mar 2005 17:40:36 EDT
Subject: FW: Earn money
To: <abuse@example.net>
MIME-Version: 1.0
Content-Type: multipart/report; report-type=feedback-report;
     boundary="part1_13d.2e68ed54_boundary"

--part1_13d.2e68ed54_boundary
Content-Type: text/plain; charset="US-ASCII"
Content-Transfer-Encoding: 7bit

This is an email abuse report for an email message received from IP
10.67.41.167 on Thu, 8 Mar 2005 14:00:00 EDT. For more information
about this format please see http://www.mipassoc.org/arf/.

--part1_13d.2e68ed54_boundary
Content-Type: message/feedback-report

Feedback-Type: abuse
User-Agent: SomeGenerator/1.0
Version: 0.1
Original-Mail-From: <somespammer@example.net>
Original-Rcpt-To: <user@example.com>
Received-Date: Thu, 8 Mar 2005 14:00:00 EDT
Source-IP: 10.67.41.167
Authentication-Results: mail.example.com
               smtp.mail=somespammer@example.com;
               spf=fail
Reported-Domain: example.net
Reported-Uri: http://example.net/earn_money.html
Reported-Uri: mailto:user@example.com
Removal-Recipient: user@example.com

--part1_13d.2e68ed54_boundary
Content-Type: message/rfc822
Content-Disposition: inline

From: <somespammer@example.net>
Received: from mailserver.example.net (mailserver.example.net
     [10.67.41.167]) by example.com with ESMTP id M63d4137594e46;
     Thu, 08 Mar 2005 14:00:00 -0400
To: <Undisclosed Recipients>
Subject: Earn money
MIME-Version: 1.0
Content-type: text/plain
Message-ID: 8787KJKJ3K4J3K4J3K4J3.mail@example.net
Date: Thu, 02 Sep 2004 12:31:03 -0500

Spam Spam Spam
Spam Spam Spam
Spam Spam Spam
Spam Spam Spam
--part1_13d.2e68ed54_boundary--
```

B. Status of This Document [To Be Removed Upon Publication]

B.1 Discussion Venue

Discussion about this document should be directed to the ABUSE-FEEDBACK-REPORT mailing list <http://mipassoc.org/mailman/listinfo/abuse-feedback-report> which is also reachable via <mailto:abuse-feedback-report@mipassoc.org>. Of course, comments directly to the authors are always welcome (you can send them via email to <ietf@shaftek.org> and <drafts@domain-assurance.org>).

B.2 Document Repository and Public Website

Copies of this and earlier versions including multiple formats can be found at <http://www.shaftek.org/publications/drafts/abuse-report/>. A public website regarding this draft and related efforts is located at <http://mipassoc.org/arf/>.

B.3 Document History

Changes from draft-shafranovich-feedback-report-01-pre1 to draft-shafranovich-feedback-report-01:

- Added an "Outstanding Issues" section.
- Minor spelling mistakes and clarifications.
- Added links to previous work and more examples.
- Added three new types: "fraud" for phishing, "opt-out-list" for a single list opt out, and "other" as a catch-all.

Changes from draft-shafranovich-feedback-report-00 to draft-shafranovich-feedback-report-01-pre1:

- Changed the introduction section to clarify specific points that are out of scope for this document
- Added pointers to a public mailing list for discussion and public web page
- Clarified the intent section and added some extra points to it
- Added a reference to RFC 2119 and changed the document to comply
- Made it clear that the requirements section) is not the one defining the standard
- Clarified the main format section to make all three parts mandatory
- Changed section 4f regarding subject lines to mandate that subject lines should be left intact. Removed the convention for subject lines that was defined in the previous version
- Added text to the the machine readable section clarifying its intent. Also added RFC 2119 references, reorganized fields, indicated whether specific header fields can appear more than once and provided references as to how they should be formatted.
- Removed "Original-Message-ID", "Authenticated-Domain:" and "Authenticated-Domain-Method" from the draft including related IANA registries. Added "Version", "User-Agent", Original-Mail-From", "Original-Rcpt-To", "Reported-Uri", "Reported-Domain" and "Authentication-Results".
- Example has been updated to reflect new headers.
- Added a new section on extensibility and changed the IANA section to reflect that.

Changes from draft-shafranovich-abuse-report-00 to draft-shafranovich-feedback-report-00:

- Name of the format and report changed to 'feedback-report'
- Minor spelling corrections

- Added authentication headers and registry
- Added feedback-type header and registry.

Changes from draft-shafranovich-feedback-report-00 to draft-shafranovich-feedback-report-01:

- None significant (just a freshening)

Changes from draft-shafranovich-feedback-report-01 to draft-shafranovich-feedback-report-02:

- Much editorial cleanup
- Added John Levine and Paul Hoffman as co-authors
- Made the line lengths in Appendix A appropriate for RFCs
- Switched to symbolic names for references
- Reduced duplication of reference calls
- Removed text that specified the type of RFC and approval type that is expected
- Removed the requirement for an RFC to update the IANA registries; both are now designated expert approval only
- Added two new categories to the initial values for the "Feedback-Type" registry: miscategorized and not-spam
- Beefed up the Security Considerations section a bit
- Removed resolved questions from outstanding issues

B.4 Outstanding Issues

Here is a list of some outstanding issues for this document that have not been finalized:

- Whether encoding of the machine readable part should be limited to 7-bit
- Whether there is a need for both "opt-out" and "opt-out-list", and whether this format should be used for opt-outs at all.

Full Copyright Statement

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org[1].

Acknowledgement

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

[1] mailto:ietf-ipr@ietf.org