Internet Engineering Task Force (IETF)
Request for Comments: 7073
Category: Standards Track
ISSN: 2070-1721

N. Borenstein
Mimecast
M. Kucherawy

November 2013

A Reputation Response Set for Email Identifiers

Abstract

This document defines a response set for describing assertions a reputation service provider can make about email identifiers, for use in generating reputons.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7073.

Table of Contents

1. Introduction

This document specifies a response set for describing the reputation of an email identifier. A "response set" in this context is defined in [RFC7070] and is used to describe assertions a reputation service provider can make about email identifiers as well as metadata that can be included in such a reply beyond the base set specified there.

An atomic reputation response is called a "reputon", defined in [RFC7071]. That document also defines a media type to contain a reputon for transport, and creates a registry for reputation applications and the interesting parameters of each.

2. Terminology and Definitions

This section defines terms used in the rest of the document.

2.1. Key Words

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 [KEYWORDS].

2.2. Email Definitions

Commonly used definitions describing entities in the email architecture are defined and discussed in [EMAIL-ARCH].

2.3. Other Definitions

Other terms of importance in this document are defined in [RFC7070], the base document for the reputation services work.

3. Discussion

The expression of reputation about an email identifier requires extensions of the base set defined in [RFC7070]. This document defines and registers some common assertions about an entity found in a piece of [MAIL].

3.1. Assertions

The "email-id" reputation application recognizes the following assertions:

abusive: The subject identifier is associated with sending or handling email of a personally abusive, threatening, or otherwise harassing nature

fraud: The subject identifier is associated with the sending or handling of fraudulent email, such as "phishing" (some good discussion on this topic can be found in [IODEF-PHISHING])

invalid-recipients: The subject identifier is associated with delivery attempts to nonexistent recipients

malware: The subject identifier is associated with the sending or handling of malware via email

spam: The subject identifier is associated with the sending or handling of unwanted bulk email

For all assertions, the "rating" scale is linear: a value of 0.0 means there is no data to support the assertion, a value of 1.0 means all accumulated data support the assertion, and the intervening values have a linear relationship (i.e., a score of "x" is twice as strong of an assertion as a value of "x/2").

3.2. Response Set Extensions

The "email-id" reputation application recognizes the following OPTIONAL extensions to the basic response set defined in [RFC7071]:

email-id-identity: A token indicating the source of the identifier; that is, where the subject identifier was found in the message. This MUST be one of:

- dkim: The signing domain, i.e., the value of the "d=" tag, found on a valid DomainKeys Identified Mail [DKIM] signature in the message
- ipv4: The IPv4 address of the client
- ipv6: The IPv6 address of the client
- rfc5321.helo: The RFC5321.HELO value used by the client (see [SMTP])
- rfc5321.mailfrom: The RFC5321.MailFrom value of the envelope of the message (see [SMTP])
- rfc5322.from: The RFC5322.From field of the message (see [MAIL])
- spf: The domain name portion of the identifier (RFC5321.MailFrom or RFC5321.HELO) verified by [SPF]

sources: A token relating a count of the number of sources of data that contributed to the reported reputation. This is in contrast to the "sample-size" parameter, which indicates the total number of reports across all reporting sources.

A reply that does not contain the "identity" or "sources" extensions is making a non-specific statement about how the reputation returned was developed. A client can use or ignore such a reply at its discretion.

3.3. Identifiers

In evaluating an email message on the basis of reputation, there can be more than one identifier in the message needing to be validated. For example, a message may have different email addresses in the RFC5321.MailFrom parameter and the RFC5322.From header field. The RFC5321.Helo identifier will obviously be different. Consequently, the software evaluating the email message may need to query for the reputation of more than one identifier.

The purpose of including the identity in the reply is to expose to the client the context in which the identifier was extracted from the message under evaluation. In particular, several of the items listed are extracted verbatim from the message and have not been subjected to any kind of validation, while a domain name present in a valid DKIM signature has some more reliable semantics associated with it. Computing or using reputation information about unauthenticated identifiers has substantially reduced value, but can sometimes be useful when combined. For example, a reply that indicates a message contained one of these low-value identifiers with a high "spam" rating might not be worthy of notice, but a reply that indicates a message contained several of them could be grounds for suspicion.

A client interested in checking these weaker identifiers would issue a query about each of them using the same assertion (e.g., "spam"), and then collate the results to determine which ones and how many of them came back with ratings indicating content of concern, and take action accordingly. For stronger identifiers, decisions can typically be made based on a few or even just one of them.

3.4. Query Extensions

A query within this application can include the OPTIONAL query parameter "identity" to indicate which specific identity is of interest to the query. Legal values are the same as those listed in Section 3.2.

4. IANA Considerations

This memo presents one action for IANA, namely the registration of the reputation application "email-id".

4.1. Registration of ’email-id’ Reputation Application

This section registers the "email-id" reputation application, as per the IANA Considerations section of [RFC7071]. The registration parameters are as follows:

- Application symbolic name: email-id
- Short description: Evaluates DNS domain names or IP addresses found in email identifiers
- Defining document: [RFC7073]
- Status: current

- Subject: A string appropriate to the identifier of interest (see Section 3.2 of this document)

- Application-specific query parameters:

  identity: (current) as defined in Section 3.4 of this document

- Application-specific assertions:

  abusive: (current) as defined in Section 3.1 of this document

  fraud: (current) as defined in Section 3.1 of this document

  invalid-recipients: (current) as defined in Section 3.1 of this document

  malware: (current) as defined in Section 3.1 of this document

  spam: (current) as defined in Section 3.1 of this document

- Application-specific response set extensions:

  identity: (current) as defined in Section 3.2 of this document

5. Security Considerations

This document is primarily an IANA action and doesn't describe any protocols or protocol elements that might introduce new security concerns.

Security considerations relevant to email and email authentication can be found in most of the documents listed in the References sections below. Information specific to use of reputation services can be found in [CONSIDERATIONS].

6. References

6.1. Normative References

[DKIM] Crocker, D., Ed., Hansen, T., Ed., and M. Kucherawy, Ed., "DomainKeys Identified Mail (DKIM) Signatures", STD 76, RFC 6376, September 2011.

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

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

[RFC7070] Borenstein, N., Kucherawy, M., and A. Sullivan, "An Architecture for Reputation Reporting", RFC 7070, November 2013.

[RFC7071] Borenstein, N. and M. Kucherawy, "A Media Type for Reputation Interchange", RFC 7071, November 2013.

[SMTP] Klensin, J., "Simple Mail Transfer Protocol", RFC 5321, October 2008.

[SPF] Wong, M. and W. Schlitt, "Sender Policy Framework (SPF) for Authorizing Use of Domains in E-Mail, Version 1", RFC 4408, April 2006.

6.2. Informative References

[CONSIDERATIONS] Kucherawy, M., "Operational Considerations Regarding Reputation Services", Work in Progress, May 2013.

[IODEF-PHISHING] Cain, P. and D. Jevans, "Extensions to the IODEF-Document Class for Reporting Phishing", RFC 5901, July 2010.

[MAIL] Resnick, P., Ed., "Internet Message Format", RFC 5322, October 2008.

Appendix A. Positive vs. Negative Assertions

[CONSIDERATIONS] some current theories about reputation, namely that it will possibly have more impact to develop positive reputations and focus on giving preferential treatment to content or sources that earn those. However, the assertions defined in this document are all clearly negative in nature.

In effect, this document is recording current use of reputation and of this framework in particular. It is expected that, in the future, the application being registered here will be augmented, and other applications registered, that focus more on positive assertions rather than negative ones.

Appendix B. Acknowledgments

The authors wish to acknowledge the contributions of the following to this specification: Scott Hollenbeck, Scott Kitterman, Peter Koch, John Levine, Danny McPherson, S. Moonesamy, Doug Otis, and David F. Skoll.

Authors' Addresses

Nathaniel Borenstein
Mimecast
203 Crescent St., Suite 303
Waltham, MA 02453
USA

Phone: +1 781 996 5340
EMail: nsb@guppylake.com

Murray S. Kucherawy
270 Upland Drive
San Francisco, CA 94127
USA

EMail: superuser@gmail.com