Network Working Group
Internet-Draft
Intended status: Informational
Expires: June 1, 2012

N. Borenstein
Mimecast
M. Kucherawy
Cloudmark

November 29, 2011

A Model for Reputation Interchange
draft-ietf-repute-model-00

Abstract

This document describes the general model underlying a set of proposals for the exchange of reputation information on the Internet, and provides a roadmap to the four additional documents that collectively define a reputation interchange protocol. It is intended roughly to follow the recommendations of RFC4101 for describing a protocol model.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on June 1, 2012.

Table of Contents

1. Introduction

Traditionally, most Internet protocols have taken place between unauthenticated entities. For example, when an email message is submitted via [SMTP], the server typically trusts the self-identification of the sender, and the sender trusts that the [DNS] has led it to the right server. Both kinds of trust are easily betrayed, leading to spam, phishing, and a host of other ills.

In recent years, stronger identity mechanisms have begun to see wider deployment. For example, the [DKIM] protocol permits a much higher level of trust in the identity of the sending domain of an email message. While this is a major step forward, by itself it does little to solve the problem of bad actors on the Internet. Even if you can be sure a message comes from a domain called "trustworthy.example.com," you don't really know whether or not that domain is trustworthy. As a practical matter, the bad actors seem to have adopted DKIM even more rapidly than the good ones, in the hope that some receiving domains will naively confuse a confirmation of identity with trustworthiness.

The next step, which could usefully be undertaken only in the presence of such stronger identity mechanisms, is to establish a mechanism by which mutually trusted parties can exchange information about other parties. Such information is known as reputation information.

While the need for reputation information has been most clear in the email world, where abuses are commonplace, it is easy to think of additional uses for such information. It could also be useful in rating the security of web sites, the quality of service of an Internet Service Provider (ISP) or Application Service Provider (ASP), the customer satisfaction levels at e-commerce sites, and even things unrelated to Internet protocols, such as rating plumbers, hotels, or books. Just as human beings traditionally rely on the recommendations of trusted parties in the physical world, so too they can be expected to make use of such reputation information in a variety of applications on the Internet.

Accordingly, this protocol is designed to facilitate a wide range of reputation applications. However, not all such reputations will need to convey the same information. An overall reckoning of goodness versus badness can be defined generically, but specific applications are likely to want to describe reputations for multiple attributes; an e-commerce site might be rated on price, speed of delivery, customer service, etc., and might receive very different ratings on each. Therefore, this protocol defines a generic mechanism and basic format for reputation information, while allowing extensions for each

application.

Omitted from this specification is the way by which an agent that wishes to report reputation information regarding something goes about collecting such data. The protcol defined in this document and its companion documents is merely about asking a question and getting an answer; the remainder of the overall service provided by such an agent is specific to the implementation of that service and is out of scope here.

2. Document Series

This memo represents the base specification, introducing a series of others that define the overall service and introduce the initial exemplary applications. The series is as follows:

1. RFCxxxx: A Model for Reputation Interchange (this memo)

2. RFCxxxx+1: Using the DNS for Reputation Interchange

3. RFCxxxx+2: Using UDP for Reputation Interchange

4. RFCxxxx+3: Using the DNS for Reputation Interchange

5. RFCxxxx+4: Using HTTP/XML for Reputation Interchange

6. RFCxxxx+5: A Reputation Vocabulary for Email Identity Reputation

7. RFCxxxx+6: A Reputation Vocabulary for Email Property Reputation

3. Terminology and Definitions

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

3.1. Keywords

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].

3.2. Vocabulary

A "vocabulary" comprises those data that are returned in response to a reputation query about a particular entity. The vocabulary is specific to an application; the data returned in the evaluation of email senders would be different than the reputation data returned

about a movie or a baseball player.

Vocabularies have symbolic names, and these have to be registered with IANA to prevent name collisions. The IANA registries are created in a separate memo.

4. Information Represented in the Protocol

The basic information to be represented in the protocol is fairly simple, and MUST include:

- the identity of the entity providing the reputation information;
- the level of confidence in that identity being genuine;
- the identity of the entity being rated;
- the overall rating score for that entity; and
- the number of data points underlying that score.

Beyond this, arbitrary amounts of additional information might be represented for specific applications of the protocol. Such information is called the "vocabulary" for that application. The general protocol defines a syntax for representing such vocabularies, but each application will define its own vocabulary. Thus, the basic information MUST also include:

- the name of the application for which the reputation data is being expressed.

For example, a subsequent document will define the reputation vocabulary for the application "email-sending-domain" which will be used to combat spam and other abuses of email. Additional documents define a [MIME] type for reputation data, and protocols for exchanging such data.

5. Information Flow in the Protocol

The basic reputation data represented in the new [MIME] media type can be transported in any number of ways, like any MIME object. However, it is anticipated that the typical use of the protocol will be a simple request/response. One entity will ask a second entity for reputation data about a third entity, and the second entity will respond with that data.

It is anticipated that a few applications, at least including the email-sending-domain application, will need a small, lightweight protocol for such queries and responses, while other applications will need to be able to retrieve larger and more complex responses. For this reason, two subsequent documents define two such protocols, one based on DNS queries and a terse representation, and one based on [HTTP] queries with an XML representation.

6. IANA Considerations

This memo presents no actions for IANA, though later memos in this series are likely to do so.

7. Security Considerations

This memo introduces an overall protocol model, but no implementation details. As such, the security considerations presented here are very high-level. The detailed analyses of the various specific components of the protocol can be found in the subsequent documents enumerated in Section 2.

7.1. Biased Reputation Agents

As with [VBR], an agent seeking to make use of a reputation reporting service is placing some trust that the service presents an unbiased "opinion" of the object about which reputation is being returned. The result of trusting the data is, presumably, to guide action taken by the reputation client. It follows, then, that bias in the reputation service can adversely affect the client. Clients, therefore, should be aware of this possibility and the impact it might have. For example, a biased system returning reputation information about a DNS domain found in email messages could result in the admission of spam, phishing or malware through a mail gateway.

Clients might also seek to interact only with reputation services that offer some level of transparency into the computation of the results they return. How this might be evaluated, however, is not specified here.

Similarly, a client placing trust in the results returned by such a service might suffer if the service itself is compromised, returning biased results under the control of an attacker without the knowledge of the agency providing reputation service. This might result from an attack on the data being returned at the source, or from a man-in-the-middle attack. Protocols, therefore, should be designed so as to be as resilient against such attacks as possible.

7.2. Malformed Messages

Both clients and servers of reputation systems need to be resistant to attacks that involve malformed messages, deliberate or otherwise. Failure to do so creates an opportunity for a denial-of-service.

8. Informative References

[DKIM] Allman, E., Callas, J., Delany, M., Libbey, M., Fenton, J., and M. Thomas, "DomainKeys Identified Mail (DKIM) Signatures", RFC 4871, May 2007.

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

[HTTP] 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.

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

[MIME] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies", RFC 2045, November 1996.

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

[VBR] Hoffman, P., Levine, J., and A. Hathcock, "Vouch By Reference", RFC 5518, April 2009.

Appendix A. Public Discussion

Public discussion of this suite of memos takes place on the domainrep@ietf.org mailing list. See https://www.ietf.org/mailman/listinfo/domainrep.

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
Cloudmark
128 King St., 2nd Floor
San Francisco, CA 94107
USA

Phone: +1 415 946 3800
Email: msk@cloudmark.com