Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 23, 2012

N. Borenstein
Mimecast
M. Kucherawy
Cloudmark

October 21, 2011

Reputation Data Interchange using HTTP and XML
draft-kucherawy-reputation-query-http-03

Abstract

This document defines a mechanism to conduct queries for reputation information using the Domain Name System.

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 April 23, 2012.

Table of Contents

1. Introduction

This memo defines a method to query a reputation data service for information about an entity, using the HyperText Transfer Protocol (HTTP) as the transport mechanism and XML as the payload format. It is part of a series defining the overall reputation query/response structure as well as the concept of reputation "vocabularies" for particular applications.

2. Document Series

This memo represents the media type registration, part of a series of documents that define the overall service and introduce the initial exemplary applications. The series is as follows:

1. RFCxxxx: A Model for Reputation Interchange

2. RFCxxxx+1: A Media Type for Reputation Information

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 (this memo)

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. Other Definitions

Other terms of importance in this memo are defined in RFCxxxx, the base memo in this document series.

4. Description

4.1. Query

A reputation query made via [HTTP] encodes the question being asked partly in the [URI] and partly within the GET instruction of the protocol.

The components to the question being asked comprise the following:

- The subject of the query;
- The name of the host, or the IP address, at which the reputation service is available;
- The name of the reputation application, i.e., the context within which the query is being made;
- Optionally, name(s) of the specific reputation assertions or attributies that are being requested.

The name of the application MUST be one registered with IANA. A server receiving a query about an unregistered application or one it does not explicitly support MUST return a 404 error code.

The syntax for the URI portion of the query is constructed using a template as per [URI-TEMPLATE]. The following variables MUST be available during template expansion:

application: The name of the application reputation in whose context the request is being made.

scheme: The transport scheme the client will be using for the query.

service: The hostname or IP address being queried.

Which scheme(s) can be used depends on how the reputation service provider offers its services. Thus, the template could include a specific schema as a fixed string in the template, or it might offer it as a variable in the template. If it is a variable, it is up to the client and server to negotiate out-of-band which schemes are supported for client queries. Implementers should be aware that the template could include a fixed scheme not supported by the client.

The following variables are OPTIONAL, but might be required by the template presented for a specific service:

assertion: A list of one or more specific assertions of interest to the client. If absent, the server MUST infer that all available assertion information is being requested.

passwd: The "password" portion of a client credential.

user: The "user" portion of a client credential.

Other required or optional query parameters might be defined by documents that register new vocabularies with IANA.

The template is retrieved by requesting the [WELL-KNOWN-URI] "repute_template" from the host providing reputation service using HTTP. If the template cannot be retrieved, the query should be aborted and/or retried at a later time.

For example, given the following template:

{scheme}://{service}/{application}/{subject}/{assertion}

A query about the use of the domain "example.org" in the "email-id" application context to a service run at "example.com", where that application declares a required "subject" parameter, requesting the "SENDS-SPAM" reputation assertion using HTTP to conduct the query with no specific client authentication information would be formed as follows:

http://example.com/email-id/example.org/sends-spam

Matching of the attribute name(s) MUST be case-insensitive.

4.2. Response

The response is expected to be an XML document. The "format" parameter of the "application/reputon" media type MUST be "xml" when used in this mode.

The XML schema definition describing the format of that response is included below.

4.2.1. XML Schema

The following XML schema describes the format of the reply:

```
<?xml version="1.0" encoding="ISO-8859-1" ?%gt;
 <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

 <!-- definition of local types -->
 <xs:simpleType name="exttype">
  <xs:restriction base="xs:token">
   <xs:pattern value="\w+(-\w*)*:\s?[\w\p{P}]+"/>
  <xs:/restriction>
 <xs:/simpleType>

 <!-- definition of simple elements -->
 <xs:element name="rater" type="xs:token"/>
 <xs:element name="rater-authenticity" type="xs:decimal"/>
 <xs:element name="assertion" type="xs:token"/>
 <xs:element name="extension" type="exttype"/>
 <xs:element name="rated" type="xs:token"/>
 <xs:element name="rating" type="xs:decimal"/>
 <xs:element name="sample-size" type="xs:positiveInteger"/>
 <xs:element name="updated" type="xs:positiveInteger"/>

 <!-- definition of complex elements -->
 <xs:complexType name="assertiontype">
  <xs:sequence>
   <xs:element ref="rater" minOccurs="1"/>
   <xs:element ref="rater-authenticity" minOccurs="1"/>
   <xs:element ref="assertion" minOccurs="1"/>
   <xs:element ref="extension"/>
   <xs:element ref="rated" minOccurs="1"/>
   <xs:element ref="rating" minOccurs="1"/>
   <xs:element ref="sample-size" minOccurs="1"/>
   <xs:element ref="updated" minOccurs="1"/>
  <xs:/sequence>
 <xs:/complexType>

 <xs:complexType name="reporttype">
  <xs:sequence>
   <xs:element name="reputon" type="assertiontype"
                 maxOccurs="unbounded" minOccurs="1"/>
  <xs:/sequence>
 <xs:/complexType>

 <xs:element name="reputation" type="reporttype"/>
</xs:schema>
```

The elements that comprise an "assertion" are used as follows:

rater: The identity of the agent making the assertion.

rater-authenticity: An expression by the rater of its confidence in the report it is giving. Expressed as a decimal value between 0 and 1 inclusive.

assertion: The assertion being made. This MUST be an assertion registered within the specified application by IANA.

extension: (OPTIONAL) One or more application-specific vocabulary extensions and their corresponding values. If present, each of these MUST be a vocabulary extension registered with IANA.

rated: The identity about which an assertion is being made.

rating: The value of the assertion. This is a decimal number from 0 to 1, with 0 meaning the assertion is completely false (according to the agent making the assertion) and 1 meaning the assertion is completely true.

sample-size: The count of data points the asserting agent used to produce the value provided in the previous element.

updated: The time at which the current rating was computed. Expressed in number of seconds since 00:00:00 UTC, January 1, 1970.

4.2.2. Example Reply

The following is an example reputon generated using the above schema, including the media type definition line:

```
Content-Type: application/reputon; app="email"; format="xml"

<?xml version="1.0" encoding="US-ASCII"?>

<reputation>
 <reputon>
  <rater>rep.example.net</rater>
  <rater-authenticity>0.95</rater-authenticity>
  <assertion>SENDS-SPAM</assertion>
  <extension>IDENTITY: DKIM</extension>
  <rated>example.com</rated>
  <rating>0.0012</rating>
  <sample-size>16938213</sample-size>
  <updated>1317795852</updated>
 </reputon>
</reputation>
```

Here, reputation agent "rep.example.net" is asserting within the context of email that "example.com" appears to send spam 1.2% of the time, based on just short of 17 million messages analyzed or reported to date. The identity "example.com", the subject of the query, is extracted from the analyzed messages using the [DKIM] "d=" parameter for messages where signatures validate. The reputation agent is 95% confident of this result. (See [RFCxxxx+5] for details about the registered email vocabulary.)

5. IANA Considerations

This memo presents no actions for IANA. Registration of the well-known URI "repute_template" will be done as defined in [WELL-KNOWN-URI] which is not a function of IANA.

6. Security Considerations

This memo describes security considerations introduced by the media type defined here.

6.1. General

This memo is part of a series introducing a reputation query and response system (see Section 2). The Security Considerations sections of the other memos should also be consulted.

7. References

7.1. Normative References

[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.

[URI] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", RFC 3986, January 2005.

[URI-TEMPLATE] Gregorio, J., Fielding, R., Hadley, M., Nottingham, M., and D. Orchard, "URI Template",

I-D draft-gregorio-uritemplate, September 2011.

[WELL-KNOWN-URI]
Nottingham, M. and E. Hammer-Lahav, "Defining Well-Known Uniform Resource Identifiers (URIs)", RFC 5785, April 2010.

7.2. Informative References

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

Appendix A. Acknowledgements

The authors would like to thank the following for their contributions to this work: Mark Nottingham.

Appendix B. 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