

RUCUS Exploratory Working Group  
Internet-Draft  
Intended status: Experimental  
Expires: August 26, 2008

D. Wing  
Cisco  
S. Niccolini  
M. Stiernerling  
NEC  
H. Tschofenig  
Nokia Siemens Networks  
February 23, 2008

Spam Score for SIP  
draft-wing-sipping-spam-score-02

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on August 26, 2008.

Abstract

This document defines a mechanism for SIP proxies to communicate a spam score to downstream SIP proxies and to SIP user agents. This information can then be used as input to other decision making engines, for example, to provide alternate call routing or call handling.

Internet-Draft

SIP Spam Score

February 2008

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">3</a>
<a href="#">2.</a>	Terminology . . . . .	<a href="#">3</a>
<a href="#">3.</a>	Calculation of the Spam Score . . . . .	<a href="#">3</a>
<a href="#">4.</a>	Information passed downstream . . . . .	<a href="#">4</a>
<a href="#">5.</a>	Operation of Spam-Scoring Proxy . . . . .	<a href="#">5</a>
<a href="#">6.</a>	Operation of Downstream Proxy or User Agent . . . . .	<a href="#">6</a>
<a href="#">7.</a>	Grammar . . . . .	<a href="#">7</a>
<a href="#">8.</a>	Examples . . . . .	<a href="#">8</a>
<a href="#">9.</a>	Security Considerations . . . . .	<a href="#">8</a>
<a href="#">10.</a>	Acknowledgements . . . . .	<a href="#">9</a>
<a href="#">11.</a>	IANA Considerations . . . . .	<a href="#">9</a>
<a href="#">12.</a>	References . . . . .	<a href="#">9</a>
<a href="#">12.1.</a>	Normative References . . . . .	<a href="#">9</a>
<a href="#">12.2.</a>	Informational References . . . . .	<a href="#">9</a>
<a href="#">Appendix A.</a>	Changes . . . . .	<a href="#">10</a>
<a href="#">A.1.</a>	Changes from -00 to -01 . . . . .	<a href="#">10</a>
<a href="#">A.2.</a>	Changes from -01 to -02 . . . . .	<a href="#">10</a>
	Authors' Addresses . . . . .	<a href="#">10</a>
	Intellectual Property and Copyright Statements . . . . .	<a href="#">12</a>

Internet-Draft

SIP Spam Score

February 2008

## 1. Introduction

It is desirable for SIP proxies to insert a spam score so that downstream SIP proxies and downstream SIP user agents can use a high score to decide that special handling is required. For example, a score above 20 might cause one of the spam avoidance techniques described in [[RFC5039](#)] to be triggered for this call.

This specification allows each SIP proxy to contribute spam scoring information that can be useful to downstream SIP proxies and the SIP user agent (UA). The downstream SIP proxies or SIP UA might ignore that information (e.g., it doesn't trust the SIP proxy that generated the spam score) or might use it.

Note that this document does not make the attempt to define how the spam score was derived nor to distribute information that could be used to verify the spam score generation. Furthermore, this document does not attempt to cryptographically bind the identity of the entity generating the score to the value itself. Hence, its usage is likely to be useful only between neighboring administrative domains communicating such a score.

One may wonder why bother marking a message that appears to be SPAM when the same process that detected the SPAM can also automatically block it. The answer is that contractual as well as regulatory issues may prevent blocking and in these cases while not able to block, the detecting proxy can nonetheless notify downstream elements of the potential threat.

## 2. Terminology

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 [[RFC2119](#)].

### 3. Calculation of the Spam Score

A SIP proxy evaluates an incoming SIP request and generates a spam score using a local mechanism. In order to allow for whitelisting as well as blacklisting the scoring is between 0 and 100, 0 indicating absolute acceptance (e.g., whitelist), 100 indicating absolute SPAM (e.g., blacklist) and scores between 0 and 100 can be considered to represent the percentage likelihood of spam.

The actual calculation is governed by algorithms (one example is found in examples section below) which MAY be agreed upon by the

upstream and downstream domains. The algorithm MAY be conveyed by the downstream domains to the upstream one out of band prior to the upstream domain marking a message for transport downstream. Alternatively, a default algorithm can be used if no alternate algorithm was established a-priori between upstream and downstream domains. The mechanism for conveyance of algorithm to upstream domain is out of scope for this document but can be seen as an extension to [[I-D.tschofenig-sipping-framework-sip-reduction](#)].

### 4. Information passed downstream

In addition to the score the following other pieces of information should be passed downstream as well:

Realm - Indicating the upstream domain or realm making the claim

Algorithm - The name of the agreed upon algorithm

Strength - An integer indicating the confidence of the score (0-100)

Info - A text field containing any arbitrary information

Param[1..3] - 3 general purpose parameters for futureproofing

IsSpam - A boolean for convenience purposes alone.

The Realm is shared by all proxies in a domain and enables the

downstream proxy to determine whether or not the score is to be trusted.

The Algorithm is the name of the algorithm used and can be standardized in the same way encryption algorithm names (e.g. sha-1) have been standardized.

The Strength field indicates the confidence level of the Score. This is generated by the same entity which generates the Score, and is an output from the algorithm, which is based on a number of factors, including for example volume of calls, call type etc. It is meant to quantify the score. A calling party that makes many calls and has an average score of 85 is preferable to calling party who made 1 or 2 calls only and has an average score of 95.

Info is a text field (which can be limited to 64 bytes if we are concerned about the MTU) that is there to provide more complete information on the SPIT scoring etc, and can be used for diagnosis etc. (which is useful for off line analysis reporting etc.). For

example, it could provide an explanation of the score such as is done in one particular email spam package.

In this email example the threshold for SPAM is 7 and this message scored a 9.5. The sample text info below could be used to explain how the score was calculated.

points	rule-name	rule-description
-----	-----	-----
2.5	SUBJ_CAPS	Subject line is capitalized
1.8	INLINE_GIF	inline GIF detected in message
3.0	NON_PREF_LANG	Non default language
2.0	HONEYPOT	Honeypot address in cc
0.2	KEYWORD_MATCH	suspicious words in message

Figure 1: Email Spam Rule example

As expected in this example, different tests have different scores depending on their contribution to the potential of SPAM. Similarly in the case of SPIT there can be many rules applied and having this info can enable the receiving party to analyze the results (non realtime)

Params 1,2 and 3 are just general purpose placeholders (containing Param\_Desc, Param\_Value pairs) for future proofing capabilities. Since so much is still unknown about IP communications SPAM this is seen as a wise approach.

Finally, as an option it may be wise to have a boolean yes/no kind of indicator for all those downstream who do not care to know why the upstream element assumes it is spit only that it is.

## [5.](#) Operation of Spam-Scoring Proxy

A SIP proxy evaluates an incoming SIP request and generates a spam score using a local mechanism. This score is between 0 (indicating the message is not spam) and 100 (indicating the message is spam). Values between 0 and 100 indicate the 'likelihood' that the SIP request is spam, with higher values indicating a higher likelihood the message is spam.

This spam score is inserted into the new "Spam-Score" header. This header field contains a summary spam score and optionally contains detail information. The detail information is implementation dependent. The detail information is valuable for debugging and to provide the SIP user agent or SIP proxy with additional information regarding how the spam-scoring SIP proxy's local mechanism arrived at

the summary spam score.

## [6.](#) Operation of Downstream Proxy or User Agent

A downstream proxy or the SIP user agent MAY use the spam score or spam-detail information to change call routing or call handling. It is envisioned that some form of policies indicate the trusted proxies in order to decide which spam scores to consider for special call treatment.

In some jurisdictions, the end user needs to authorize call handling, including rejection of a call based on a spam score. Mechanisms to allow users to authorize such policies are, however, out of scope of this document.

The behavior of the SIP proxy or user agent when the spam score is above a certain value is a local policy matter. Examples of behavior include:

- o a SIP request with a high spam score might cause a proxy or user agent to redirect the SIP request to company's main telephone extension or to the user's voicemail
- o a user agent might alert the user by flashing the phone (without audible ringing)
- o a user agent might allow calls with a spam score below a certain value during daylight hours, but deny such calls at night.
- o a proxy might challenge the caller to complete a Turing test.

## [7.](#) Grammar

ABNF using the ABNF syntax of [[RFC3261](#)]:

```
extension-header = "Spam-Score:" SP
                  spam-score *[";" spam-detail ]
```

```

spam-score      = score SP "by" SP hostname
score           = 1*3DIGIT [ "." 1*3DIGIT ]

spam-detail     = spam-strength / spam-algorithm / spam-param

spam-algorithm  = "spam-algorithm" EQUAL quoted-string

spam-strength   = "spam-score-strength" EQUAL strength
strength        = 1*3DIGIT [ "." 0*3DIGIT ]

spam-info       = "spam-info" EQUAL info-value
info-value      = quoted-string

spam-param1     = "spam-param1" EQUAL param-value
param-value     = quoted-string

spam-param2     = "spam-param2" EQUAL param-value
param-value     = quoted-string

spam-param3     = "spam-param3" EQUAL param-value
param-value     = quoted-string

spam-isspam     = [ "isSpam" ]

```

Figure 2: ABNF



The following example shows a SIP score generated and inserted by two SIP proxies, sip.example.com and sip.example.net. In this example, sip.example.com is owned by a spammer who is trying to fool downstream systems with their low spam score (0). However, the example.net proxies and user agents only pay attention to spam scores from Spam-Score headers generated by example.net proxies, so example.com's attempts to fool the downstream proxies (with its low spam score) are in vain. Note also the sample Session Duration Time (SDT) algorithm SDT [[I-D.malas-performance-metrics](#)] simply compares the given callers previous session duration time with the expected session duration time over all destinations.

```
INVITE sip:bob@example.net SIP/2.0
Via: SIP/2.0/UDP sip.example.net;branch=z9hG4bKnashds8
    ;received=192.0.2.1
Spam-Score: 75 by sip.example.net
    ;detail="SIPfilter-1.0;call_volume=75"
    ;spam-algorithm="SDT"
    ;spam-score-strength=50
    ;spam-info="High call volume"
    ;spam-isSpam
Via: SIP/2.0/UDP sip.example.com;branch=z9hG4bKfjzc
    ;received=192.0.2.127
Max-Forwards: 70
To: Bob <sip:bob@example.net>
From: Alice <sip:alice@example.com>;tag=1928301774
Call-ID: a84b4c76e66710@pc33.example.com
CSeq: 314159 INVITE
Contact: <sip:alice@pc33.example.com>
Content-Type: application/sdp
Content-Length: 142

[... SDP elided from this example...]
```

Figure 3: Example with spam scores

## [9.](#) Security Considerations

SIP proxies and SIP user agents need to ignore spam scores generated by proxies that aren't trusted. As the spam scores are inserted along with Via: headers, the last Via header inserted by a trusted proxy indicates the last trusted spam score.

In addition, the entire issue of securing the channel between the upstream and downstream domains MUST be addressed via mechanisms such

as TLS.

## 10. Acknowledgements

Thanks to Joachim Charzinski, Daniel Quinlan, and S. Moonesamy for their suggestions to improve this document. Thanks to David Schwartz for his contributed text and to Eli Katz for editing assistance.

## 11. IANA Considerations

[[This section will be completed in a later version of this document.]]

## 12. References

### 12.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", [RFC 3261](#), June 2002.

### 12.2. Informational References

- [RFC5039] Rosenberg, J. and C. Jennings, "The Session Initiation Protocol (SIP) and Spam", [RFC 5039](#), January 2008.
- [I-D.tschofenig-sipping-framework-spit-reduction]  
Tschofenig, H., Schulzrinne, H., Wing, D., Rosenberg, J., and D. Schwartz, "A Framework to tackle Spam and Unwanted Communication for Internet Telephony", [draft-tschofenig-sipping-framework-spit-reduction-02](#) (work in progress), November 2007.
- [I-D.malas-performance-metrics]  
Malas, D., "SIP End-to-End Performance Metrics", [draft-malas-performance-metrics-08](#) (work in progress), December 2007.

Internet-Draft

SIP Spam Score

February 2008

## [Appendix A](#). Changes

Note to RFC Editor: please remove this section prior to publication.

### [A.1](#). Changes from -00 to -01

- o Changed scoring from positive/negative to 0-100 range.
- o Moved score from a "Via:" extension to a new header "Spam-Score:".
- o Changed from Standards Track to Experimental.

### [A.2](#). Changes from -01 to -02

- o Describe how spam score could be computed
- o Added more descriptive text describing how the header is passed downstream towards the user agent

## Authors' Addresses

Dan Wing  
Cisco Systems, Inc.  
170 West Tasman Drive  
San Jose, CA 95134  
USA

Email: [dwing@cisco.com](mailto:dwing@cisco.com)

Saverio Niccolini  
Network Laboratories, NEC Europe Ltd.  
Kurfuersten-Anlage 36  
Heidelberg 69115  
Germany

Phone: +49 (0) 6221 4342 118  
Email: [saverio.niccolini@netlab.nec.de](mailto:saverio.niccolini@netlab.nec.de)

URI: <http://www.netlab.nec.de>

Wing, et al.

Expires August 26, 2008

[Page 10]

---

Internet-Draft

SIP Spam Score

February 2008

Martin Stiemerling  
Network Laboratories, NEC Europe Ltd.  
Kurfuersten-Anlage 36  
Heidelberg 69115  
Germany

Phone: +49 (0) 6221 4342 113  
Email: [stiemerling@netlab.nec.de](mailto:stiemerling@netlab.nec.de)  
URI: <http://www.netlab.nec.de>

Hannes Tschofenig  
Nokia Siemens Networks  
Linnoitustie 6  
Espoo 02600  
Finland

Phone: +358 (50) 4871445  
Email: [Hannes.Tschofenig@nsn.com](mailto:Hannes.Tschofenig@nsn.com)  
URI: <http://www.tschofenig.com>

Internet-Draft

SIP Spam Score

February 2008

#### Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

#### 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](mailto:ietf-ipr@ietf.org).

#### Acknowledgment

This document was produced using xml2rfc v1.33pre8 (of <http://xml.resource.org/>) from a source in [RFC-2629](#) XML format.