SIPPING Working Group	J. Seedorf
Internet-Draft	S. Niccolini
Intended status: Informational	NEC
Expires: August 21, 2008	H. Schulzrinne
	Columbia University
	February 18, 2008

TOC

Spam score for SIP: a proposal for semantics
draft-seedorf-sipping-spam-score-semantics-00

Status of this Memo

Abstract

This document reports a proposal for semantics of SIP spam scoring in order to achieve a flexible signalling standardization allowing an incremental adoption of the scoring mechanism. This approach can give early experimental implementers the possibility to start using spam scoring extensions in an explorative fashion without running into interoperability problems.

Table of Contents

---

1. Introduction

TOC

Latest discussion in the IETF demonstrated that there is still a lack of consensus how to address the general topic of SPIT mitigation. In particular, many controversial discussions have been centered around the SIP spam score draft [I-D.wing-spam-score] (Wing, D., Niccolini, S., Stiemerling, M., and H. Tschofenig, “Spam Score for SIP,” February 2008.), as well as the mechanisms and the rationale behind it.
The main issues raised were:

- *uncertainness on the appearance of the threat;
- *unknown effectiveness of mitigation algorithms;
- *lack of semantics and transmission of SPIT estimation scores;

Even though the spam threat is not fully defined today, the recommendation of [RFC5039] (Rosenberg, J. and C. Jennings, “The Session Initiation Protocol (SIP) and Spam,” January 2008.) is to not wait until it is too late (i.e., providers should not ignore the spam problem until it happens); there is the need for some work in this area.
Even if [RFC5039] (Rosenberg, J. and C. Jennings, “The Session Initiation Protocol (SIP) and Spam,” January 2008.) indicated a possible path, spam is such a multifaced problem that this cannot be regarded as the only one; multiple paths must be explored and standardization bodies should give implementers the possibility to experiment with solutions before the problem gets too big (as it got in the email case).

Given that something needs to be done, this document details a proposal for dealing with the remaining two issues, namely how to give implementers a chance to start experimenting with SPIT mitigation mechanisms and to communicate spam score results across different entities in the network in an interoperable and incremental way.

2. SIP spam score semantics proposal

TOC

2.1. Proposal motivation

TOC

The main points that motivated us to write such a proposal and made us believe that spam score is an important mechanisms for spam mitigation were:

- Whether a message is spam or not is not a binary proposition, both in terms of identifying it (mechanisms will have false positives and false negatives) and even in judging it (e.g., depending on user preferences).

- Different SIP routing elements have different types of information available that are useful for spam identification, but are not necessarly the ones that should be making call handling decisions.

- End systems or user services implemented in proxies should be judging this information and make a decision as to how to handle the call, i.e, whether to reject, forward or present the call to the user and what user interface indications to provide to the user.

For interoperability of such spam scores being exchanged among SIP entities it is absolutely necessary to have semantics defined. If no clear semantics are defined for spam scores, there is the risk of entities falsely interpreting scores they receive. Since every SPIT mitigation technique works differently, we propose to have semantics defined "per-method" and not in general.

2.2. Proposal details

TOC

We propose to have SIP signalling extensions allowing the binding of SIP spam scores to well defined semantics. Such a solution would allow

the possibility of making network-wide distributed decisions across multiple entities involved in SPIT mitigation decisions.
Even though spam is not a binary proposition, some currently suggested mitigation mechanisms give a 0/1 result when being applied to a message. Still, such an outcome is only an indicator for a message being spam or not. Defining semantics for SPIT mitigation mechanisms with such a 0/1 output (i.e., a binary output) is a matter of assigning 0 and 1 to specified outputs.
Thus, methods giving a "binary output" can have very straightforward semantics:

*blacklist/whitelist: 0 means "not on list", 1 means "on list";

*timecontext: 0 means "the caller initiated a session inside the user-defined interval for receiving calls", 1 means "the caller initiated a session inside the user-defined interval for receiving calls";

*captcha: 0 means "the caller passed the CAPTCHA test", 1 means "the caller did not pass the CAPTCHA test".

*etc.

Each binary method would need to standardize the "methodID" (e.g., the name) and the corresponding "semantic" (the meaning of 0/1, respectively). In principle, these methods could well be mapped to policies, see [I-D.tschofenig-sipping-spit-policy] (Tschofenig, H., Wing, D., Schulzrinne, H., Froment, T., and G. Dawirs, “A Document Format for Expressing Authorization Policies to tackle Spam and Unwanted Communication for Internet Telephony,” July 2008.) and reference within.
Other mechanisms currently proposed for SPIT mitigation have a more detailed output (which still only gives an indication for SPIT). These mechanisms need well-defined semantics as a basis for interoperability as well. Such methods giving a "non-binary output" could have more elaborate semantics based on statistics:

*statistics based on user feedback; such methods would give an estimation of a score x that could be the percentage of messages with the same method-result that were marked as SPIT by users;

*statistics based on anomaly detection; such methods would give an estimation of a score x that could be the percentage of previous messages were below/above (depending on the method) the result for this method compared to the current message.

*etc.

Also in this case, each non-binary method would need to standardize the "methodID" (e.g., the name) and the corresponding "semantic" (the meaning of the x score).
The proposal allows a per-method score in order to have different methods with different ranges. This flexibility enables the use of new detection methods without changing the standard which defines the SPIT estimation scores. In addition, methods can report the parameters used for computation (e.g., the call-rate method could have a parameter defining the length of the time-frame used as a basis for the score in milliseconds). Also these parameters would need to be agreed and standardized together with the methodID and the semantics.
In principle a single node can append a number of scores equal to the number of mechanisms that it applied to the message.
Once such semantics are defined and standardized it would be easy to start experimenting with these extensions avoiding interoperability problems.

2.3. Examples of combinations of SIP spam scores

TOC

Examples of combinations of SIP spam scores would be

- an ingress spam filter performs call rate analysis and appends a score, a filter near the callee's UA combines this knowledge with the callee's black and white lists (using a secret magic algorithm that is completely out of the scope of standardization discussion).

- an ingress spam filter performs call rate analysis and appends a score, a filter near the proxy server of the callee perform a CAPTCHA test because the call rate score was suspicious, the final decision is taken by the UA based on the time of the day taking into account the previous tests performed (the final filter is on the UA).

In these examples we assume a multi-operator and multi-vendor scenario where the spam score semantics would play a fundamental role.

---

3. Objective of the proposal

TOC

The objective of the proposal is to show a solution space to the issues raised in the SPIT mitigation discussion recently happening in the IETF.
This proposal would allow standardization of SIP spam scoring extensions that could be standardized and adopted incrementally giving

early experimental implementers the possibility to start using spam scoring extensions in an explorative fashion without running into interoperability problems.
According to the authors' opinion this proposal allows to address all the three issues raised in section Section 1 (Introduction) and it is therefore to be considered as legitimating the progress of the spam score draft [I-D.wing-spam-score] (Wing, D., Niccolini, S., Stiemerling, M., and H. Tschofenig, "Spam Score for SIP," February 2008.) inside IETF.

---

4. SPIT mitigation mechanisms overview and feasibility study

TOC

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

---

5. Security Considerations

TOC

There are issues related to integrity, confidentiality, and trust of SPIT-related information, but they are not direclty related to the definition of semantics for SIP spam score mechanisms.

---

6. IANA Considerations

TOC

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

---

7. Informative References

TOC

[I-D.tschofenig-sipping-spit-policy]	Tschofenig, H., Wing, D., Schulzrinne, H., Froment, T., and G. Dawirs, "A Document Format for Expressing Authorization Policies to tackle Spam and Unwanted Communication for Internet Telephony," draft-tschofenig-sipping-spit-policy-03 (work in progress), July 2008 (TXT).
[I-D.wing-spam-score]	Wing, D., Niccolini, S., Stiemerling, M., and H. Tschofenig, "Spam Score for SIP," draft-wing-sipping-spam-score-01 (work in progress), February 2008.
[RFC5039]	Rosenberg, J. and C. Jennings, "The Session Initiation Protocol (SIP) and Spam," RFC 5039, January 2008 (TXT).

Authors' Addresses

TOC

	Jan Seedorf
	NEC Laboratories Europe, NEC Europe Ltd.
	Kurfuersten-Anlage 36
	Heidelberg 69115
	Germany
Phone:	+49 (0) 6221 4342 221
Email:	seedorf@nw.neclab.eu
URI:	http://www.nw.neclab.eu
	Saverio Niccolini
	NEC Laboratories Europe, NEC Europe Ltd.
	Kurfuersten-Anlage 36
	Heidelberg 69115
	Germany
Phone:	+49 (0) 6221 4342 118
Email:	saverio.niccolini@nw.neclab.eu
URI:	http://www.nw.neclab.eu
	Henning Schulzrinne
	Dept. of Computer Science, Columbia University
	1214 Amsterdam Avenue
	New York, NY 10027
	US
Email:	hgs@cs.columbia.edu

Full Copyright Statement

TOC

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.