

Spam Opt-In

The following information is a pre-read for the Departmental Computing Forum next Tuesday at 1:30 PM in the Peter/Wright Room, UC. The content is meant to inform and provoke your questions and feedback; it **should not** be construed as fully-formed communications. Ideas within may be less baked than usual and may represent our attempts to share information as quickly as we know it rather than when it's ready for prime-time. Please treat it accordingly, and **PLEASE DO NOT FORWARD THIS CONTENT AS-IS TO YOUR USERS**. Instead, bring your comments, suggestions, and questions to the Forum. Thanks!

Background

Approximately 30% of the email received by the Andrew mail service is tagged as probable SPAM. Our current implementation strives to deliver all email to the end-users. As such, the Andrew email service expends a great deal of resources to process mail messages that are most likely never looked at by our users.

Every piece of Email that is destined for an Andrew user is processed by a filter which uses numerous rules and heuristics to assess the likelihood that the message is Spam. The filter assigns to each Email a score in the range 0-100; 100 corresponding to definitely Spam, 0 corresponding to definitely not Spam ("Ham"). This score is recorded in the message in the form of an additional header.

Currently, we give users the ability to use this recorded score to file messages determined to be Spam into a separate mailbox. We want to provide new and greater controls for end-users to manage their Spam.

One of the new controls is the ability to opt-out of Spam delivery (or opt-in to Spam delivery). If a message is categorized as Spam, it will be dropped before ever getting to the user's INBOX; if the user would prefer to receive their Spam, they can still do so.

The Spam opt-out/opt-in will be exposed to the user via the my.cmu.edu portal, under the My Accounts/Email category. We are planning on making "No spam delivery" the default for new accounts starting with the incoming fall class.

Technical details

To enable users to control their spam more closely, we will:

- Add a new control to the Spam Java applet to enable/disable discarding of Spam
- Add new functionality to Spam portlet backend (getspam.pl) to get/set Spam discard switch in LDAP
- Modify mail filter servers to add a X-Spam-Score header to messages.
- Modify MX servers to discard Spam based on LDAP attributes.

The perl backend for the portlet will be extended to accommodate the new features. The script will be extended to get and set the new LDAP attribute to control discarding Spam and feed it to and from the Java applet.

Sendmail on the MX machines will have a mail filter (milter) added which will fetch the new LDAP attribute for each recipient. If a recipient has enabled discarding of Spam via the discard LDAP attribute, Spam messages will not be delivered and important header information from the messages will be logged.

Once the new features are implemented and put into production, the processing of a message will look like the following:

1. Message arrives at MX server, is passed to mail filter machine and is scored for spaminess. An X-Spam-Score header is added to the message.
2. We look up the recipient in LDAP and retrieve the "SpamHandlingMethod" (SMH) LDAP attribute. The SMH is a binary flag: Y to deliver, N to drop Spam.
3. If SMH=Y and the message is scored as Spam, then the message is dropped (not delivered to a backend), we log the fact that the message was dropped and processing stops. Otherwise, we proceed to step 4.
4. Message is delivered to recipient's backend. (The recipient may have a Sieve script in place which will further process/file the message.)

Common concerns

Will Spam be logged?

- We will log the fact that we dropped a Spam, and its relevant header information, but we will not log the message itself.

What about people who want their Spam?

- Current accounts will not be changed automatically.
- People with new accounts can change this setting via the portal.

How will people know that their Spam is being discarded?

- User education is an important part of this project. Since current account holders will have to change this setting themselves, we want to make sure that they understand the implications of what they are doing.