



A Multi-Layered Approach to Effective
Outbound Spam Protection

MAILCHANNELS

Executive Summary

Although inbound spam is a serious and perennial problem, outbound spam is a more rapidly growing problem, primarily for service providers who act as the unwilling hosts of this content.

–Michael Osterman, Messaging Analyst

Outbound spam filtering is all about ensuring reliable email delivery. If your organization counts on email delivery, then you should invest in outbound spam filtering.

As the war on spam enters its eleventh year, the role of email sender reputation becomes increasingly important both for email receivers and email senders. Receivers assign a reputation score to senders that is based on their past email sending history. Senders who have previously delivered unwanted email are rate limited or blocked. Senders with a good history are permitted to send high volumes of email.

Until the last few years, only a few very large email receivers – services like Gmail and AOL – used reputation heavily to rate limit or block senders. But with more companies moving their email into the cloud, reputation is now something that everyone needs to pay a whole lot more attention to if they want to get their email delivered reliably.

This technical briefing describes the necessary techniques that sending networks must implement to defend their email sending reputation and ensure highly reliable email delivery for their users.

What Senders Must Learn From Receivers

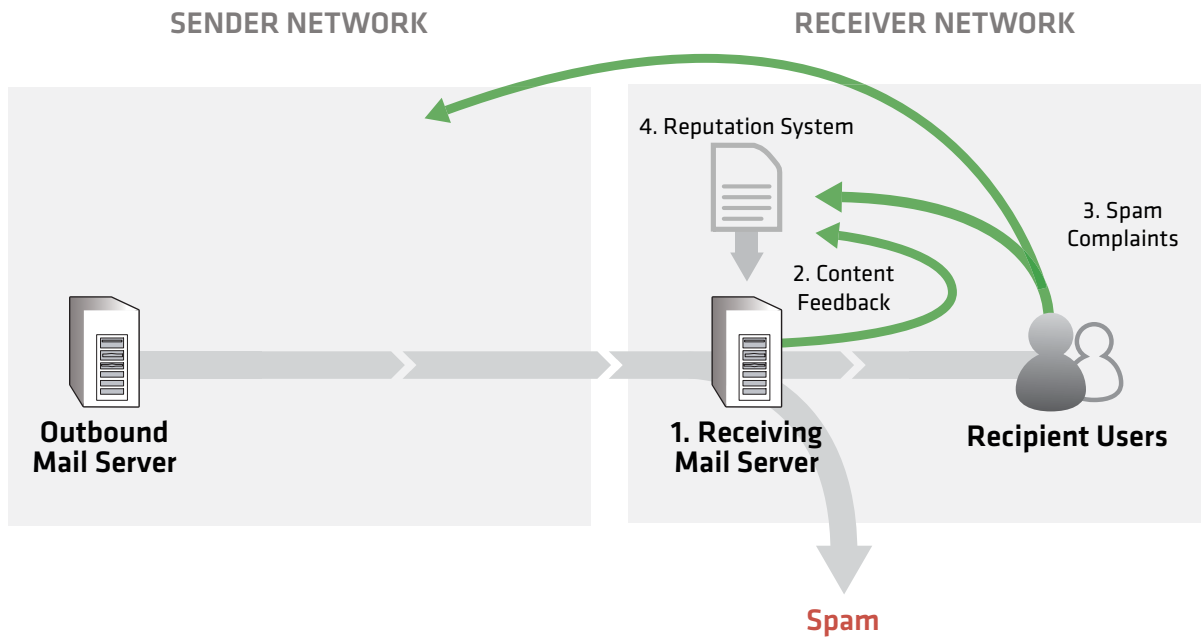


Figure 1 - Email receivers have set up sophisticated systems to detect and eliminate spam. This diagram shows the most common systems in place at large receivers, and increasingly at smaller receivers as anti-spam vendors improve their product offerings.

In designing an effective outbound spam filtering system, it's critical to understand the complex, multi-layered approach that email receivers take to eliminate spam on behalf of their users. If we understand the techniques that receivers are using to filter out spam, we can design sending systems that ensure reliable delivery for legitimate email from our users.

In Figure 1, a receiving system's first line of defense is its receiving mail server (1), which blocks and/or rate limits connections from sources identified in its sender reputation system (4). Blocks are usually applied by consulting a blacklist such as Spamhaus. Any messages that make it past the blocking stage are then scanned for spam content, and a report about the content of each message is fed back (2) into a local reputation system.

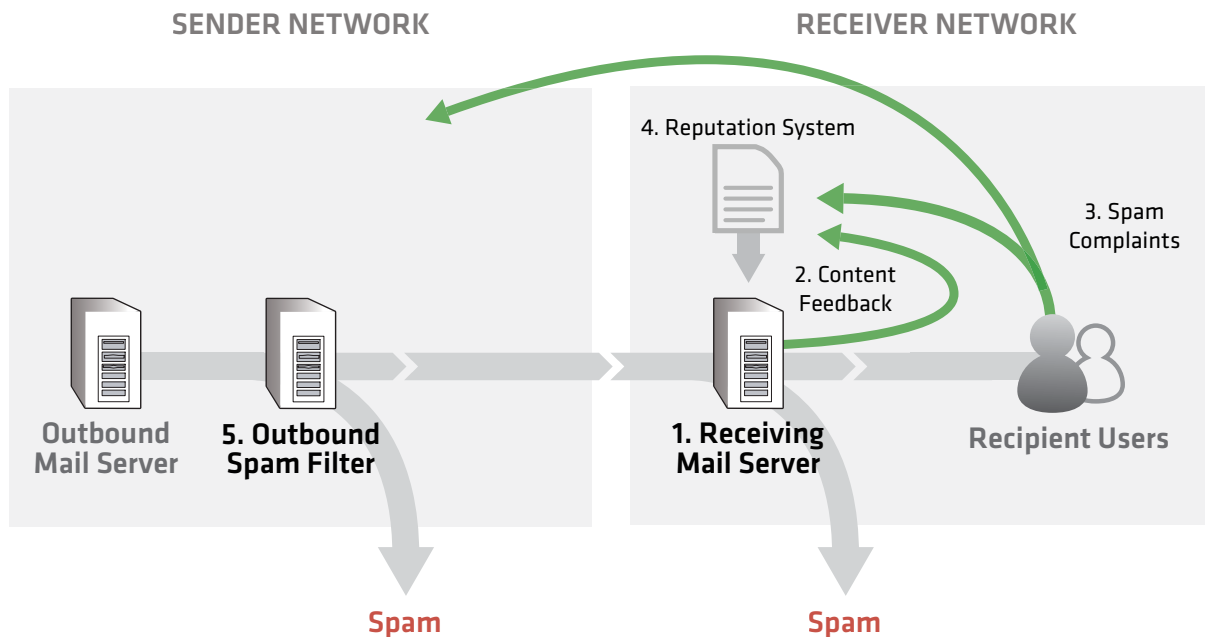
Messages that pass through the spam filter are delivered to recipient users, who then have the option of reporting unwanted content by hitting the "spam" button in their email client or web browser. These spam reports (3) are delivered both to the receiver's local reputation system, and to the sender's network through a standard abuse reporting system known as the Abuse Reporting Format (ARF).

Now that we understand how receivers filter out spam, we can begin to design an effective outbound spam filtering system to ensure reliable delivery for our users.

Inbound Filtering Layers: Fast and Slow Systems

Good inbound spam filtering systems make use of both “slow” and “fast” filtering mechanisms. Fast mechanisms identify well-known spam campaigns accurately and quickly, and tend to be responsible for removing the bulk of spam from the inbox. Slow mechanisms study feedback from users and look for longer-term patterns to mop up the rest. Large receivers like Yahoo! have made public some of their important “slow” mechanisms, which perform extensive data mining to identify dangerous or offensive content long after messages have already been delivered to users. Anti-spam vendors have their own “slow” mechanisms, analyzing feedback from their customer base to identify new spam signatures that can help to block emergent spam campaigns.

Technique #1: Real-Time Content Filtering



Accurate content filtering is the first layer in any outbound spam filtering solution. Content filters pick up on the latest spam campaigns by monitoring global spam traffic patterns and distilling the essence of each campaign as a set of rules or digital fingerprints. Ideally, an outbound spam filter should receive close to real-time updates about the latest global spam campaigns, because spammers quickly change their campaigns to evade filtering.

Content filters typically rank each message using a numerical score or category. In score-based systems, a high score indicates a high probability that a message is spam. In category-based systems, the category indicates the type of message – whether it be bulk mail, newsletter traffic, one-to-one email, or confirmed spam. In either case, if the filter thinks the content is undesirable, then it should be blocked. If the content is suspicious but not definitely undesirable, then it can be deferred for later delivery.

In the diagram above, we see that by adding an outbound spam filter (5), we can reject spam content before it reaches the receiver network, ensuring that the receiver and its users will have less to complain about.

Technique #2: Local Reputation Management

Receivers know much less about your users than you do. When someone in your network sends an email message to another system on the Internet, the receiver has to rely on the sender's IP address and the content of the message in order to make a determination as to whether to block or permit the message delivery. As the operator of the sending network, you probably know the sender's identity – for instance, on a mobile network, the mobile operator knows the account to which a particular IP address has been assigned, as well as the identity of the device that is sending the message at that time.

Because they have more information about the identity of their users, sending networks can control their users, and therefore their outbound email traffic much more effectively than receiving networks.

How Receiving Networks Use Reputation to Filter Email

Large email receivers record the historic email sending history of each IP address that has sent email to them. Each sender has a report card, in a sense, which notes the types of things each sender has done over time. When a sender wants to connect and send a message, the report card for that sender is fetched and examined, and if the report card indicates that the sender is a threat, then the sender is either blocked or rate limited to prevent abuse.

Example of a Sender Report Card

The table below shows what a sender report card might look like for a sender (192.168.32.14) that has been sending a mixture of good and bad email:

Report Card for 192.168.32.14

	10:00am	9:00am	8:00am		Friday	Thursday
Total Connections						
Non-Spam Messages	67	83	102		1204	1195
Spam Complaints	2	0	0		0	0
Spam Detections	5	3	0		3	1
Invalid Recipients	17	2	1		2	0

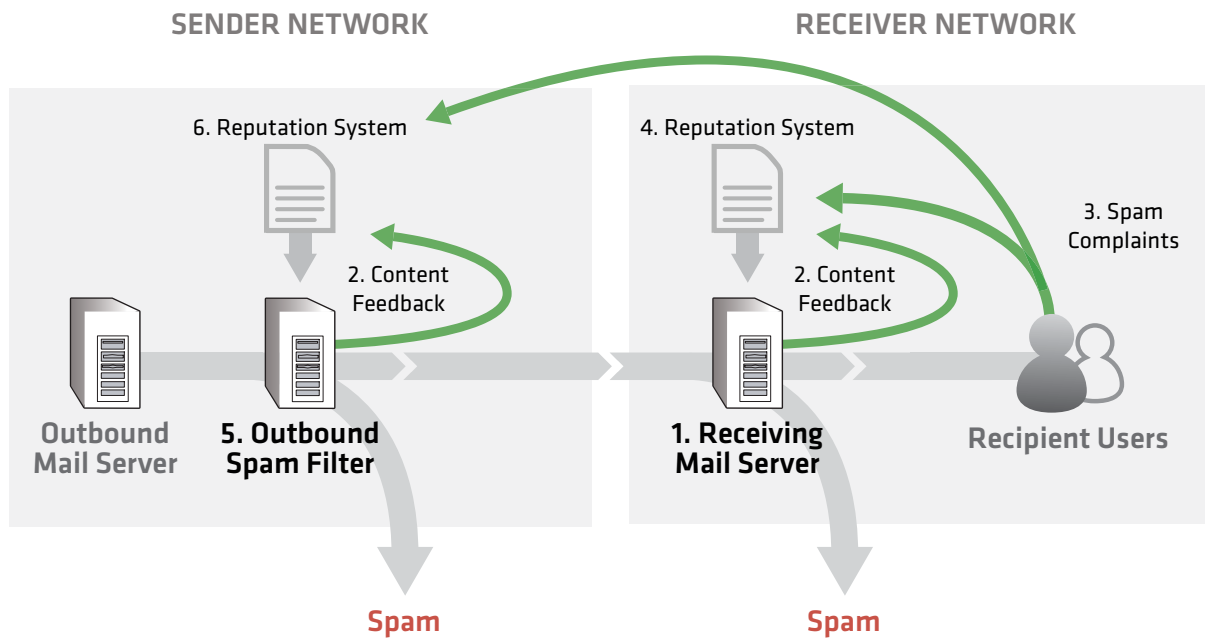
In this example, we see how the sender was behaving quite well on Thursday and Friday (no spam complaints and minimal detections); however, today around 10:00am things started to go wrong. The levels of spam and invalid recipients from this sender suddenly jumped, and users complained about two of the messages it sent. A receiving system seeing a sender report card like this might decide to start rate-limiting connections from this sender to avoid further spam.

But What is a "Sender"?

We think of the email sender as the person writing the email messages. If we go down a layer, we could view a sender as being the person's email address. Down one more layer we have the sender's domain name. At the bottom rung we have the IP address of the mail server used to send the messages on behalf of the user.

Large email receivers know nothing about the person sending the email. Domain names and email addresses can be forged. They also don't know whether a trustworthy administrator runs the system they are receiving messages from. Or whether spammers recently compromised that system. In fact, the only thing that a receiver knows for certain about your mail server is its IP address. Everything else can be forged or manipulated by spammers.

Tracking Sender Reputation to Protect IP Address Reputation



By adding a local reputation system (6 in the diagram above) to our outbound spam filtering solution, we can track the historical behavior of our users and prevent “bad” users from sending as much email as “good” users. The local reputation system maintains a “sender report card” for each user, and enables the outbound spam filter to rate limit or block email traffic from users whose history indicates they are likely to harm the sending network’s IP address reputation.

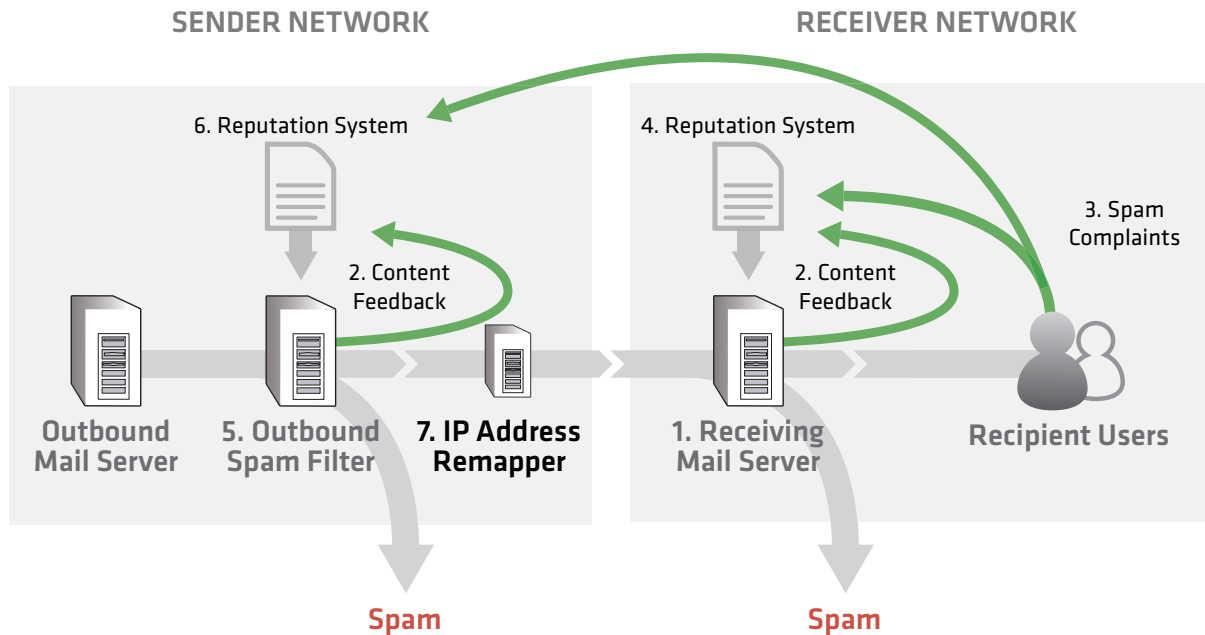
Sending networks can rely on one or more of the following types of sender identity within their local reputation system, depending on the type of sending network:

Types of Local Sender Identity

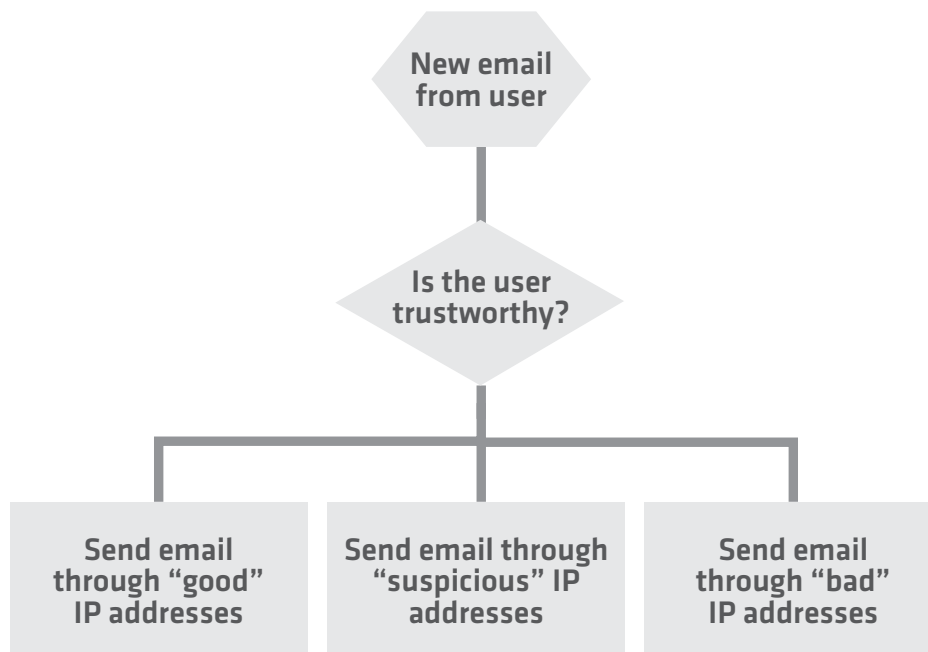
Armed with these extra piece of identification, outbound email sending systems should maintain Sender Report Cards for the following types of sender identities

Local Sender Identity	Examples	Description
Customer Name, Number or User Name	Sharon Smith Customer #453992 User "ssmith"	An identifier such as a name, telephone number, or number that uniquely identifies an individual person who sends email through the system.
Sender's Original IP Address	10.1.3.4	The IP address of the customer or user's computer from which messages are originally submitted to the outbound email gateway.
Sender's Email Address	ssmith@example.com	The email address found in the MAIL FROM SMTP command argument when email is submitted by the user's email client to the outbound email gateway.
Sender's Email Domain	example.com	In so-called multi-tenant systems such as shared hosting servers, the domain part of the MAIL FROM SMTP command argument can often be relied upon as a unique and trustworthy sender identity.
Sender's Device Identifier	+1-415-555-3299	In mobile networks and ISP networks, RADIUS can be used to map the user's IP address to their telephone number or other unique device identifier.

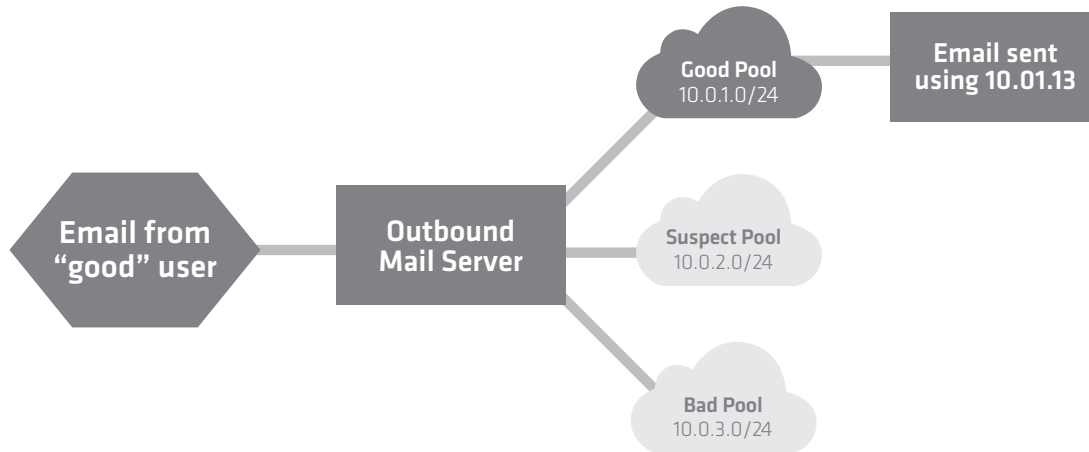
Technique #3: IP Address Management



The final tool in our outbound filtering arsenal leverages information from the local reputation system to move outbound email traffic across different IP addresses depending on the historical behavior of each sender. All the major email senders – such as Google, AOL, and Microsoft Live Mail – employ this technique as a way of protecting the good reputation of IP addresses they use to send “legitimate” email from trustworthy users. Here’s how it works:



When it comes time to send a message on behalf of a user, the user's identity is looked up in the local reputation system (6). If the user has a very positive reputation (i.e. the user has sent nothing but clean email and no spam), then the user's email traffic is sent out through one or more IP addresses that have been set aside for sending trustworthy email traffic. If the user has a somewhat less positive reputation, or if the user is simply new or has an unknown history, then their email traffic is sent out through a second IP address or set of addresses. Finally, if the user has a bad reputation, then their email is sent out through a third IP address or set of addresses designated for "bad" senders.



Most outbound filtering systems will be well protected by implementing a three-cloud configuration for IP address remapping. However, additional remapping clouds can be added to send out traffic with further granularity. It's important to mark each cloud in the DNS and by using WHOIS records so that receivers know "this is our good cloud" and "this is our bad cloud".

Conclusions

The most effective outbound spam filtering solution involves a combination of several techniques, including accurate content filtering, reputation-based policy management, and active IP address management. By leveraging the additional sender identity information available on the sending side, sending networks can police their own customers and ensure that the reputation of their IP address space remains positive for reliable email delivery.