Email Authentication Summit- Comments
Matter Number P044411

Federal Trade Commission
Office of the Secretary, Room 159-H, Annex V
600 Pennsylvania Avenue NW.
Washington, DC 20580
authenticationsummit@ftc.gov

Key Points:

- An authenticated HELO domain is the only sufficiently strong name identifier within a mail stream to safely identify the message source and allow a reputation assessment of those accountable.
- The HELO domain is analogous to the letter postmark, but is currently unreliable due to a protocol defect resulting from limitations in DNS information. Without changes to RFC 2821, the HELO domain must be allowed to fail authentication.
- Adding a DNS Service record to assure HELO domain authentication and additionally validating authorization can be implemented without any negative impact upon existing applications.
- Making the authenticated HELO domain name visible to the user would be an effective and safe deterrent against phishing and spoofing, as this identity would be the most difficult to spoof as a means to provide false assurances.
- The authenticated HELO domain identifies the server holding the SMTP log needed for criminal enforcement.
- Authenticated HELO domains enable a simple and safe means to associate authorized mail transfer agents with the mailbox domain through the publishing of a simple name list. The alternative text scripts, as used with SPF or Sender-ID, to obtain addresses for a large array of hosts is inherently perilous, and must be discouraged.

Reduction in Spam

A strategy predominately used to reduce spam and preserve network resources is called the Real-time Black-hole List (RBL). This list is based upon the IP addresses of the connecting client where, when a record is returned after the address is queried, the client is refused with a reply that often names the list responsible for the rejection. The address is used as the identifier because there are currently no validated names associated with a mail stream. The lack of a validated name makes it difficult to know when an address has been reassigned and the reverse DNS name directory is too poorly maintained to be of use for this purpose as well.

An address is currently the only strong identifier that is authenticated by way of interaction within the transport protocol within a mail stream. With the potential number of addresses being fewer than the potential number of names, tracking addresses that are sending abusive mail is effective at making access difficult for abusers. The success of the RBL, as a tracking mechanism, has led to tactics such as Trojan programs as a means to commandeer new addresses. Introduction of IPv6 addressing also changes the number of addresses compared to names. To move away from the address based methods for tracking abuse, a strongly authenticated name is required.

The strength of this authenticated name must be as strong as that of the IP address due to litigation risks. The strength of the abuser's identity is paramount for defending a reputation assertion. The selected name must identify the domain administrator which controls the specific mail transfer agent and is accountable for the requisite security. For the purpose of abating spam, there is only a single domain name within the mail stream which meets this requirement. This would be the HELO domain name.

All other domain names within the mail stream are based upon an unverifiable assumption of the mail stream integrity, regardless of any association within domain name server records. Any such association which serves to authorize the sending of mail on behalf of a specific mailbox domain, does not imply the mailbox domain administrator accepts accountability for the performance of this authorized mail transfer agent when controlled by a different administrator. The administrator of the mailbox domain not in control of the mail transfer agent can not be expected to make assurances or take corrective measures. The consumer, as well as the reputation service, would be threatened by an overly broad application of accountability where litigation becomes the only recourse for an unfair reputation assessment. Accountability must be constrained to the domain administrator, identified by the HELO domain, in control of the mail transfer agent. This administrator is expected to monitor SMTP error logs, track abuse@ complaints, maintain security, and disable problematic accounts as a means to control access.

The SMTP protocol has all the necesary components to assess this accountability. SMTP presents a client name in the HELO exchange at the beginning of an SMTP session that is recorded within the message RECEIVED headers. Unfortunately, failure to authenticate the HELO domain name must not be used to refuse mail from the client, as required by the SMTP standard RFC 2821. Any name may be presented in the HELO exchange, where authentication, although vital, is often missing. The steps to rehabilitate the HELO name to ensure it will authenticate also enables an assertion of authorization to protect machines that may be compromised. When this domain name is both authenticated and the authorization for sending mail is validated, then this domain can be safely accredited "by name" for traffic emitted.

Having authenticated and verified the authorization of the SMTP client by name, should there then be criminal activity detected, the server logs can be readily uncovered. This would be a substantial improvement over just having an IP address, as now an authenticated domain has made expressed authorization for the mail sent. The registration of the domain also offers the name of the applicant and a possible money trail from the purchase of the domain.

Prior to the break up of the IETF MARID working group, there was a suite of standards proposed under the title of Client SMTP Validation, CSV by Dave Crocker, Douglas Otis, and John Leslie. A new working group will carry forward these standards, as the chairs of the now defunct MARID working group held back consideration of these standards. The salient aspect of this suite of proposed standards was aimed directly at repairing the problem which prevents the HELO domain from being authenticated. This suite further requires the domain to expressly authorize the client for sending mail.

Modification of standards

There is one aspect of SMTP which needs to be addressed. The authentication of the HELO domain name. This can be done easily by introducing a DNS record specifically tasked to ensure a complete list of addresses will be available when queried. This can be done within a single binary query of a DNS Service record. This solution will be addressed by a new work group formed within the IETF to carry forward the Client SMTP Validation (CSV) specification suite that was started in the now defunct MARID work group. See:

http://www.csvmail.org/
http://www.ietf.org/internet-drafts/draft-ietf-marid-csv-intro-01.txt
http://www.ietf.org/internet-drafts/draft-ietf-marid-csv-csa-01.txt

If there is a desire to constrain the mail transfer agents authorized to send mail on behalf of a mailbox domain, there is also an IETF draft named “Mail Policy Record (MPR)” that illustrates how a simple list of names can define the authorized sources of mail. The use of this name list does not invite exploitation, as it will not normally act as a gate-keeper. This list can alert the user to a message carried outside the nominal mail channel. Financial institutions likely the subject of phishing would be well advised to use consistent names for their HELO domains and to publish their nominal sources by way of this name list. See:

http://www.ietf.org/internet-drafts/draft-otis-marid-mpr-00.txt

Should there be a need for an immediate solution that rejects obviously spoofed mail, then the use of the Mail Policy Record draft name list used in combination with Client SMTP Validation, offers a safe alternative to either Sender-ID or SPF. Use of the name list to indicate nominal sources for a specific mailbox domain avoids the perilous use of domain name server text scripts that attempt to compile all the addresses for a vast array of hosts. The text scripts used to implement both the SPF and Sender-ID schemes promoted by POBOX.COM and Microsoft represent a hazard from several perspectives. The time needed to process these SPF and Sender-ID address lists can be exceeding long. If these drafts adhered to timeouts for domain name server transactions, the time required may exceed hours. By not adhering to these timeouts, the User Datagram Protocol (UDP) exponential back-off is violated and thus does not provide the requisite congestion avoidance. Even with the early timeout of 200 seconds as specified, malicious scripts could easily effect a denial of server attack aimed at disabling the checks.

Because both of these script schemes require a specialized parsing program, the source port used for domain name server queries may be constrained by the application running the parser. The hundreds of potential records and different domain name servers referenced allows a "birthday" attack staged using just a single DSL network to poison the records held in the domain name server's cache. This risk is acute when the domain name server does not aggregate pending queries as with Bind 8, a popular version of the domain name server.

DNS cache security overview by Joe Stewart:
http://www.securityfocus.com/guest/17905

Compatible with Legacy

CSV offers true consumer protections when the authenticated HELO domain name is visible and when those responsible for security receive the reputation assessment. Until the consumer is using a mail client able to present the authenticated HELO domain name, there would be no expectations of increased assurances. Unlike SPF or Sender-ID, the use of CSV will not interfere with the normal use of mail and allows rapid deployment, as the impact to legacy systems would be negligible. Unlike SPF or Sender-ID, the user is not expected to forgo use of their favorite mailbox address. Unlike SPF or Sender-ID, it is the provider that receives the reputation allowing their customers the freedom to find other providers in the event the provider is blocked for allowing abusive mail. Unlike SPF or Sender-ID, older mail clients can not be used to provide false assurances.

Handling of unauthenticated messages

Initially, the phase-in of CSV moves from the reliance upon the client's address to the authenticated and authorized HELO domain. Unauthenticated messages that fail to provide the newer HELO domain authentication records may receive a "slow path" approach as a means to limit the extent of potential damages. By making only the authenticated HELO domain names visible to the user, this too would be an incentive to deploy CSV as a means of providing this increased assurance. This added assurance may also be used to reduce the chance of being "filtered."

Erroneous results

Unlike Sender-ID or SPF, there is little opportunity for erroneous results with CSV. With either Sender-ID or SPF, a shared mail transport agent opens up the possibility of spoofing the mailbox domain. Both of these schemes depend upon the unverifiable assumption that the sending mail transfer agent has performed the needed checks and that the lists are "closed." With the information published for Sender-ID or SPF, a malicious attack may only require an assertion of the mailbox domain. CSV records are unique for each mail transport agent and make no assumptions of the mail stream integrity. Unlike Sender-ID or SPF, there is little risk of the domain name provided by the HELO transaction being spoofed. This provides consumers an identity which can be safely relied upon whereas SPF and Sender-ID do not. Institutions would be well advised to be consistent with their naming conventions within the HELO transaction to ensure reliable recognition of the domain name.

Compatible Enhancements

The use of the Mail Policy Record proposal, which allows the listing of the nominal mail channel for a mailbox domain, can be added without incurring any significant overhead. The use of the name list can be used to alert users of a possible spoofing without exposing the domain name system to the risk of a denial of service attack or being poisoned as is possible with Sender-ID and SPF. The alternative Mail Policy Record proposal permits a simple name list be obtained within a single domain name server query to ensure proper source recognition.

Proprietary Nature

Unlike Sender-ID, there is no proprietary algorithm for sorting the field being checked with CSV. By always using the same HELO parameter there is less risk of differences in the application of a complex algorithm that could allow a sneak path for spoofing.

Proprietary Services

CSV does not depend upon any proprietary services. Right Hand Side Black-hole Lists (RHSBL) have been in general use and are based upon structures established by MAPS, which provided the original RBL services.

Effect on forwarding

CSV will not interfere with forwarding. SPF and Sender-ID do interfere.

Effect on mobile use, roving users, and Mailing lists

CSV will not interfere with the operation of any of these mail applications or uses. SPF will seriously impact mobile and roving users. Both Sender-ID and SPF may potentially impact the operation of mailing lists and may require changes to the mail submission agents.

Effect on Outsourced mail

CSV allows a provider to request a CSV record be made to allow a recognized HELO domain to be used. Such a strategy would only be likely with financial institutions where phishing has become a problem. Otherwise, the outsourced mail source can accept accountability for the mail and have their service visible to the consumers.

Effect on multiple domain use or web-mail

CSV does not affect the use of multiple domains nor impact those using mail services that do not offer a mailbox. Although CSV allows the use of a nominal mail source list, this list should be viewed as informational to ensure mail still functions and messages are not lost inadvertently.

Scalability

CSV does not reduce the scalability of SMTP. Once widely deployed, CSV will actually reduce the overhead often expended in attempts to validate the sending client.

Identity costs

The overhead and maintenance associated with CSV records would be extremely small. These records are name based and do not require a double entry system of addresses as is the case with SPF and Sender-ID. CSV simply adds a stable record to DNS and does not require any digital certificates or other outside services. CSV does enable the use of reputation services, but these services would be optional. Those that do not use these reputation services still benefit indirectly from those that do use these services.

Impact on Non-participants

Those that fail to participate will not offer their message recipients assurances of the source for these messages. These messages will likely also run a greater risk of being filtered as being spam. Some providers may wish to utilize a slow path algorithm to limit the number of messages from participants that fail to authenticate the source of the mail.

Adoption time-frame

Although nearly all types of domain name servers are able to support Service records, providers offering web-based mail services may need to create the necessary user interfaces to add the CSV record. As a CSV record has so few requirements, the record could be included transparently without any inputs from those setting up the service.

Microsoft has supported these record types, established in 1996, since their departure from WINS in Windows 2000. Those older domain name servers that predate this time frame, but are expected to support SMTP, should be upgraded out of security concerns anyway.

As CSV does not impact the operation of mail, the adoption time-frame can be exceedingly rapid with minimal risk. Wide adoption of CSV will actually lower the overhead associated with SMTP.

Burden of mechanism

Unlike SPF or Sender-ID that wish to impose a tremendous burden upon the SMTP mail exchange, CSV potentially represents a reduction in the overall overhead. Once widely deployed, reliance upon the CSV record allows the reverse DNS, address record, and MX record lookups to be skipped which are typical DNS lookups often done in vain while attempting to validate the client. The single record and single domain name server minimizes the risk involved when performing lookups from potentially hostile sources.

Ensure anonymous use

CSV does not change the paradigm that the sending domain controls account access and sets permissions for the acceptable range of mailbox addresses. Both SPF and Sender-ID impose a risk on those that leave their mailbox domain sources open-ended. These types of open-ended records actually invite address spoofing as a means to side-step an exclusion based on a lack of recorded association. Although CSV protects those domains using this freedom responsibly, SPF and Sender-ID punishes those wishing to leave their records open, through these invited exploits.

Anti-trust concerns

CSV should be open to the world to implement without any corporate approvals.

Exposure to Attack

CSV does not increase the susceptibility for attack by not increasing the number of records or the number of servers sequentially queried. SPF and Sender-ID however open up many avenues for attack. A simple domain name server waiting 4 seconds for each response creates a type of denial of service attack, as these SPF or Sender-ID scripts often need to resolve hundreds of records. The quantity of records needed for either SPF or Sender-ID also enables relatively lazy cache poisoning attacks. As both Sender-ID and SPF expect to allow script extensions without changes to the script revision, the complexity needed to step-over unknown statements adds to the risk. In the case of Sender-ID, this process is also based upon the complex selection of the message header. A difference in processing of complex algorithms for SPF or Sender-ID may also allow an undetected type of spoofing.

Prevent Phishing

CSV implements authentication for the strongest name within the mail stream, the HELO domain. There is no other name better suited to offer consumer assurance of the message's origin. SPF only checks the invisible RFC2821 MAIL FROM parameter. MAIL FROM need not be related in any manner to the RFC2822 FROM header that the consumer sees. Sender-ID suffers the same problem as the unseen RFC2822 RESENT-FROM header is also not related to the RFC2822 FROM either. The HELO domain represents the mail provider used by the author of the message and thus is more closely related to the author than any other header within the message. CSV also allows a means to safely obtain a name list referenced by the RFC2822 FROM mailbox domain as a means to alert the recipient when the message appears to be outside the nominal mail channel. See the Mail Policy Record (MPR) draft for details.

An authenticated HELO domain offers the equivalent of the postmark. If you received a letter from Uncle George, but from an unexpected postal location, you could write Uncle George and ask about his trip. If he responds saying "What trip?" then the postmark served a valid service. By the same token, if your bank statements always arrive carrying the postmark "mx01.big-bank.com", but this time you notice that the postmark is "mx01.big-bank-one.com", again the recipient would be advised to call and ask about this odd postmark before considering the message valid. Given the needed resources, people are capable of recognizing how a postmark is related to the author of the message. By allowing people to see information that is not easily spoofed, the ability to perpetrate a hoax will drop significantly. Neither Sender-ID nor SPF offer any notable protections from being fooled as to the source of the message. With Sender-ID, the header could be one of half a dozen headers selected that have nothing to do with the service provider nor the author.

The Mail Policy Record proposal provides a means to request messages for particular mailbox domain to be rejected when not from approved mail transfer agents. This would not be a suitable mode for most mail, but, for those experiencing a significant amount of phishing attacks, they may opt for this approach. When declaring exclusive sources for mail, such institutions would need to alert their customers not to provide mailbox addresses that are being forwarded to different mail accounts.

Complexity of implementation

CSV is by far the easiest to implement. It only requires the change from a lookup of an Address record, to that of a Service record. The mail user agent should recognize the domain of the mail delivery agent and display the HELO domain obtained outside this domain when it confirms the validity of the referenced record.

National concerns

SENDER-ID AND SPF MUST NOT BE DEPLOYED! Both of these mechanisms put both mail and the domain name server system at risk. Theses script based schemes attempting to provide addresses for a vast array of hosts and overwhelm DNS at an unsuitable scale which may lead to the following:

- Inordinately high levels of UDP traffic.
- UDP exponential timeout back-off circumvention with higher rates of packet loss.
- DNS cache poisoning.
- Denial of service attacks upon the mail system as many are connection limited.
- Denial of service attacks upon DNS due to loads created by "invited" exploits.
- Increased equipment costs for mail processing due to greatly increased overhead.
- Consumers hurt by unfair reputation assertions based upon weak identities.
- Consumers hurt by unfair reputation assertions caused by lax providers.
- Consumers hurt with mailbox addresses not being independent of service providers.

Cryptographic requirements

CSV does not use any cryptographic mechanisms unless DNSSEC is used. DNSSEC however is an independent issue.

How can the summit support the effort

- Recommend deployment of the Client SMTP Validation (CSV) standard.
- Recommend against use of either the SPF or Sender-ID mechanism out of security, network integrity, and consumer protection concerns.
- Recommend the display of authenticated HELO domain names in mail user agents.
- Recommend the consistent use of HELO domains by financial institutions.

Future measures to combat spam

Create a salted hash based "Do Not Email List" with plenty of "planted" addresses from either web pages, newsgroups or other potential sources. Provide a service to allow bulk emailers a means to vet their lists by submitting them to be processed. If the submitted list contains one of the "planted" addresses or contains invalid domains, then the results of the comparison should be squelched where the sender is asked for an interview to discern why they are using "planted" or "bad" addresses. The "Do Not Email List" should never be shared nor stored in the original form to further protect those submitting their mailbox address from their address being used as a result of their submission. Even releasing the hash values would unfortunately enable a dictionary type of scanning to discern the original email address.

Create a clearing house for domain name applicants as a means to prevent an increased flood of names intended to overwhelm name based reputation services. Initially the only means a name based reputation system will endure a name flood would be through the categorization of known good, a rapidly aged known bad, and those new to the system.

Douglas Otis
Senior Engineer
Research and Development
Mail Abuse Prevention System (MAPS)
1737 North First Street, Suite 680
San Jose, CA 95112-4234

(408) 453-6277 x170
(408) 453-6222 fax
dotis@mail-abuse.org
http://www.mail-abuse.com/