Trust in Email Begins with Authentication

Issued by the
Messaging Anti-Abuse Working Group (MAAWG)
June 2008

Edited by Dave Crocker
Brandenburg InternetWorking

Abstract

The Internet's growth allows us to interact with people all over the world. Unfortunately, some of those people do not make good neighbors. Along with the effort to detect and filter the problematic traffic they generate, there is a complementary effort to identify trustworthy participants. In security technology parlance, the first seeks to identify *Bad Actors* whereas the second creates ways of distinguishing *Good Actors*. At its simplest, identifying Good Actors can be divided into two activities: A safe means of identifying a participant–such as an author or an operator of an email service–and then a useful means of assessing their trustworthiness. The first activity is called *authentication* and the second is usually called *reputation assessment*. This white paper considers the first step: authenticating the identity that asserts responsibility for an email. In it, recent developments in standardized authentication mechanisms are reviewed that have been tailored for use in email anti-abuse efforts.

This white paper provides background on authentication as a foundation for understanding current efforts to protect Internet mail. It then looks at the most popular mechanisms currently in use. The paper is intended for a general readership that has basic familiarity with Internet mail service. While this single document is unlikely to be the final word on the topic, MAAWG has striven to capture the current best practices and leading theories regarding email authentication. As a complement to enabling identification of Good Actors, authentication is expected to aid efforts in protecting business' brands from forgery and phishing attacks. The Executive Summary provides a one-page overview that can be used independently.

Contents

Executive Summary

The disruptive impact of spam and other email abuse has generated two lines of response by the email services industry. One focuses on detecting and filtering problem messages. A complementary, but quite different, response seeks a basis for trusting a message rather than for mistrusting it. Is someone trustworthy responsible for the message? This approach has three steps: 1) assign a person or organization's name to the message – *identification*, 2) then verify that the use of this identifier is authorized and correct – *authentication*, and 3) finally determine the trustworthiness of the identity – *reputation assessment*. This paper discusses current industry efforts to satisfy the first two requirements.

Internet mail is extremely flexible in the types of identities that can be referenced. Most recipients only know about the `From:` header field, containing a displayable name of the message author and their email address. However email also can separately name the agent that posted the message for sending, the agent that receives return handling notices, or agents that handle the message at different stages of transmission. Each of these is often referred to as a *sender* of the message, so the term *sender* can be ambiguous.

When authentication mechanisms are applied, both the originating and receiving systems are able to correctly and reliably validate who is accountable for the message. This is generally described as "knowing who the message came from." When an authentication mechanism becomes widely popular, it opens the door to a variety of assessment products and services that can rely on it.

Three candidates have emerged for authenticating who is accountable for a message while it is in transit. They cover two technical paradigms, use different identifiers, and have different limitations and flexibilities: *Sender Policy Framework* (*SPF)* and *Sender Identification Framework* (*Sender ID)* use a path registration approach, whereas *DomainKeys Identified Mail* (*DKIM)* uses digital, cryptographic signing. A simple way to distinguish these two paradigms is that the one builds authentication into the email-handling infrastructure along the path that a message travels, whereas the other wraps authentication information into the message itself and is independent of the infrastructure. Other authentication technologies are used for email, but they do not deal with email transit accountability.

SPF uses the underlying network address (*IP Address*) of the email-sending neighbor that is closest to the validating server and the machine name (*Domain Name*) in the message's return address or the Domain Name in the protocol transfer greeting between email handling hosts. It queries the Domain Name System (DNS) to perform a mapping between the name and the address. Hence, the IP Address of any email relay that might be a neighbor, for SPF evaluation, must be known beforehand and pre-registered in the DNS by the sender. SPF validation does not function as desired when a message is sent through a forwarding service or independent posting service. An enhancement, called *Sender Rewriting Scheme (SRS),* recruits intermediaries to modify the problem address and make it acceptable for SPF validation. SPF also permits the DNS mapping entry to contain assertions about mail that comes under SPF registration.

Sender ID is the same mechanism as SPF, except that it chooses from a different array of Domain Names, specifically the *Purported Responsible Address (PRA)* domain in a message's `From: or Sender:` header fields, or its return address' Domain.

DKIM uses digital, cryptographic signatures, attaching information to a new header field in the message. A DKIM signature can withstand minor message modifications without becoming invalid, including some that are made by forwarding services and mailing list software. Any Domain Name can be used for signing the message. As with SPF and Sender-ID, the queried entry in the DNS self-validates the name's use. Associating it with an existing identifier, such as the `From: or Sender:` header fields, is a separate step. An enhancement to DKIM will permit DNS publication of a Domain's *Signing Practices (SP)* that is intended to aid assessment of the possible legitimacy of unsigned messages.

Introduction

There are two lines of industry response to protecting users against abusive email. One focuses on deciding that a message is from a Bad Actor; it is based on detecting and filtering out problem messages. It analyzes email sources, traffic patterns and content. Mail cannot cause problems for recipients if is not delivered to them. Although essential as a first line of defense, this approach is at best approximate. Filtering junk email can be prone to error, with *false positives* – legitimate email classified as junk – and *false negatives* – junk email classified as legitimate. The distressing yet inevitable result is that this approach produces an ever-escalating arms race of counter-techniques, locking the abuse and anti-abuse communities in a constant struggle. For example, as the anti-abuse community has gotten better at analyzing textual content, the advent of *image spam* has become a creative vector of attack to defeat such analysis. Some service providers have become extremely proficient in protecting users, but the cost is very high and the protection is very fragile. Few providers can ensure this level of protection. With email abuse estimated to be as high as 90 percent of Internet messaging traffic, it has become critical to find ways of restoring user confidence in email.

The second approach, which complements problem email detection, is to find a basis for trusting the message rather than a basis for mistrusting it. A message determined to be from a *Good Actor* does not need to be subject to the stringent analysis that would be applied to mail from an unknown source. The usual method of accomplishing this is to associate a confirmable identity with the message and to obtain an assessment about that identity. In other words, provide the recipient with a means of deciding that someone trustworthy is responsible for the message. For example, provide information answering the questions: who is the author and are they known to write legitimate messages?

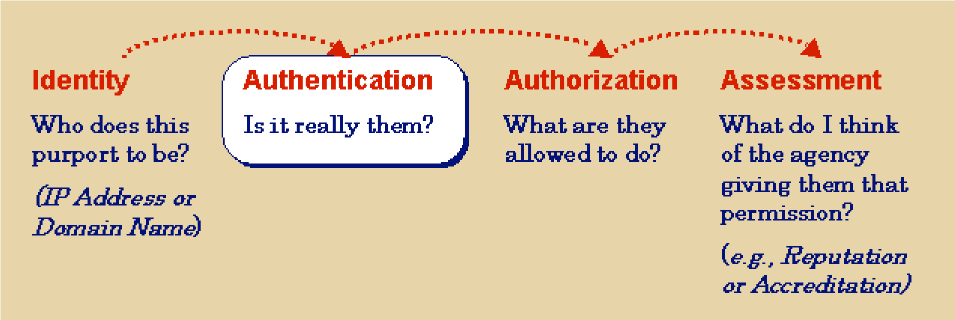

Identity
Who does this purport to be?
(IP Address or Domain Name)
Authentication
Is it really them?
Authorization
What are they allowed to do?
Assessment
What do I think of the agency giving them that permission?
(e.g., Reputation or Accreditation)

Figure 1: The Assessment Framework

As shown in Figure 1, if we are to trust the claimed identity of a message sender, then we first need a mechanism that validates the identity's use. This is called *authentication*. Only when we know that the identity is valid can we assess its trustworthiness. Without authentication, we could know a person's or an organization's reputation information, but could not be certain that they are indeed responsible for this message.

Authentication-based assessment follows these three steps:

1. Assign an *identifier* to a message, which refers to an *identity* – the name of a person or organization.
2. Provide a means of validating the use of that identifier – an *authentic* identity that is *authorized*, for this use.
3. Assess the *reputation* or trustworthiness of that identity.

This paper focuses on the first two steps, which produces an authenticated reference to an identity. The technologies discussed here describe how to assign an identifier to the message. The second step checks whether this use of the identifer is permitted, and the third step assesses the trustworthiness (*reputation*) of the responsible person or organization (*identity)* for that use.

Outside of Internet technology, an analogy of the distinction between the processes of identifier verification described in the first two steps above and assessing reputation in the third step can be illustrated by the role a driver's license plays in commercial transactions. A license contains a name that refers to a person; the name is an identifier. A license is generally accepted as validating a person's identity, but it does not indicate the person is approved for a loan. Rather, the person's credit score is the measure of his reputation that is used to determine credit-worthiness.

Moreover, a person or organization can have multiple reputations, each within its own context. Reputation analysis begins by determining the authorized scope for using an identifier and then assesses the associated identity within the current context. For example, in the case of the loan, the context analysis might start with whether the person is applying for a personal loan or a loan for their company. In the case of email reputation assessment, the starting point for the context analysis might be whether the individual is sending a personal message or one purporting to represent their company.

Underlying Technologies

Email

The global architecture for Internet mail is shown in Figure 2. There is a simple split between the user world, in the form of *Mail User Agents (MUA)*, and the transmission world, in the form of the *Mail Handling Service (MHS)* composed of *Mail Transfer Agents (MTA)*. An MTA that sends an organization's mail directly into the public Internet, or receives mail from it, is called a *Boundary,* or *Border,* MTA. An MTA that sends messages handles *Outbound* mail and one that receives messages handles *Inbound* mail.

The MHS is responsible for accepting a message from one user, the *author*, and then delivering it to one or more other users, the *recipients*. This creates a user experience of apparently-direct MUA-to-MUA exchange, without a user having to be cognizant of the intervening steps performed by the MHS infrastructure. The first component of the MHS is called the *Mail Submission Agent (MSA)* and the last is called the *Mail Delivery Agent (MDA)*.

Internet mail is often referred to by one of its standards, *SMTP* (Simple Mail Transfer Protocol), but is really a collection of specifications, with the core being RFC 2821 (SMTP) for transfer, RFC 2822 for the message header and RFC 2045 (MIME) for the message body and attachments. A recipient retrieves a delivered message using *POP* (Post Office Protocol) or *IMAP* (Internet Message Access Protocol) or proprietary protocols such as Microsoft's MAPI.

Internet mail standards specify the meaning of the identities such as author and sender that are used in sending a message but does not mandate or enforce particular choices for them. Although some software and some originating organizations choose to constrain the use of identifiers, the reality is that the underlying Internet standards permit an author to

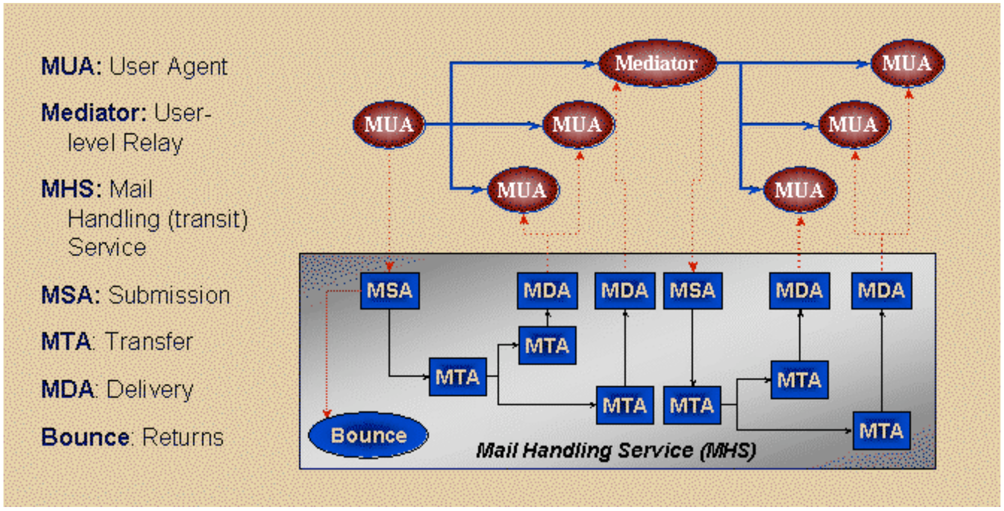


Figure 2: Internet mail Service Architecture

claim to be anyone. While this is the same as for telephone and postal communications, we have developed adequate techniques such as … and … for dealing with abuse issues in these older communications mechanisms. Because these techniques will not work for the Internet, we need new ones that are designed for our new communications mechanisms.

Many Roles in Handling a Message

The concept of an (online) identity has a rich and confusing history. One source of confusion is the difference between a thing itself, versus the name of the thing. In society, a person or an organization can be referred to by one or more names, such as a legal name or a nickname. Both refer to the same person or organization, but with different labels. In technical parlance, the term *identity* refers to the person or organization itself. The term *identifier* refers to a label that is used for that identity.

In the case of a driver's license, there is one *identifier* – the name on the license – and one *identity* – the person to whom the identifier refers. A bank might use a driver's license to confirm identity, but it needs additional information, such as a credit score, to determine a person's financial reputation and credit worthiness. The better the reputation, the more financial and other privileges the driver enjoys, such as lower interest rates. Similarly, the better an email sender's reputation, the fewer restrictions that might be placed on their email. Restrictions that are commonly placed on email include limitations on attachments or attachment types, the number of messages that are permitted to be sent into an organization in a given amount of time, "spamminess" of a message before it gets blocked, and so on.

There can be a number of different identifiers associated with each message, as listed in Figure 3. Most recipients recognize the `From:` header field, containing a displayable name of the message author, along with their email address to which replies are usually sent. However there are other identities, primarily to assist the system in distinguishing different actors, such as who posted the message for sending, who will receive return receipt (*Return*) messages, and who handles each stage of message transmission. Each of these identities is often referred to as a *sender* of the message. So the term *sender* can be confusing.

Reference	Meaning of Identification
Peer MTA Host IP Address	Neighbor SMTP Client Host
SMTP `EHLO` Command	Neighbor SMTP Client Organization
Peer Network IP Address	Neighbor SMTP Client Provider
SMTP `MAIL FROM` Command	Notification Return Address
RFC 2822 `From:` Header Field	Content Author
RFC 2822 `Sender:` Header Field	Message Posting Agent
RFC 2822 `Received:` Header Field	Transit Handling Organizations

Figure 3: Internet Mail Identities

Among these identities, there is a distinction between responsibility for content, that is the *author*, versus responsibility for message handling, that is *posting, transmission* and *returns*. Each can be useful to consider when deciding whether to accept an authenticated message, and different authentication techniques use different identities. But unfortunately, answering the key question of who is "responsible" for a message, and then validating their identity, is not always easy.

Domain Name System (DNS)

The DNS provides a mapping service, from *domain names* to associated information, such as the *IP Address* of the Internet hosts that are known under the registered name. The name itself specifies a registration hierarchy, such as `marketing.example.com`, with the right-most field being the top of the hierarchy. In addition, the DNS query service defines a hierarchy of hosts to contact, for resolving the domain name to its associated information. The query hierarchy does not have to match the boundaries of the naming fields. That is, a sequence of fields can be resolved by a single server, such as `marketing.example` being fully resolved by a single host that was referenced by the `.com` DNS server.

The term *mapping* distinguishes the DNS from more general "search" services. It takes the exact name and produces either an exact match or a failure, whereas a search service explores for approximate matches. An aggregation of domain names that are resolved by a single host is called a *zone*. This construct is administrative rather than being visible to users of domain names. So a host might resolve a single level of the name's hierarchy (*sub-tree*) or multiple levels.

The information that is associated with a domain name is listed under a set of *Resource Records* (`RR`). Each type of `RR` has its own format. The `TXT RR` has a general text format that is further defined by various applications that use it for recording different information. To distinguish which application defines the `TXT` contents, the record either must contain a defining string, or the record must be stored under a special DNS naming subtree, typically define by a name string that begins with an underscore.

The Promise of Authentication

Email authentication specifically addresses the problems caused by Internet mail's flexibility in choosing identities to use. With authentication, both originating and receiving email systems gain a mechanism for validating who is responsible for the message and is generally described as "knowing who the message came from." This ability can ensure that legitimate mail reaches its intended recipients. It also creates an opportunity for presenting recipients with a visible indication that a message can be trusted. Some commercial experiments are exploring user acceptance and the benefits of such an indication.

In this white paper, the term *authentication* describes a technology for determining that an identifier is being employed by, and for, the organization (the *identity*) that it belongs to. Is an IP Address being used by the organization it is assigned to? Is a domain name being used by the organization that registered it? Address and domain name registrations derive from the global Internet administrative authority, ICANN.

The owner of the identifier being used is declaring that they are accountable for the message. This means that their reputation is at stake. Receivers who successfully validate the identifier can use information about its owner as part of a program to limit spam, spoofing, phishing or other undesirable behavior.

Whether this information will assist in improving the deliverability of a message or in bypassing filters is entirely at the discretion of the validating receivers. When a message is authenticated, a receiver uses their knowledge about the owner of the identifier – that is, the identity – to determine the most appropriate treatment of the message. It is generally assumed that messages associated with an identity that has a good reputation will be subject to less scrutiny by the receiver's filters. Although the assumption is often correct, it bears repeating that it is not a guarantee. Absent contractual agreement, a sender's use of one particular authentication technique or another does not ensure delivery or recipient viewing.

Today, there are seven deployed and openly available specifications for authenticating email:

- IP (Internet Protocol)
- OpenPGP
- S/MIME
- BATV
- Sender Policy Framework (SPF)
- Sender ID Framework (Sender ID)
- DomainKeys Identified Mail (DKIM)

IP (Internet Protocol) is the glue that holds the Internet together. It is the underlying mechanism for transferring data; all applications are built on top of it. In email, an *IP Address* identifies the source host system that is the direct SMTP transfer neighbor to the one that is validating the message. In networking parlance, the source is *one hop* away from the receiver and the receiver uses the IP Address of that source. The Address is obtained from this lower layer of Internet technology, independently of the message content. Email transit from author to recipient normally entails many IP hops, as it is handed from one MTA to the next, where each hop has a different source IP Address.

Until recently, this IP Address was the only method available for confirming a responsible identity during message transit. However it has significant limitations, both with the stability of the reference and with organizational alignment. Addresses change as machines move and as new Internet service providers are used. In addition, the identity actually is intended for the operator of an access system rather than the content author or their organization. These limitations with IP Addresses are what motivated development of more recent techniques.

OpenPGP and *S/MIME* use cryptographic techniques to provide long-term authentication and privacy for message content. They were not tailored to use during email transit and they have not developed any momentum for that use. Consequently, they are not considered further in this white paper.

BATV **(Bounce Address Tag Validation)** provides authentication for messages sent to the SMTP `MAIL FROM` Return Address. It is used by the recipient of an email handling notice to verify that the notice is likely to be valid. As such, this mechanism is for a different concern than authenticating the responsible handling agent or the original message and it will not be considered further in this paper.

The remaining three mechanisms cover two major technical paradigms, use different identities and different administration models, and have different limitations and flexibilities:

- *SPF* and *Sender-ID* use a *path registration* approach
- *DKIM* uses *cryptographic* authentication

A simple basis for distinguishing these two paradigms is depicted in Figure 4, with one using *channel authentication* built into the email-handling infrastructure along the path that a message travels, and the other adding *object authentication* information to the message itself. That is, one authentication mechanism is associated with the transmission channel whereas the other is carried along with the content.

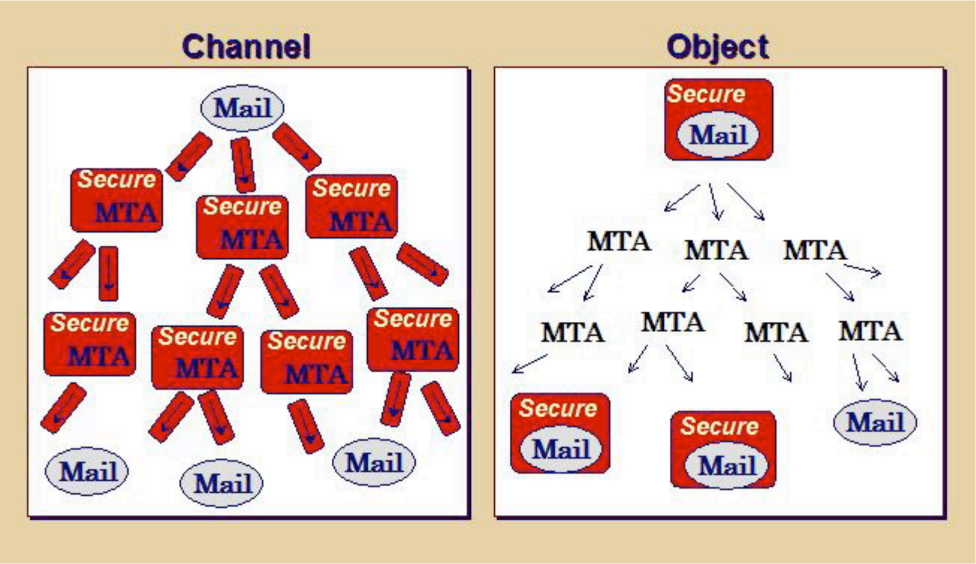


Figure 4: Approaches to Security

SMTP does support some additional authentication methods (SASL, SMTP AUTH and SSL). However these are primarily used for authentication and privacy during initial message submission of new mail, rather than during later relaying and delivery steps. Consequently, these mechanisms are useful for establishing a responsible author, under the originating organization, but are not directly useful for analysis by a receiving system.

Channel Authentication (Path Registration) Basics

Path Registration schemes build upon the established idea of using the IP Address of the neighboring SMTP MTA sending client. They extend it by mapping that address to a Domain Name, like `example.com`, and querying the DNS to see whether the IP Address is associated with that Domain Name as shown in Figure 5. This gives an organization a means of publicly registering the MTAs that are authorized to send mail on its behalf. The validating site checks the address of the neighboring server that is sending the email, against a registered list of servers that the domain owner has authorized to send email.

Note:

> Formal system security discussions that use the words *spoof* and *forge* pertain to impersonation through the unauthorized use of an identifier. Within email path registration authentication discussions, the terms are used to refer to any mismatch between the domain name in one email field and the domain name in another, even when it is authorized. For example, the SMTP `MAIL FROM` Return domain (usually displayed as "`Return-Path:`") might not match the author's `From:` address, or the author's `From:` address might not match the posting agent's `Sender:` address. These occurrences can be entirely valid, but they can cause problems for path registration schemes. In the email newsletter header in the example below you can see that even though the message is valid there are numerous, different domain names being used.

Example:

```
Received: from mta1.esp1234.com (HELO mta1.esp1234.com) (10.0.0.1)
  by mailserver.company.com with SMTP; 28 Mar 2008 19:53:28 -0000
Date: Fri, 28 Mar 2008 14:53:27 -0500 (CDT)
From: "Author" <author@authorscompany.com>
To: Recipient@company.com
Subject: March Newsletter
Sender: authorscompany@esp1234.com
Return-Path: bounce-4101674@authorscompany.esp1234.com
...
```

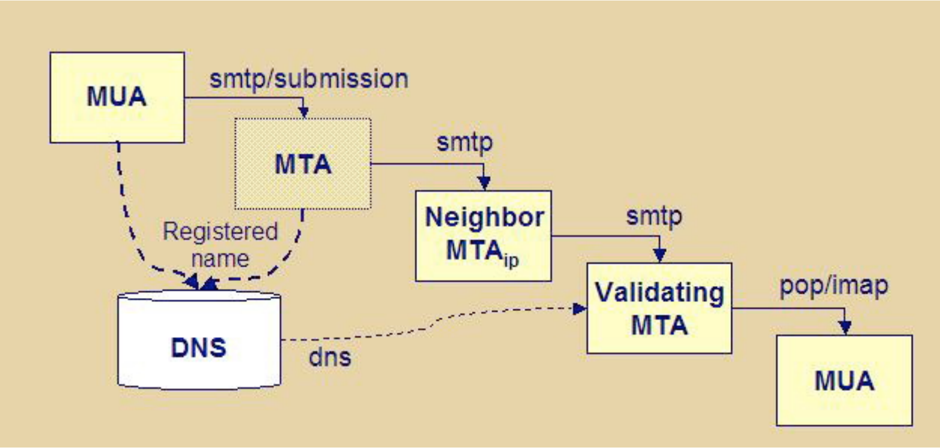
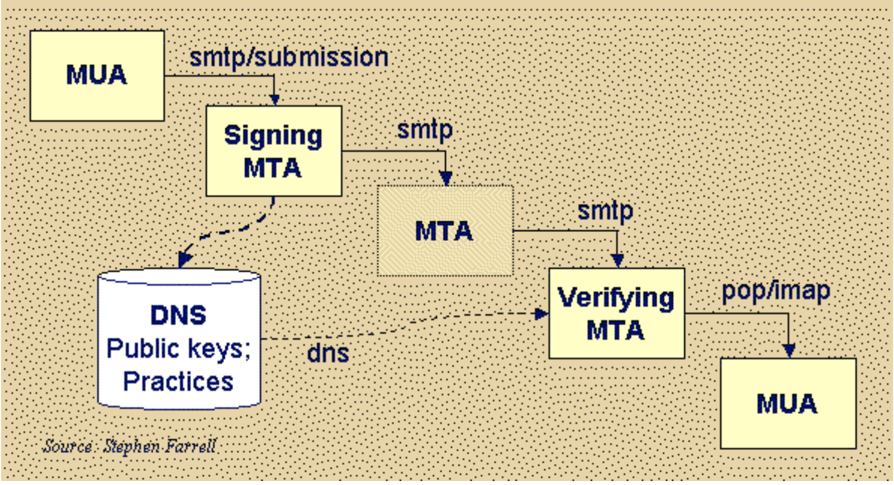


Figure 5: Path-Based Authentication DNS Use

Object (Cryptographic) Authentication Basics

Cryptographic authentication performs mathematical calculations on selected message content. These produce a tiny digital summary, called a *hash*, of the message content that is extremely difficult to reproduce without a copy of the content. Additional calculations are then performed to protect the hash, so it cannot be modified without detection. The signer uses a private (secret) cryptographic key to create the signature and the validating side uses a corresponding public key to verify it. The public key can be circulated openly, such as in the DNS, under a Domain Name of the signer, as indicated in Figure 6. In fact, this form of



Source: Stephen Farrell

Figure 6: Cryptography-Based Authentication DNS Use

publication is self-certifying; by virtue of finding it under the domain name for which it is used, the public key is known to be valid. When validating a signature, the hash is re-calculated. If the message has been modified, the hash will not be correct and the validation will fail.

Considering the Alternative Schemes

This white paper uses a template for discussing the basic characteristics of the three authentication technologies:

- Identity that is authenticated
- Authentication mechanism
- DNS query mechanism
- Effort to add to a service for the message origination side
- Effort to add to a service for the message reception side
- Technical limitations – when does the technique not work?

By virtue of there being multiple authentication techniques, a very real question presents itself: What is the effect of using more than one? That is, will it be helpful for message delivery or could the presence of multiple technologies in the same message interact in counter-productive ways? The current expectation is that using more than one mechanism will not cause problems and well might have some benefits.

Sender Policy Framework (SPF)

Sender Policy Framework (SPF) was developed out of the Internet mail anti-abuse operations community. It is published as an IETF Experimental Specification (http://tools.ietf.org/rfc/rfc4408.txt).

Identity that is authenticated

SPF uses the IP Address of the SMTP neighbor and maps it to the Domain Name in the `MAIL FROM` Return command of SMTP (a.k.a "`Return-Path:`") and/or the `HELO/EHLO` SMTP command. The latter name is explicitly provided by the neighboring SMTP client host to label itself.

It is probably more helpful to view the IP Address as the identity, with the mapping being useful for aggregating a number of different MTAs' IP Addresses under the same organizational reputation.

Authentication mechanism

SPF uses path registration. A site that is validating a message receives it from a neighboring MTA. It uses the IP Address of that neighbor and the Domain Name in the SMTP `MAIL FROM` Return and/or the `HELO/EHLO` commands for the message. Validation consists of finding the IP Address registered under the Domain Name.

Querying on the `MAIL FROM` Return command is mandatory, according to the SPF specification. Querying the `HELO/EHLO` command is recommended.

DNS query mechanism

The owner of the MAIL FROM and/or HELO/EHLO Domain Name registers a record in the DNS that contains the IP Address of each MTA that will be the closest neighbor to a validating MTA that is authorized to send mail on behalf of the domain. When the validating site queries the DNS for the domain in the MAIL FROM Return command, it will find that the neighboring MTA's IP Address is registered under it. This means that the Address is authorized to send email containing that Domain Name in the MAIL FROM and/or HELO/EHLO commands.

SPF defines two choices for recording information in the DNS. One defines a format to be applied to the existing, general-purpose TXT RR record. The other is a new SPF RR record. Because it has proved challenging to obtain widespread deployment and use of new DNS RR records, it is common to define an interim alternative, such as the SPF specification has done. One issue with having different standards specify different definitions for TXT content is distinguishing which application service applies to a particular record. SPF identifies its TXT records by including a v=spf1 parameter inside. Hence TXT records not containing it are ignored. (This same v=spf1 record may be interpreted slightly differently under Sender ID, as described in the next section.)

Whether coded as a TXT or SPF RR, the SPF DNS record is intended to be a rather flexible means of publishing a variety of email service practices information rather than only for registering authorized systems. This includes registering addresses that are not authorized, alternate mechanisms that are authorized, and even recursive references that derive authorization information from other records. This flexibility can make it challenging to create records that accurately reflect the policies of a registering organization. Consequently administration software has been developed to facilitate the process of specifying a SPF DNS record for the most common configurations.

Example:

Header from an email newsletter:

```
Received: from mta1.esp1234.com (HELO mta1.esp1234.com)(10.0.0.1)
  by mailserver.company.com with SMTP; 28 Mar 2008 19:53:28 -0000
Date: Fri, 28 Mar 2008 14:53:27 -0500 (CDT)
From: "Author" <author@authorscompany.com>
To: Recipient@company.com
Subject: March Newsletter
Sender: authorscompany@esp1234.com
Return-Path: bounce-4101674@authorscompany.esp1234.com
...
```

The validating MTA will extract the MAIL FROM and HELO/EHLO domains as authorscompany.esp1234.com and esp1234.com respectively. When the validating MTA queries DNS for these domains it receives the following SPF record for both domains:

```
v=spf1 ip4:10.0.0.1 mx ~all
```

The validating MTA then compares the neighboring MTA's IP address from the Received: header to the IP address or addresses in the SPF record and determines that the neighboring MTA is authorized to send email for the Domain Name in the MAIL FROM and/or HELO/EHLO commands.

Effort to add to a service for the message origination side

The owner of the Domain Name that is specified in the MAIL FROM and/or HELO/EHLO command must register each MTA's IP address that might send mail to another MTA – which in turn might evaluate the sending MTA's reputation. Typically, Outbound Border MTAs are registered, so that Inbound Border

MTAs can do the validation. As the IP Addresses of registered MTAs change, the registration records must be updated.

On the origination side, the only software required for supporting SPF is a DNS administration tool that permits creating the necessary types of DNS records.

Effort to add to a service for the message reception side

A validating site obtains the IP Address of the neighboring MTA and the Domain Name contained in SMTP `MAIL FROM` command. It performs the DNS query to determine whether the MTA is registered. Software must be added to the recipient component that performs the validation.

Limitations – when does the technique not work?

The `MAIL FROM` Return address tends to match the domain name of the author or of the posting agent, but this is not required. Valid uses of the `MAIL FROM` Return address can be an entirely different domain, such as when processing bulk mail requires that error notices go to a special address. In such cases, SPF validation will fail unless the domain of the address for processing error notices also has SPF the appropriate SPF record.

Email often is sent directly from an originating organization to a receiving organization. SPF works for these cases, as long as:

- The originating organization is able to identify all of its Boundary MTAs.
- It keeps them correctly registered.
- Validation is performed at the Inbound Border MTA or the Inbound MTA safely adds the IP Address of the neighboring Outbound Boundary MTA to the message – so it can be evaluated later.

However, email that is sent through an intermediary, such as a forwarding service or other re-posting agent, will fail an SPF validation, unless the intermediary is explicitly registered under the SPF domain name. Mail that is sent through unregistered MTAs will fail validation. One example of this occurs when a mobile user is forced to post mail through an access provider's MTA, rather than being able to post through their home organization's email service. Another popular example is mail that is sent through an outsourced serviced, such as when subscription bulk email is handled by a contractor.

SPF defines a mechanism for dealing with `MAIL FROM` Return domain names that do not conform to SPF registration requirements, called *Sender Rewriting Scheme (SRS)*. It modifies the Return address to point to the new, registered domain, but permits recovering the original Return address. This requires that the intermediary system be modified to participate in the SPF mechanism. SRS is not commonly used today.

Sender IDentification Framework (Sender ID)

Sender Identification Framework (Sender ID) is an authentication protocol from Microsoft that was created by the merger of Sender Policy Framework (SPF) and Microsoft's Caller ID for Email. It is published as an IETF Experimental Specification (http://tools.ietf.org/rfc/rfc4406.txt).

Identity that is authenticated

Sender ID permits mapping a neighboring MTA's IP address to a *Purported Responsible Address (PRA)* domain, or to an SPF `MAIL FROM` Return domain. The receiving site decides which of these they will use for analysis.

The PRA is obtained through a multi-step analysis that inspects a number of fields within the message header. The approximate hierarchy of sources for the PRA is:

1. `Resent-Sender:` header field
2. `Resent-From:` header field
3. `Sender:` header field
4. `From:` header field

The PRA can be carried through an optional *Responsible Submitter* SMTP protocol enhancement.

The principal in having a PRA domain, versus `MAIL FROM` Return domain, is that the PRA covers fields that are typically viewed by users, such as the `From:` field. It is hoped that this will improve protection against some social engineering exploits, including phishing and deceptive email campaigns.

Authentication mechanism

Sender ID uses path registration. Within Sender ID, SPF DNS records list outbound mail servers, and other servers, associated with a domain and authorized to send mail on their behalf along with their respective IP Addresses. An organization publishes the SPF record to its associated DNS servers, which is then checked by recipient mail servers.

Sender ID obtains the sending host's IP Address from the incoming SMTP neighbor's TCP/IP connection. Sender ID acquires the neighbor's domain name from the SMTP Return address or from a message header field determined by the PRA. Sender ID then queries the DNS for the SPF record that lists all authorized SMTP hosts. The authentication is authoritative because the DNS server that hosts the queried SPF record is controlled and secured by the registering organization.

DNS query mechanism

The Sender ID use of the DNS is derived from the SPF use. It permits use of SPF's existing `v=spf1` records, albeit choosing the domain names according to Sender ID rules rather than SPF rules. In addition, Sender ID defines a `v=spf2.0` record to contain additional parameters.

The scope parameter defines the hierarchy of SPF, PRA and other naming mechanisms from which to choose. That is, each registered name can specify to which identifiers in a message it can be applied.

Effort to add to a service for the message origination side

Sender ID has essentially the same requirements as SPF, although one difference is that Sender ID defines two different DNS records. The first is the same as the SPF record and the second is a revision to it.

As with all security mechanisms, to maintain the SPF record it is necessary to employ a change management process. That process will define how new sources of SMTP messages from your domain are communicated and approved for addition to the SPF record. It is recommended that such changes to the SPF records should be made 72 hours in advance of any major email campaign, to allow for DNS replication and remote DNS record cache deployments to be updated.

Effort to add to a service for the message reception side

Sender ID has essentially the same requirements as SPF, except for determining which domain name to select for querying the DNS.

Limitations – when does the technique not work?

Because its primary operation does not employ the domain name in the SMTP `MAIL FROM` command, Sender ID does not encounter problems when that domain does not match the content `From:` or `Sender:` domain. However it has the same challenges as SPF with respect to re-posting, such as by mailing lists and forwarding services.

DomainKeys Identified Mail (DKIM)

DKIM is based on cryptographic content signing. It was produced through a merger of Yahoo!'s DomainKeys and Cisco Systems Inc.'s Identified Internet Mail (IIM) specifications. First developed by an informal industry consortium, it was then revised and has been published by the IETF as a Proposed Standard (http://tools.ietf.org/rfc/rfc4871.txt).

Identity that is authenticated

DKIM allows the signer to choose *any* Domain Name, which is indicated in the `DKIM-Signature:` header field of the message. Whether that Domain Name is related to another identifier in the message, such as the `From:` or `Sender:` fields, is a separate decision. As DKIM usage develops, common practices for this choice will become apparent and are likely to be standardized. The IETF is specifying a means of publishing these kinds of email signing practices, but DKIM can be useful without them.

Authentication mechanism

The responsible organization adds a digital signature to the message, associating it with a domain name of that organization. Typically, signing will be done by a service agent that is part of the message author's organization or delegated by them. Signing might be performed by any of the components in that environment, although most often it is the MSA or MTA that adds the signature today. DKIM permits signing with a particular domain name to be performed by authorized third parties, such as having an originating organization obtain a signature by an independent assessment (reputation) organization and affixing the signature to the message.

DNS query mechanism

As with SPF and Sender ID, DKIM envisions a new DNS resource record but defines a `TXT` record for initial use. It is placed under a special sub-domain of DNS, which is underneath the domain name declared in the `DKIM-Signature:` header field. Any `TXT` records under that sub-domain name are only for DKIM use.

The special portion of the sub-domain has a field, called the *selector,* which is used for key management. So the DNS query string has multiple fields, with only a portion intended to be used for actual *reputation* assessment. That is, a core domain name represents the organization. It then is combined with an administrative sub-name so that keys can be assigned more conveniently. This is necessary for control over signing by different individuals or systems, as well as for migrating to a new key.

The DKIM DNS record has some parameters for constraining its use to particular services or addresses. However the record only validates an existing signature. Publishing additional email signing practices (SP) for the domain is the subject of separate follow-on work.

Example:

Header from an email newsletter:

```
Received: from mta1.esp1234.com (HELO mta1.esp1234.com) (10.0.0.1)
  by mailserver.company.com with SMTP; 28 Mar 2008 19:53:28 -0000
DKIM-Signature: v=1; a=rsa-sha1; c=relaxed/relaxed; s=esp1234;
d=authorscompany.com;
 h=From:To:Subject:Mime-Version:Message-ID:Content-Type:Date;
i=author@authorscompany.com; bh=EMR7D1qC7ykz41K8ArLCt++IWxM=;
 b=TGkNEq7fW4OIno/5DlX2qHDQeRmzhY+uiTzEcxu2KIKC+4B7+i2olIWGZP9JBnOR4Ck6iAiidnRj
   DLuc2QJh3ifDNPWJ6xYjiuE73ilCZfbtN0r2MVke9pRU4aydBQ5DSCFS7YhUFB22CT70MutZkaDF
   SZZpqI5vTlSWm9MI8PM=
Date: Fri, 28 Mar 2008 14:53:27 -0500 (CDT)
From: "Author" <author@authorscompany.com>
To: Recipient@company.com
Subject: March Newsletter
Sender: authorscompany@esp1234.com
Return-Path: bounce-4101674@authorscompany.esp1234.com
Mime-Version: 1.0
Message-ID: <20080324040103985572.328428@mx12.emailroi.com>
Content-Type: multipart/alternative;
boundary="============_emailROI_============"...
```

Corresponding DKIM DNS record:

```
nslookup -type=txt esp1234._domainkey.authorsdomain.com
Non-authoritative answer:
esp1234._domainkey.authorsdomain.com       text =

        "g=; k=rsa;
p=QIdfMA0GCSqGSIb3DQEBAQUAZ4GNADCBiQKBgQC61RrUNTIcNbf/+f5Co2V37GMvPQdbUVyjgvLXrU
KAXeJDwYVumAtE9BovuDZNYxcgG2oy7mkcZX/3rBF5SJX9Cp5yw0axuMpzkuzPQq26h+2+MLuvtJtfDI
aHgNeEJOjMeq7s9RFQHRr9g26lkZQTRAob8YevaA1KHiNNyIaZuQIDAQAB;"
```

Effort to add to a service for the message origination side

DKIM signing can be performed anywhere in the originating or relaying path. A common example is a department MTA or an organization's Outbound Boundary MTA. Signing requires addition of a software module that can:

- Obtain the private key
- Perform the necessary calculations
- Affix the signature information to a `DKIM-Signature:` message header field

Effort to add to a service for the message reception side

DKIM validation can be performed anywhere along the message transit path after the message is signed. Validation requires a software module that can:

- Obtain the public key
- Perform the necessary calculations
- Affix the results to the message in a trusted manner or hand the results to an evaluation engine

Notably this means that the signature can be used by the recipient organization's filtering software, rather than requiring the recipient end-user to make an assessment.

Limitations – when does the technique not work?

Some mail transit behavior can modify message content in a way that breaks an existing signature. DKIM has some features that provide robustness against some common modifications; however this robustness is (intentionally) limited, since it would otherwise open the door for abuses.

The DKIM *selector* mechanism in the DNS uses a special keyword `_domainkey`. The use of the underscore can cause confusion. Although not allowed in the use of names that refer to Internet hosts, the underscore is entirely valid within domain names more generally. It is increasingly used to distinguish special sub-names, underneath host names, and for storing parameters associated with the host name.

Conclusion

Receivers' efforts to detect abusive email are confounded by the efforts of Bad Actor senders who are seeking to avoid detection and accountability; therefore, analysis techniques are forced to deal with fuzzy information. In contrast, trust-oriented systems permit senders to provide explicit, accurate and reliable information with the intent that it will permit receivers to expedite message handling.

Authentication is the foundational component of trust-based message processing because it provides a confirmable identifier. SPF, Sender ID and DKIM all use Domain Names as identifiers that are coupled to an organization taking responsibility for a message. SPF uses an identifier from the underlying email transport mechanism. Sender ID is based on author information. DKIM uses an identifier that is independently specified, with an emerging Sender Practices extension allowing it to be coupled to the author.

References

Email:	Overview	– Internet Mail Architecture
	SMTP	– RFC 2821
	Internet Mail Format	– RFC 2822
	Body Format (MIME)	– RFC 2045
	POP	– RFC 1939
	IMAP	– RFC 2060
	SASL	– RFC 2222
	SSL	– RFC 2246
DNS:	Overview	– RFC 1034
	Specification	– RFC 1035
	ICANN	– icann.org
SPF:	Home Page	– www.openspf.org
	Overview	– www.openspf.org/Introduction
	Specification	– RFC 4408
Sender ID:	Home Page	– www.microsoft.com/senderid
	Overview	– Sender ID Framework Overview
	Specification	– RFC 4406
DKIM:	Home Page	– dkim.org
	Overview	– DomainKeys Identified Mail (DKIM) Service Overview
	Specification	– RFC 4686