Network Working Group
Internet-Draft
Intended status: Informational
Expires: September 12, 2019

S. Blank
Valimail
N. Kumaran
Google
J. Levine, Ed.
Standcore LLC

March 11, 2019

An Overview of the Design of BIMI
draft-bkl-bimi-overview-00

Abstract

Brand Indicators for Message Identification (BIMI) provides a mechanism for mail senders to publish a validated logotype that mail receivers can display with the senders' messages. This document provides a brief overview of BIMI and examines some of the trade offs and decisions in its design.

Discussion venue

Comments on this draft may be directed to the BIMI list at bimi@ietf.org.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 12, 2019.

Table of Contents

1. Introduction

Brand Indicators for Message Identification (BIMI) provides a mechanism for mail senders to publish a validated indicator, which is typically a logotype, that mail receivers can display with the senders' messages.

For senders, an indicator published with BIMI is available to any recipient without having to make individual arrangements with each recipient system. BIMI indicators are only displayed with messages that are validated with DMARC [RFC7489]. For senders that want to have their indicator displayed ("brand impression") this may provide an incentive to make their mail streams DMARC compliant.

For mail systems that display indicators next to messages, BIMI provides an automatic way to obtain indicators that have been validated by third parties, without having to individually research or contact each sender.

2. High level architecture and data flow

A sender publishes its indicator by putting a file containing the indicator's image on the web, or possibly several images if they want to show it in various forms. The sender may arrange for a third party to verify that the sender is authorized to use the indicator.

The sender then publishes one or more DNS records with names like selector._bimi.example.com, containing tags with the URIs of the indicator(s), and URIs of validation information if any. If it wants to associate different indicators with different messages, it can publish several DNS records with the various indicator URIs, and the message can pick a selector (see below.)

Then the sender sends mail the usual way, arranging for it to be DMARC validated. If it wants to pick a particular indicator, a message can include a BIMI-selector header to pick a particular selector. Otherwise it implicitly uses the "default" selector. (Note that this allows mail to use BIMI without having to include any new BIMI material at all.)

The recipient system receives the mail and does DMARC validation as usual. If DMARC passes, it does a DNS lookup of the BIMI record using the DMARC validated domain and the appropriate selector. It fetches the indicator(s) using the URIs in the DNS record, may check the third party validation, and may make other checks to decide whether the indicator is usable.

Assuming that all worked, the recipient system adds the usual Authentication-Results header with the DMARC and BIMI validation results, and a new BIMI-Location header that contains the URIs of the indicator files to the message. An MUA, either one integrated into a webmail system or a separate one, can use the BIMI-Location header URIs to fetch the indicators to display to the user, per its own policy.

3. Risks and problems of BIMI

BIMI inherently assumes that showing indicators with mail messages is desirable, and standardizing how this is done is of value for senders and perhaps for recipients. Some mail systems have been showing logos and other per-sender images such as avatars or pictures of the sender for a long time.

Since indicator validation requires human effort, there will always be avenues to show indicators that are confusing or malicious. BIMI attempts to contain those abuse vectors in its validation process and ensure they are mitigatable at scale. Even assuming abuse is successfully contained within the validation mechanisms, there are other inherent risks and potential problems that could arise out of the usage of BIMI if it is improperly architected or implemented. Some of the most concerning are detailed below. The trade offs between potential problems and benefits remains unclear.

3.1. Private club

Currently large mail systems show indicators of large senders using fragmented methods to identify the senders' mail and the indicators to use. One of the goals of BIMI is to open up the process to organizations that don't have access to representatives of receiving systems, or the ability to go through a separate process for each receiver, but if implemented poorly, it just changes the secret handshake. While many domains may publish validated logos via BIMI, others will be unable to or intentionally decline to participate. Similarly BIMI can identify some set of indicators to tie to senders, but will never handle a complete set of validated indicators either.

Depending on DMARC means that only senders with their own domain can use BIMI. It can't identify many small senders that share their

providers' mail domains, e.g., a girl scout troop using a single address from their provider, such as troop42@bigisp.com, even though the troop has a license to use the Scouts' trefoil logo. A modified design could address this, described later.

Conversely, if the indicator has to be a registered trademark, that limits BIMI to organizations big enough to pay for and maintain a trademark registration (in the US typically several thousand dollars to get a trademark, and several hundred dollars every few years for maintenance.) Some countries have no trademark registries, or do no meaningful examination, which means that people in those countries are out of luck. On the other hand, if the indicator doesn't have to be registered, see the next section.

3.2. Inconsistent validation

To the extent that BIMI has multiple validation methods, it will produce different results on different mail systems. If systems use their own reputation data to decide whose indicators to show, indicators that appear on one system may not appear on another. If there are third party validators (certificate authorities or the like), there's no reason to assume that everyone will accept the same set of validators, any more than every HTTPS client trusts the same set of CA's.

Determining whether an indicator in an SVG file is the same as one in a trademark registry is non-deterministic. The picture in a trademark registry is typically a low resolution monochrome image. (For example, see the copy of the IETF logo at https://www.tmdn.org/tmview/trademark/thumbnail/US500000078755496). There is no mechanical way to decide whether two images in different resolutions and different sizes and colors drawn different ways are "the same". At some point, people will have to look at them and decide, and they will not always decide the same way. There is limited existing practice to work from, primarily trademark examination to determine whether a proposed trademark is "substantially similar" to one registered elsewhere. Trademark legal disputes are invariably about whether marks are likely to be confused, not whether they are the same, and in any event the resolution of the dispute generally involves lawyers and expert witnesses. Dispute resolution about indicators is unlikely to work at Internet scale.

For the vast numbers of organizations that don't have registered trademarks, the problem is even worse since there's no agreed way to find the reference mark to compare the indicator to.

And, of course, there's no guarantee that the third party validators will be consistently honest, competent, and diligent. Experience with certificate authorities suggests that is improbable.

3.3. Pay to Play

DMARC started with limited use to protect heavily phished domains, with paypal.com being the usual example. As DMARC rolled out in other fashions (such as AOL and Yahoo! using it to deal with address book theft), it soon became widely used and increasingly required. This is because DMARC is a fairly effective tool to deter exact domain phishing attacks and email-borne fraud. However, DMARC has proved difficult to deploy for many organizations. While there are some free resources to help people get their DMARC set up, in practice it means that businesses pay someone to make their DMARC work. Similarly, although the current plan is for BIMI to be totally optional, it's not hard to see it going down the same path, especially if logo validation is perceived by senders as a signal to receiving systems to help disambiguate fraudulent mail streams. Even if there are paths to getting a validated BIMI indicator without a full trademark registration, it's likely to be complex enough that most organizations will need a consultant. Hence it would end up as another expensive hoop to jump through just so you can send some mail to your clients.

Even if you do your BIMI implementation yourself, third party validators will not be free. Validation is labor intensive, both the process of determining that your domain matches the organization that owns the indicator, and the process of comparing the reference indicator to the proposed indicator. If it's a derivative or otherwise modified mark (the girl scout troop combines the trefoil with a picture of a local landmark) the process is even more labor intensive. The work is similar to what trademark examiners do, and they are attorneys with experience in trademark law.

3.4. User Security

In the current draft, every time BIMI causes an MTA or MUA to fetch a remote resource, whoever runs that resource can tell something about the recipient. As noted below, fetches in the MTA right after the message has been received leak relatively little, particularly on large systems where the resources are likely to be cached locally. Fetches from the URL in the BIMI-Location header are a definite web bug. This can be fixed relatively easily as described later. The risk here is with smaller mail systems or MUAs that do not have means to cache the logos.

4. Indicator Publishing Options

The indicator(s) are published on the web as SVG files or potentially other formats.

For a domain owner that wishes to assert a BIMI indicator, the process must be as lightweight and transparent as possible. There are multiple options, each with its own inherent value and threats, that are possible to make this assertion.

4.1. Message header

BIMI indicators could be published by adding them in a new header field in participating messages, along the lines of a validated X-Face header. This mechanism is simple, but has several disadvantages and threats:

- It requires all participating messages to have the header field; which is neither lightweight nor transparent.
- It requires per-message validation of the header field, instead of being cached at the domain level.
- It does not allow for the indicators to be pre-fetched and validated in advance, which opens attack vectors from CGI scripts and the like.

4.2. S/MIME signatures

In principle Indicators could be asserted through S/MIME certificates. They don't share the same attack vectors as asserting through a naked header field, and they're self-validating to the extent the MUA trusts the signer. S/MIME is unlikely to work at scale for a variety of reasons:

- Many senders don't have the skill to do the signing,
- major webmail systems don't interpret S/MIME signatures,
- and S/MIME certificates have all of the process issues of any other certificates.

Requiring S/MIME for BIMI to function is neither lightweight nor transparent for domain owners.

4.3. DNS assertion records

Publishing an assertion record in DNS has several advantages:

- There is exactly one mechanism for both simple and complex policies to be published.
- Operational complexity is reduced, and MTAs only need to check a single record per sending domain in a consistent manner to enforce policy.
- Indicators can be verified and/or cached in advance, so that malicious headers cannot be used as an attack vector.
- By publishing a single DNS record, the indicator can be consumed by all receiving mail systems without needing to make any modifications to the email in transit, allowing for transparency regardless of the source of email.
- This looks and feels similar to a DMARC record, meaning a domain owner that is already using DMARC is familiar with the mechanism already.

The downsides and threats of publishing an indicator in the DNS are that this requires BIMI be a domain-based standard, and can be compromised if a domain's DNS is compromised or hijacked.

5. Validation Requirements

For an indicator to be considered validated, there must be a confirmed mapping among organization, domains, and indicators. To prevent fraudulent representation of any of this mapping, all of the following criteria must be met:

- The organization is legitimate
- The domain names are controlled by the organization
- The party requesting validation is currently authorized to do so by the organization
- The party requesting the validation is who they say they are
- The organization has current rights to display the indicator

The final requirement, that the organization has rights to display the indicator, is fraught with nuance, and BIMI could become a vector for phishing unless this is handled properly. What follows are the

different types of indicators an organization might wish to prove they have the rights for, and how they can attest to those rights.

None of the below attestation of rights are full proof; meeting these requirements does not mean a mail system will consider that sufficient evidence to display the indicator. It is expected that both validators and receiving systems will choose which of these use cases they will support and which they will not.

5.1. Indicator Usage and Rights Validation Scenarios

There is a wide variety of ways that an indicator might be associated with an organization and its domain(s), and how the association can be verified.

The intent is not for a domain owner to know which of these use case their indicator falls into and make a request accordingly, but for a validator to make sure they use appropriate validation criteria, given the type of indicator they are provided with.

Further, in the short term, validators may choose to only validate indicators that can be verified using the highest standards, waiting to verify more complicated use cases until they're certain of their verification processes and confident they can prevent fraudulent requests against new types of indicators.

5.2. Registered Mark

Usable when an indicator has been registered in a well known and trustworthy jurisdiction.

To prove a logo is associated with the organization, the registration, jurisdiction, issuance date, and expiration must be provided, and the registrant of that mark must match the organization requesting indication validation.

The problem with registered marks, beyond what is discussed in Section 3.3, is that trademarks are not anti-phishing mechanisms. Rather, they are anti-confusion mechanisms within tightly scoped silos and jurisdictions, and they see frequent abuse and have slow and costly remediation processes.

5.3. Registered Mark (untrusted jurisdiction)

When a registered mark in a jurisdiction that is untrusted is provided, all the evidence should be reviewed, but validation should proceed as if a Common Use Mark had been provided, not a Registered one.

5.4. Common Use Mark

A widely used but unregistered indicator. Without providing for Common Use Marks, BIMI will certainly become a private club. This is, however, the most difficult use case to provide rights for. You must prove active usage of the indicator in the real world, and provide documentation showing that the rights to the logo have been properly assigned to your organization.

The threat here is cousin or lookalike marks. While BIMI has no direct protections for lookalike marks, allowing for Common Use Marks makes it far easier to validate such a mark.

5.5. Common Use Mark (untrusted jurisdiction)

Validation similar to above.

The threat in untrusted jurisdictions for lookalike or cousin domains is much higher.

5.6. New/Rebranded Mark

A Registered or Common Use Mark that is not in wide enough use to pass the Common Use requirements.

Many organizations come out with new products or wholly rebrand from time to time. These common changes should be able to be reflected in BIMI, or systems that displayed BIMI indicators could be out of date for significant period of time until Common Usage could be proven or even longer until a registration was completed.

Validation is the same as for Common Use, but instead of proving usage in the real world, you must prove intent to use.

The threat here is that there is that bad actors could attempt to use this use case to validate arbitrary logos.

5.7. Mildly Altered Mark

For an indicator to look good within a BIMI context sometimes mild changes are required (i.e. to take a long logo and stack it so that it will fit within a square, to remove a background color, etc.). Specifically, a Mildly Altered Mark is a logo of one of the preceding types of marks that has been altered minimally (recolored, cropped, rotated, stretched, or reformatted or sliced and repositioned to fit in a different aspect ratio). These mild modifications aren't new indicators worthy of registration or wouldn't pass common usage tests of their own.

To validate a mildly altered mark, you must validate the unaltered mark, and the validator must further attest that the modifications are minimal per the above list and only for the purpose of working better within a BIMI context.

The threat with mildly altered marks is that any defenses depending on automated detection will be confused by the alterations, which bad actors would presumably take advantage of. Further, this relies on a human saying that an alteration is mild, which different people will do differently.

5.8. Multiple Marks

A logo comprised of more than one of the preceding types of marks, all from the same organization, with one or more marks potentially obscured by others.

Validation requires validating each primary mark individually, and the validator attesting that any obscured mark is still distinguishable at its original self. Again, automated detection could be confused by this, which bad actors can take advantage of. It is also possible that obscured marks could be a vector to create confusable indicators.

5.9. Derivative Mark

A logo comprised of more than one of the preceding types of marks, to which new imagery has been added, potentially obscuring the initial mark(s). Derivative marks are common, as organizations frequently slap other images on top of their logos, such as in seasonal or geographic campaigns. Providing this functionality dramatically expands the approachability of BIMI for many organizations' needs.

As with Multiple Marks, anything that obscures the original mark could create a new confusable indicator.

5.10. Co-marketing

Multiple marks or derivative marks, but from separate organizations, where there is joint permission for use. This occurs frequently with partnerships.

Validation is the same as for multiple marks, but each organization must prove all validation requirements for its mark, and all organizations must provide proof that their indicator can be used in conjunction with the other. The expiration of the co-marketing must also be provided for.

Validating a co-marketing indicator would likely depend on validating the underlying marks first.

Threats here are again related to one indicator obscuring another, or two marks being put together in such a way as to become confusable with a more well known one. .

5.11. Franchisee

It is common for an organization to have the rights to display an indicator without being the owner of the indicator.

To validate this, the initial indicator must be validated, and the franchisee must prove they have the rights to display and prove they are in good standing with the licensor and those rights have not expired.

The indicator would likely have to be validated by its owner before franchisees can be granted their own validated indicator.

Threats are franchisees who were legitimate when they got the indicator validated, but then lost standing and continued to use the indicator.

6. Validator options

There are multiple ways that a recipient system might validate a logo.

- Third party validation: a mutually trusted third party attests that the domain is authorized to use the logo. There are various ways the third party could publish the attestation.

- Local reputation: large systems track sender reputation, and could accept logos for senders with good reputations.

- Remote reputation: There might be a shared reputation service, or receivers could use sender characteristics like a domain in a TLD such as .BANK or .NGO that limits registrants to known Organizations.

- Self validation: some systems know their own customers and/or are whitelist based, and may not require validation for published BIMI indicators so long as DMARC passes.

6.1. Pros and cons of validation approaches

Third party validator

Pros:

- Scalable solution that takes much of the burden off of senders and receivers
- Senders bear a predictable cost for vetting, rather than receivers bearing the bulk of the burden, which allows smaller receivers to participate
- More flexibility and lower technical bar for domain owners

Cons:

- Some cost component (e.g. senders still have to pay.)
- Does not exist today; (almost) entirely new.
- Smaller receivers may not be able to tell which third parties are reliable.

6.2. Receiver Managed Reputation

Pros:

- If a receiver already trusts a brand, validation may be more straightforward
- Easy to set up; no additional steps on the sender side

Cons:

- Impractical for smaller receivers due to lack of adequate local or shared reputation data.
- Reputation can be gamed.
- Lack of transparency and consistency
- Smaller senders may not send enough mail to generate trust data.

6.3. Remote Reputation

Some registries limit registrants to verified members of a group such as NGOs, co-ops, or banks.

Pros:

- Leverage existing validation at TLDs or similar groups.
- Easy to use, just see if the domain's in a validated name tree.

Cons:

- Limited coverage
- Unknown and highly variable quality of remote reputation validator. (Some TLDs only validate at application time, never check again later.)

6.4. Centralized system

Similar to government, e.g. USPTO, JIPDEC

Cons:

- Risk of too much concentration of power, too slow in terms of responsiveness to customers and ability to execute, lack of incentives.

6.5. Self validation

Pros:

- No overhead to publishing an indicator

Cons:

- No validation; no assurances about indicator
- Too easy to publish fraudulent indicators
- No meaningful way to evaluate as a receiving system unless the domain is on a whitelist, in which case it's like remote reputation.

6.6. Third Party Validation Publishing Options

Once a third party has validated an indicator, how can it tell consuming systems about it?

6.6.1. Publishing validation: Certificates

A third party validator could act as a CA and sign certificates that contain the sender's domain and logo, or for large logos perhaps a hash of the logo. See [BIMICERT].

Pros:

- Open standard, well defined and understood
- Costs borne by sender, certificates published by sender. No centralized trusted authority that may fail catastrophically
- Technical security of certificates is well understood
- Certificate Transparency could be utilized, including prepublication for challenge.

Cons:

- CAs have ongoing history of mismanagement, lack of trust and responsiveness.
- Revocation doesn't really work. (See below Section 9.2.)

6.6.2. Publishing validation: API

Third party validator could provide an API where receivers could send queries with the domain and indicator and get back yes or no, or perhaps send the domain and get back a list of indicator hashes.

Pros:

- Doesn't depend on CAs.
- Revocation is easy.
- More flexible than certificates (easier to add or remove domains that an indicator is validated for)

Cons:

- Doesn't exist yet, needs to be defined, implemented, and discoverable.
- Easily subverted into closed club or expensive (to receivers) service.
- Could end up with many of the same issues that CAs have

6.7. General Issues of Validation

Is there a reproducible way an authority can decide whether a proposed SVG image is close enough to an image in a registry that is likely lower resolution and not in color? Does the expertise exist and if so is it widespread enough to support an adequate number of validation organizations?

How if at all to audit validation. CT logs or the like could prepublish certificates, and log all issued certificates to aid in forensics in case of mistakes.

How is conflict of interest handled? If third parties are paid to validate, what is their motivation to turn away business when that business appears to be fraudulent?

7. Consuming the indicator

The MTA fetches the certificate(s) and indicator(s) using the URIs in the DNS record, and checks the validation.

Assuming that all worked, the recipient adds a new BIMI-Location header to the message that contains the URIs of the indicator files.

7.1. Issues

If MUAs fetch the indicator using a URI in the BIMI-Location header, that's a web bug with privacy issues. If the MTA put a copy of the actual indicator as a data: URI in the header rather than an https pointer, that would solve the web bug problem and would also make the MUA faster. If data: is too bulky, the MTA could cache the indicator and put a local URL in the header. The MTA still has to fetch the indicator when the message is received, but since the message has just arrived, the sender knows the message has been received. Since the indicator fetch is by the MTA before the message is delivered, the sender can't use the indicator fetch to tell when the user looks at it in an MUA. The sender may be able to use the fetch to tell that the message got far enough to be worth BIMI validation.

8. Reporting

BIMI currently has no provision for reporting its use, but indicator publishers say they want it.

- It must not be a mechanism for tracking whether messages were delivered or opened. Caching certificates and indicators helps here.

- It must not be a mechanism for exposing the internals of mail systems, e.g, by leaking data about what of messages were delivered vs not.

One possibility would be a BIMI tag similar to DMARC reporting tags that reports how many BIMI eligible messages for a domain were received, but not what the MTA did with them.

9. Threats and Remediation

BIMI should have some mechanism for remediating problems that one receiver sees to make that available to all consumers in a timely manner.

9.1. Bad indicators

An obvious threat is that a corrupt or incompetent validator approves a indicator that a domain owner shouldn't display. There are two and a half somewhat different scenarios here:

1. The actual owner of the domain gets a validation with his domain and someone else's indicator.

2. A malicious third party tricks a validator into approving a totally fraudulent validation for someone other than the owner of the domain involved.

3. A malicious third party tricks a validator into validating an indicator with a lookalike domain (ub3r.com vs. uber.com).

The second scenario doesn't seem important, because you can't use the certificate unless you can send DMARC aligned mail, which you can't do unless you control the actual domain's DNS. (If the bad guy can take over your DNS all bets are off.)

9.2. Revoking Certificates

If BIMI validations are published with certificates, sometimes the validations will be wrong and a certificate will need to be revoked. As has been widely observed, certificate revocation has never worked very well. But since this would be a fairly unusual application of certificates, maybe some other approaches would work if the validators issue certificates:

1. Taking a tip from Let's Encrypt, reissue the certificates every week. Another possibility would be some sort of stapling, but if the CA has to do something anyway, it might as well reissue

2. Serve the certificates from the CA rather than from the sender. (Put something in the certificate that has to match the domain in the URI used to fetch it.) Then if the certificate turns out to be bad, the CA can stop serving it. This is essentially the API approach, but serving certificates rather than hashes. Or under the assumption that most unexpired certificates are still good, perhaps just send a hash of the certificate and get back an answer whether it's still good, which is a lot less data.

3. Since the certificates will be validated in tens of thousands of MTAs rather than in a billion PCs and phones, some sort of revocation list that doesn't scale to a billion might still scale to tens of thousands.

9.3. Geographic scope

Trademarks and indicators have geographic scope. A BIMI certificate could presumably have a list of country or other geographic codes but what is the recipient supposed to do with them? Does the MDA attempt to know where its users live? Does the code go into the BIMI-Location header (or in a certificate the header points to) so the MUA on my phone shows me different indicators in New York or Paris? This is a user interface issue beyond the protocol, but there should be a plausible story that could calm the lawyers who'd be signing off on indicator publication.

9.4. IMAP flags

BIMI invents a flag for IMAP that the MDA can set on a message to mark that the BIMI-Location header is real. It presumably stays with the message if it's moved to other folders. This opens a narrow abuse channel if a malicious IMAP client can put messages with fake BIMI-Location headers and the IMAP flag, but if your client can do that, you will have far worse problems than bogus MUA indicators.

An alternative approach would be for the MDA to add a DKIM signature with recipient system's d= that includes the new BIMI-Location header and the MUA can check that signature to decide whether to believe it. This would require some careful sanitizing to prevent spoofing but I think it's doable.

10. Informative References

[BIMICERT] Chuang, W., Ed. and T. Loder, Ed., "Brand Indicator for Message Identification in X.509 certificates", internet-draft draft-chuang-bimi-certificate-00.txt, May 2018.

[RFC7489] Kucherawy, M., Ed. and E. Zwicky, Ed., "Domain-based Message Authentication, Reporting, and Conformance (DMARC)", RFC 7489, DOI 10.17487/RFC7489, March 2015, <https://www.rfc-editor.org/info/rfc7489>.

Authors' Addresses

Seth Blank
Valimail

Email: seth@valimail.com

Neil Kumaran
Google

Email: nmk@google.com

John Levine (editor)
Standcore LLC
PO Box 727
Trumansburg, NY 14886

Phone: +1 6465701224
Email: standards@standcore.com