Dealing with malware, spam, suspicious content

Skip right to mail filtering agenda paragraph

Over the past years, phishing and ransomware have become the most rampant form of cybercrime as its volume and sophistication exponentially grows.

Unfortunately, **CERGE-EI as a publicly well-known institution is constantly targeted by such threat, along with many other public organizations** (hospitals, schools, municipal houses, government facilities, etc).

Ransomware, a form of malware designed for the purpose of extorting money from victims; and phishing, the delivery mechanism of choice for ransomware and other malware, are critical problems that we must address.

The generously opened and heterogeneous nature of the academic and research institution is extremely vulnerable to such kind of threat. Regular enterprises and other profit-making businesses are usually much more homogenous with much simpler rules and measures against the third parties (no IMAP, no access to emails from non-business devices, strict mobile device management, blocked or limited traffic etc.).

Both areas of malicious or potentially problematic emails and regular emails are overlapping; it is not easy to distinguish between them sometimes.

The most dangerous threats are usually those of the “zero day attack” nature; they usually take advantage of badly protected or misprotected email servers and domains so they can mimic the regular sender.
That's why there are introduced tools/measures like SPF, DKIM, DMARC, graylisting, defering, reputation filters, spam databases, heuristics, IP reputation lists, on-line scanners etc.

It is important to inform recipients that there may be something suspicious or misconfigured in particullar email message and that the email should be dealt with attention. Mail filtering services usually include short warning tag like [Spam] [Suspicious] [Newsletter] to the subject and eventually prepend email with the warning text.

False positive detections are sometimes inevitable, it is often sign that the other party still do not fully understand or value the necessity of keeping high credibility and trustworthiness of the organisation.

See also “**Spam, Phishing and Malware**” in separated CERGE-EI Wiki article (details about potential risks and their nature, recommendations, explanation…)

Monitoring and filtering agenda

All incoming email traffic towards CERGE-EI mailserver is automatically monitored, classified, filtered and sometimes tagged and/or rejected. There are several anti-malware and filtering systems chained together.

It is more and more complicated to distinguish between regular emails and malicious or abusing ones. There are several standardized methods whose help with the recognition of problematic content. These methods range from strict ban of well-known vectors of attack to attempts to hint recipient to be cautious.

Some of the most common issues are explained here:

- SPF hard fail - sending server is not on the allowed list provided by domain's owner and the domain owner asks for message blocking.
- SPF soft fail - [Suspicious - SPF - soft fail] - sending server is not listed among allowed ones, but the domain owner allow message passing with warning.
- SPF bad alignment - [Covert sender] - verify the authenticity of the domain sending the email by using two diffrenent header signatures in the message.
- Bad DMARC - [Bad DMARC] - the sender's domain does not have DMARC record and SPF set properly.
- DNSBL listed - [IP reputation - DNSBL listed] - the sender's IP is listed in SPAM database.
- Suspicious Newsletter - [Newsletter] - it may be found that certain newsletters are suspicious because they may actually be spam under the disguise of newsletters.
- Bad IP reputation - [IP reputaton] - emails from IP addresses with bad reputation may be discarded or quarantined. It may be dangerous to receive emails from such IPs.
- Warning Disclaimer (prepended to email) - [Newsletter] - Anti-Phishing engine cannot decide about targeting URL link (usually concealed by click spying)
- PDF macro - PDF files include the ability to execute code on your device — and that’s where the danger lies
- Suspicious content (HTML links, docs) - [Suspicious] - HTML content and attachments may contain potentially hazardous tags and attributes
- Image Spam (images, pdf) - [Image spam] - Some spammers conceal spam text as an image or PDF document.
- Deepheader analysis - [Suspicious - header analysis] - Deepheader analysis examines the entire message header for spam characteristics.

Possible Spoof

Added subject tag: ***[IPt:Possible Spoof]***

see https://en.wikipedia.org/wiki/Email_spoofing

Email spoofing is the creation of email messages with a forged sender address.

It usually happens when a sender uses different email address in “From:” field from the envelope email address (MAIL FROM:)

Legacy “legitimate use” - In the early Internet, “legitimately spoofed” email was common. For

example, a visiting user might use the local organization's SMTP server to send email from the user's foreign address. Since most servers were configured as “open relays”, this was a common practice. As spam email became an annoying problem, these sorts of “legitimate” uses fell out of favor.

Malicious use of spoofing - Phishing and business email compromise scams generally involve an element of email spoofing. Email spoofing has been responsible for public incidents with serious business and financial consequences.

Example of spoof email:

MAIL FROM: **johndoe2@gmail.com**
From: **john.doe@cerge-ei.cz**
To: **jane.dow@cerge-ei.cz**

Such email is suspicious. Some user with an account at Gmail (johndoe2@gmail.com) set his profile to use institutional email address (john.doe@cerge-ei.cz).
Problem is that such email is not sent (hence authorised) by cerge-ei.cz email server but it is sent by some third party server(google server in this case).

SPF

SPF - Sender Policy Framework

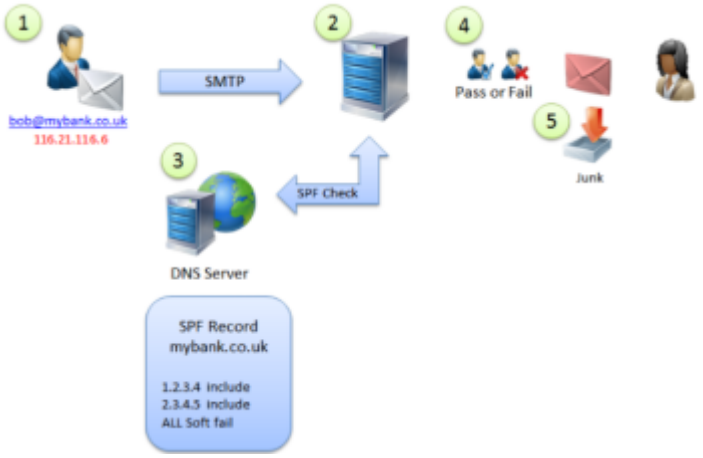
see
https://blog.zensoftware.co.uk/2014/09/02/what-are-spf-sender-policy-framework-records-all-about/

SPF main role is to define a policy that controls who, or more specifically which servers are allowed to send email claiming to be from your domain. Without SPF any email that is sent using certain domain has to be considered to be possibly valid, regardless of what server is sending it and this leads to large amounts of spoof email.

By creating a Public SPF DNS record owner of the domain can publicly announce which Internet servers are allowed to send email for our own specific domain.

If the email is not from the servers explicitly specified by the domain owner in their SPF record, there is uncertainty about the message origin and the receiving party (in this case CERGE-EI) have tool to recognize such unwanted behavior. It is solely up to the sender side to follow their own set rules, simply because they use SPF just voluntarily and hence express their intention to be respectable party with a good reputation.

Picture: How SPF works



SPF hard fail

This means that if the sending server is not on the allowed list then domain owner want the receiving server to not accept the message at all.

The point of the SPF record is to list all allowed servers either by IP, name or alternatively with simple use ‘mx’ to say all mx records for this domain.

```
example.com. IN TXT “v=spf1 mx -all”
```

SPF soft fail

Added subject tag: ***[Suspicious - bad SPF - soft fail]***

This means that the sending server is not listed among allowed ones, but domain owner wants to just inform recipient about that, not to block message completely.

The receiving server would usually accept the message but tag it as ‘suspicious’ and warn the recipient.

To allow a soft fail, domain owner use the ‘~’ tilde character rather that ‘-‘ Minus.

```
example.com. IN TXT “v=spf1 ip4:192.0.2.0/24 ip4:198.51.100.123 a ~all”
```

SPF bad alignment

Added subject tag: ***[Covert sender]*** *or* ***None***

Disclaimer:

```
***** WARNING *********************************************
This email may imitate the sender address:SPF bad alignment
Don't click any link unless you are certain it's legitimate.
If not sure or have questions, contact helpdesk@cerge-ei.cz
?: https://wiki.cerge-ei.cz/doku.php?id=public:emai:malware#spf_bad_alignment
************************************************************
```

see https://mxtoolbox.com/dmarc/spf/spf-alignment

SPF Alignment is the alignment of two headers found in an email message, meaning the value found in those two headers (a domain) needs to align with one another.

These two headers are evaluated during SPF validation testing, at which point the server that received the email will compare two headers in the email, which are:

1. The <From:> domain
2. The RFC5321 MailFrom / Return Path domain

The Alignment test for SPF is performed in order to verify the authenticity of the domain sending the email by using two header signatures in the message where the sender's domain is present.

Example: SPF NOT in alignment:

```
MAIL FROM: <sender@amazonses.com>  *This is the RFC5321.MailFrom domain.
From: sender@example.com  *This is the RFC5322.From domain.

Date: Fri, Feb 15 2002 16:54:30 -0800

To: receiver@example.org
```

DMARC

https://www.dmarcanalyzer.com/dmarc/dmarc-record/

A DMARC record is the core of a DMARC implementation in which the DMARC record rulesets are defined. This DMARC record informs email receivers if a domain is set up for DMARC. If so, the DMARC record contains the policy which the domain owner wants to use. In essence, a DMARC record a DNS (Domain Name Service) entry. One can start using DMARC by implementing a DMARC DNS record. This DMARC record will be used by email receivers which have adopted DMARC. This will result in keeping track of all the messages which have been sent to your domain taking your DMARC policy into account.

The bottom line is that this will empower the organization publishing the DMARC record to instruct how non-compliance should be handled. The messages can be monitored (and delivered), moved to the junk folder or rejected.

Bad DMARC

In case DMARC record of the sender's domain exists but there is a problem in the DMARC setup, it is signalled by the appendage of [Bad DMARC] in the email subject.

In case you get an email with such warning, it may be a good idea to inform the sender about the issue.

It is usually worse to have badly set DMARC and SPF then not to use DMARC and SPF at all.

We at CERGE-EI cannot do more than to detect such misconfigured and hence not fully trustfull party.

To check DMARC setup for any domain go to
https://www.dmarcanalyzer.com/dmarc/dmarc-record-check/

Reputation databases - Blacklists

IP reputation

Added subject tag: ***[IP reputation]***

More problematic IPs are also taged with ***[!]*** *or* ***[!!]***

Bad IP reputation - emails from IP addresses with bad reputation may be discarded or quarantined. It is usually dangerous to receive emails from such IPs.

IP reputation may be checked here: https://www.ipqualityscore.com/ip-reputation-check/lookup/

It is responsibility of the sender to have 'clean' IP address.

In case there is involved dynamically assigned address from a service provider (like Vodafone, T-mobile, O2, UPC ...) the sender's IP address may be somehow compromised just because it was mis-used by a previous user. This is up to IP address user to ask the respective service provider for removal from the bad reputation lists.

IP reputation database - DNSBL

Added subject tag: ***[IP reputation - DNSBL listed]***

see: https://www.dnsbl.info/

Domain Name System Blacklists, also known as DNSBL's or DNS Blacklists, are spam blocking lists that allow a website administrator to block messages from specific systems that have a history of sending spam.

DNSBL Information provides a single place where you anyone check that blacklist status of the mail server's IP address on more than 100 DNS based blacklists.

IP reputation database - SURBL

Added subject tag: ***[IP reputation - SURBL listed]***

see: http://www.surbl.org/

SURBLs are lists of web sites that have appeared in unsolicited messages. Unlike most lists, SURBLs are not lists of message senders

Newsletter

Mail filter tries to recognize typical newsletter signatures. Newsletters are further categorised by their inner structure clarity. Details follow:

Suspicious Newsletter

Added subject tag: ***none*** *or* ***[Newsletter]***

Suspicious newsletters are part of the newsletter category.

FortiMail may find them to be suspicious because they may actually be spam under the disguise of newsletters.

The internal structure of such email is not clear enough, usually contains obfuscated links or links to problematic third-party sites.

It is usually sent from server which differs from the sender's domain. Sometimes IP address of the sending server is identified as a mass-mail commercial service.

Mail filtering logic can hardly distinguish the regular and malicious content context - semantics understanding is mostly up to indivisual recipients.

Warning Disclaimer (prepended to email)

If you see the warning disclaimer prepended to the email text it means that the Anti-Phishing/Anti-Spam **engine cannot authoritatively decide** whether the URL links in email message leads to the harmless targets, or whether they redirect to the malicious content.

Certain sort of newsletter senders wants to track recipients clicks (to monetize and/or monitor recipient behavior) so they conceal the target URL behind their own hash. It is then undecidable whether the redirected URL is OK or not (phishing).

Example

Obfuscated/unresolvable link:If you get the newsletter from **bostonglobe.com** with links in the form https://bostonglobe.us11.**list-manage.com/track/click?u=90f9e490a86&id=0c98f5&e=e8fef** , it cannot be said what is the targeting URL. Hence the **warning about uncertain content is added**.

Regular/direct link: If the newsletter from **newyorker.com** contains links in the form https://link.**newyorker.com/view/5dc1b3fc91f4/03075c2d** , it may be tracked down to the target URL and **is considered safe**.

Macro in attachments

PDF macro

PDF files include the ability to **execute code on your device** — and that's where the danger lies!

Hence PDF files containing macro / executable code (like filling forms) are preventivelly **placed to users's quarantine** where may be carefully released by user in case content is harmless. User may "whitelist" a trustful sender so quarantine might be skipped next time.

PDF can contain the following:

Javascript – Javascripts are used in the website coding to control browser appearance and functionality. In past, it has been used to exploit multiple vulnerabilities in Adobe as well as many other PDF readers.

System Commands – Launch action in PDF can open Command window and execute commands to initiate malware. Most of the commands have now been disabled by Adobe but they might be open in other readers or earlier versions.

Hidden Objects – PDFs can have embedded and encrypted objects which prevents being analyzed by antivirus scanner. These objects are executed when file is opened by the user.

Multimedia Control – When we say PDF can have embedded objects, it could be a quicktime media or flash file. Attacker can exploit vulnerability in media players.

Suspicious

Suspicious content (HTML links, docs, macro)

Added subject tag: ***[Suspicious]***

HTML contents in email body and attachments may contain potentially hazardous tags and attributes (such as hyperlinks and scripts). MS Office and PDF attachments may contain potentially hazardous macros, active scripts, and other active contents.

FortiMail provides the capability to remove or neutralize the potentially hazardous contents and reconstruct the email messages and attachment files.

Suspicious links (phishing, spam, malware) are redirected to Click Protection. URL is rewritten to `https://gw.cerge-ei.cz/xxxxxxxxx..` (where gw.cerge-ei.cz is the address of our email security gateway) and in case the user clicks on the URL, the link is evaluated by FortiGuard and appropriate action is taken according to risk level (link is blocked or allowed)

Image Spam

Added subject tag: [Image Spam]

Some spammers conceal spam text as an image or PDF document. Mail filter tries to recognize such content and warn users about such situation.

False-positive detection is sometimes possible. That's why such email is not rejected but is just subject-tagged. In more suspicious cases the message may be put to quarantine.

Deepheader analysis

Added subject tag: [Suspicious - header analysis]

Added warning text: “Deepheader analysis examines header for spam characteristics. Don't click any link unless you are certain it's legitimate.”

Tool to examine message header (displays human-readable content):https://mha.azurewebsites.net/

Deepheader analysis examines the entire message header for spam characteristics.

More specifally - the deep header scan examines each message and **calculate a confidence value based on the results of the decision-tree analysis**.The higher the calculated confidence value, the more likely the message is really spam.

There may be sometimes very subtle difference in the specific email, which triggers the confidence value. As a best practice - it is always better to have an information that certain email may be problematic, because attackers today are able to mimick messages that are almost indistinguishable from the original messages! **So check twice such messages before you click link in it or open an attachment**.

Line *X-FEAS-DEEPHEADER:* is added to the message header that includes the message’s calculated confidence value.

Basically an email has two parts. The body (information sent to recipient) and the header containing metadata (like “from”, “to”, content type, date of delivery, message forwarding path, signatures of mailservers, certificates, system-gegerated informations like spam level, processing info etc.).

Recipient can learn a lot about the email history and nature by examining message header.