Armed for the spam battle: a technological and organizational infrastructure framework

Guido Schryen
Institute of Business Information Systems, RWTH Aachen University
schryen@winfor.rwth-aachen.de

Abstract

Spamming remains a form of Internet abuse, which burdens the Internet infrastructure, is generally regarded as an annoyance, and is said to cause economic harm to the tune of about several billion US$ per year. Many technological, organizational, and legislative anti-spam measures have already been proposed and implemented, but have not led to any substantial decrease in the number of spam e-mails. We propose here a new infrastructure framework that combines several anti-spam measures in a framework that features both a technological and an organizational facet. The key element of our infrastructure is a new organizational unit that reliably and transparently limits the number of e-mails that can be sent per day and per account. This paper first gives an overview of the framework, then it provides technological and organizational details of the infrastructure, the deployment of which depends to a large degree on its acceptance and propagation by the ICANN, the ISOC, and by large e-mail service providers. Finally, the paper discusses the limitations and drawbacks of the proposed framework.

1. Introduction

We still face in practice a high volume and a high portion of spam e-mails, although many technological, organizational, and legislative anti-spam measures have already been proposed and implemented. This makes it necessary to continue with research regarding both the development and deployment of effective anti-spam countermeasures. This far, no single measure has proved to be the silver bullet against spam, and it is doubtful whether any single, simple solution will ever be able to reduce or stop spam. Rather it seems appropriate to look for solutions that provide a complementary application of several anti-spam measures. This paper aims at the conceptual development and analysis of an infrastructural e-mail framework. This framework is intended to have the following characteristics, which we assume to be preconditions for an effectiveness in the long run and a widespread adoption by the e-mail community:

- Both technological and organizational modifications must be minor.
- An openness must be present, insofar as the framework provides for principles, and not for concrete algorithms or data formats.
- Spam should be stopped as close to its true source as possible. The prevention of spam has a higher priority than does its detection.
- Means to support the sending of solicited bulk e-mails have to be provided.
- The deployment of the key elements can be done smoothly and flexibly, i.e. the adoption of the infrastructure can occur evolutionarily.
- The infrastructure must not undertake an “arms’ race” with spammers (for example, filters do undertake such a race).

This rest of this paper is structured as follows: Section 2 provides an overview of the framework and of the interaction of its key components. The framework includes both organizational and technological elements, which are discussed in detail in Sections 3 and 4 respectively. Deployment issues and the impact on e-mail communication are covered in Section 5. The paper closes with a consideration of the limitations and drawbacks in Section 6.

2. Overview of the framework

The core ideas of the framework are (1) to limit the number of e-mails that can be sent during a specific time-window and per account[1], (2) to restrict the automatic set-up of e-mail accounts and (3) to provide means for controlling this limitation of e-mail traffic by introducing an element of centralism [2].[2] In order

[1] The implementation of rate limits on outbound e-mail traffic is part of “Best practices for e-mail service providers”, which are proposed by many organizations,such as the Anti-Spam Technical Alliance [1].
[2] In principle, the framework follows the idea that a credit of, for example, 100 messages per day is a very large number for an individual, but an inconsequential number for a spammer.

to support these ideas, a new organizational role is introduced: the Counter Managing & Abuse Authority (CMAA). The framework is intended to include several organizations, each of them taking on the full CMAA role. These organizations are either new and designated ones or established ones, such as trustworthy e-mail service providers (ESPs). In our framework, in principle, a sending organization (SO), for example an ESP, either directly transmits an e-mail to the receiving organization (RO) or sends the e-mails to a CMAA organization, which then relays the message to the SO. The former option is today's default option for sending e-mails, but is intended to be used in our framework only if the RO trusts the SO with regard to the implementation of effective anti-spam measures. Otherwise, the latter option applies, which means that the CMAA first checks whether the sender would exceed the number of e-mails he or she is allowed to send on one day. Depending on the result, the CMAA would then either bounce the e-mail or relay it to the RO. This replacement of today's direct SMTP connection between the SO and the RO by a relaying procedure represents an element of centralism, which allows for controlling and accounting the (volume of) e-mail traffic. This control is intended to strongly reduce the sending of unsolicited bulk e-mail. Solicited bulk e-mail may still be sent, if a person or organization accepts to be taken (legally) responsible for a proper use. The (anti-spam) control is also intended to make additional anti-spam measures by ROs obsolete. As the control mechanism is unlikely to prevent all spamming, it seems reasonable to complementarily provide a forum for e-mail users' complaints about unsolicited e-mails. Therefore, any CMAA organization is intended to also operate a central anti-spam abuse system. The abuse system and the relaying system are connected to each other in that numerous complaints about the spamming activities on behalf of a specific sender may lead to the blocking of the sender's CMAA account and, thus, to the bouncing of further e-mails of this sender. For the rest of this paper, we use the shorter term CMAA for "CMAA organization", unless we explicitly provide the term CMAA to designate the role.

In order to implement the accountability, on which the framework bases, the SO sets up a record for each sender's e-mail account prior to the first relaying. The records are stored in a "Counter Database (CDB)". As any CMAA is also responsible for the locking of accounts due to abuse complaints, these complaints are stored in the "Abuse Database (ADB)". A third database, the "Organization Database (ODB)", serves for the storage of information about those SOs that are registered on the CMAA for the usage of its services.

Fig. 1 illustrates the infrastructure framework. For the purpose of simplification, those infrastructure elements that are responsible for the administration of the databases are omitted. They are presented in Section 4.

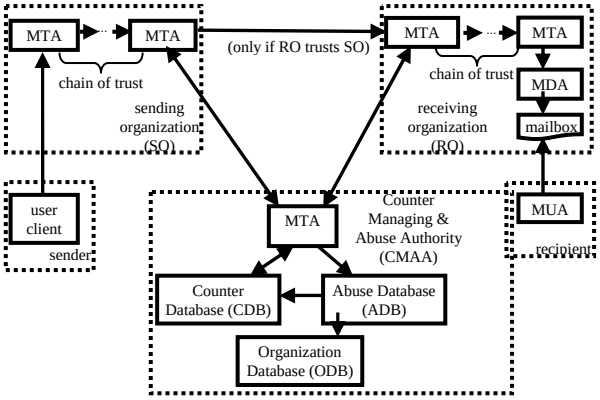

Figure 1. Overview of the infrastructure framework

The introduction of additional organizational units that are responsible for the implementation of the described tasks requires organizational, technological, and financial support and, therefore, seems to be unnecessary and even counterproductive. However, some reasons support its appropriateness:

1. The operation of CMAA (role) services is critical for the success of the framework and requires both the willingness and the technological ability to operate a CMAA properly. Organizations that reside in countries with a non-restrictive anti-spam environment may only improperly fulfill these requirements; organizations that are notorious for addressing spam only lackadaisically are likewise unqualified. A CMAA that is operated and controlled by a trustworthy organization seems to be much more appropriate for providing the required services.
2. The list of trustworthy organizations is CMAA-specific and is maintained by each CMAA. The administration of decentralized whitelists and blacklists by ROs would become obsolete. Each organization receiving an e-mail that has been relayed (and counted) by a CMAA can assume that the SO is a trustworthy one. Therefore, ROs would only be required to maintain data for all CMAAs, such as IP lists of trustworthy MTAs.
3. The infrastructure will not eradicate spam, but should support an abuse system. Currently decentralized abuse systems could be consolidated by integrating this service into the portfolio of the CMAAs.

Although the framework seems to resemble reputation-based approaches, such as LUMOS [3] or the sTDL approach of Spamhaus [4], it differs from them in two main issues: (1) The reputation-based approaches keep the e-mail communication direct. It is the RO that has to prove the reputation or accreditation of a particular SO. In contrast, our framework provides an additional organizational unit, which relays e-mails, and makes the communication indirect. Therefore, with our approach, it is not up to each RO to prove the reputation of a particular SO; this is a CMAA's task. (2) With our approach, the SO's fulfillment of requirements is not sufficient for the successful delivering of a message. In addition, a restriction on the remaining account-specific credit applies.

However, as with reputation-based approaches, it remains an important task to formulate a set of requirements for SOs, which are effective regarding the misuse of a CMAA's services and the fulfillment of which can be verified. Because of this importance, these issues are addresses in Section 3 in detail.

3. Organizational facet

The framework involves technological as well as organizational modifications to the Internet e-mail infrastructure and the e-mail processes. The organizational modifications, which are addressed in this section, result from the introduction of the CMAA as a new organizational role. As mentioned above, the framework is intended to involve several organizations each of them taking on the full CMAA role. However, a few outstanding key questions must be addressed prior to implementation and deployment: 1. Who will operate a CMAA? 2. How is a CMAA certified and by whom? 3. Which CMAA is responsible for which organization? 4. How does an organization register for the usage of CMAA services? These issues are addressed in the following subsections.

3.1. Integrating CMAAs onto the Internet

The introduction and the maintenance of a new organizational role that is as important and central as the CMAA demands a control and a policy that is independent of technological, economic, social, political, and cultural players. Therefore, we propose to entrust an established and well-accepted Internet organization, such as ISOC or ICANN, with the ruling of CMAA issues. In the following, we denote the trustworthy organization as *central organization (CO).*

It is the task of the CO to specify precise requirements for a CMAA, receive submissions, inspect the applications, officially certificate applying organizations, and withdraw CMAA certificates. It is also desirable that the CO provides standardized software for CMAAs and their customer organizations.

In principle, designated CMAA organizations may be set up. However, it seems reasonable to assume that, at least in the beginning, mainly already established network organizations, such as trustworthy ESPs, anti-spam organizations, and universities, will serve as CMAAs, because they already dispose of the technological experience, tools, and staff, all of which is helpful, if not crucial, to running a CMAA. The motivation to gain certification and serve as a CMAA can result from two goals: (1) If the CMAA services offered have to be paid for by the organizations that make use of them, then there may be an economic incentive. Furthermore, it saves the costs of registering for an external CMAA. (2) It may increase the organization's reputation.

The operation of a CMAA represents a new business segment of network operations, and in the long run, it seems unavoidable that organizations would have to pay for CMAA services. On the other hand, they may save money which they would have otherwise spent on anti-spam resources that are not needed anymore, such as filters. The fee for the usage of the CMAA services should be balanced: it should be high enough to attract potential operators, and it should be low enough to be non-cost-prohibiting for customer organizations. Especially organizations in less-developed countries may be affected by this issue.

3.2. Certificating an organization as CMAA

The effectiveness of the framework regarding the reduction of spam e-mails heavily relies on the trustworthiness of the CMAA organizations. Therefore, the requirements on organizations that apply for certification as a CMAA should be stringent. We propose that the CO considers the following evaluation criteria for CMAA applicants:

- An applying organization should have either a good reputation in the Internet community or at least references from such organizations. The reputation could include a high integrity in network-based services, an active involvement in anti-spam activities, and a good reputation with regard to anti-spam blacklists maintained by well-accepted organizations.
- The applicant should be under legislation that allows for prosecution in the case of any tolerating or supporting of spam activities. Any spam-promotive behavior, be it intentional or

negligent, must be triable. Additionally, an applicant may be obliged to pay a deposit, that is forfeited in the case of a strong violation of the requirements on a CMAA. These requirements and any case of strong violation would have to be precisely specified in the contract signed by the CO and a particular CMAA.

- The organization's data in the "whois" database must have been successfully validated.
- The implementation of technological requirements that are mandatory for the operation of a CMAA must be accomplished. These include (1) the protection of services and databases against security vulnerabilities, (2) a system redundancy in order to guarantee the operational availability of CMAA services in the case of system crashes and heavy traffic, and (3) an appropriate load balancing system for a time-efficient use of the redundant systems in order to guarantee an appropriate throughput. We propose that the CO supports applicants with standardized and certified software for the operation of tasks that each CMAA has to perform. The usage of such software could even be regulated by the CO.

The certification process is intended to involve personal contacts between the applying organization and the CO, and the agreement is formally defined by a contract.The list of certified organizations, their contact information, the CMAA policy that has to be signed by each certified organization, and organization-specific information, such as service fees, should be provided by the CO. Complaints about a violation against the CMAA policy should be directed to the CO, which can withdraw CMAA certificates if this is deemed necessary.

3.3. Mapping organizations onto CMAAs

It is the decision of each organization that sends e-mails on behalf of its users whether it should use the services of one or more CMAAs, so that we have an (m : n) relationship between SOs and CMAAs. Usually, an SO would use only one CMAA. The framework is scalable in that it allows sending organizations to bypass any CMAA and to omit the registration on any CMAA. The pressure on these organizations to register is determined by the extent to which the Internet e-mail community accepts the importance of CMAAs, i.e. to which extent the community of ROs makes the decision of whether an e-mail is accepted or rejected dependent on the participation of a CMAA (or trustworthy SO). If the CMAAs' role is widely adopted by the Internet e-mail community, an SO's omission of a registration at a CMAA results in the rejection of messages sent to users of many or even most organizations.

If an organization has decided to register for CMAA services, then the question arises as to which CMAA to choose. The mapping of organizations onto CMAAs can follow different paradigms.

Market paradigm The decision could be left to the particular SO. Then, a market emerges with CMAAs as sellers and SOs as buyers. However, in order to support the wide diffusion and adoption of the CMAAs' integration, the CO should regulate those issues that may otherwise hamper the diffusion of the usage of CMAAs.

Regulation paradigm Mapping is regulated and CMAAs are assigned to SOs. Examples of regulatory approaches can be found at ICANN.

The organizational structure of the framework is simple and illustrated in Fig. 2.

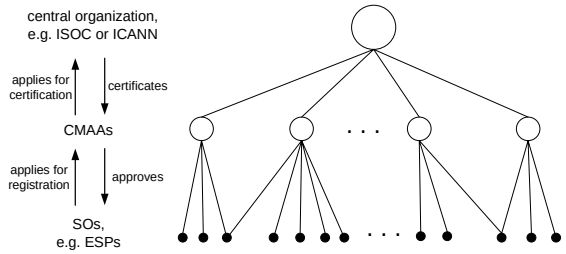

Figure 2. Organizational structure of the infrastructure framework

3.4. Registering for the usage of CMAA services

One of the most critical requirements of the proposed infrastructure is the integrity of registered organizations. Although it seems impossible to exclude all those organizations that tolerate or even support spamming in advance, a set of requirements that applicants have to fulfill may be helpful for the reduction of fraudulent or careless organizations. Similar requirements can be found in [4].

The organization's data in the "whois" database must have been successfully validated. This includes that the administrative contact has signed the application form and proved the identity by attaching a copy of a valid identity card. In case of a repeated misuse of CMAA services, the toleration or even support of spammers this contact may be prosecuted.

Each organization being registered has to sign the anti-spam policy to which it must adhere. In the case of a violation, the organization or its administrative contact can be prosecuted.

The administration client (see Fig. 3) has to be installed. Like the administration server, this software should be provided by the CO.

A public key pair must be generated, and the public key must be added to the DNS. The private key has to be stored securely.

For the purpose of authentication and authorization (when sending an e-mail to a CMAA), LMAP records have to be added to the DNS.

The component that signs messages on behalf of the organization must be protected against any misuse. The CO should provide such software and specify the requirements on the hardware to be used.

It has to be ensured that a reverse DNS query, with any name server of the applying organization as an argument, results in a FQDN whose "SLD.TLD" part is the name under which the organization is registered at its CMAA (see Equation (1)), where SLD is the second-level domain and TLD the top-level domain.

One of the most important requirements on applying organizations is the demand for a manual set-up of accounts. The automatic set-up must be prohibited because, otherwise, the limitation of the number of e-mails per account and day would be pointless. One option would be to initiate an offline registration procedure, which demands a letter-based application, that includes both user identification by signature and the provision of a valid mail address. Another option would be to implement a CAPTCHA procedure [5]. However, CAPTCHA procedures suffer from several drawbacks. We propose that the underlying algorithm has been evaluated by the CO and that the CO provides CAPTCHA software.

In order to protect e-mail accounts from easy misuse, an authentication mechanism has to be applied. If SMTP-based connection is used, then SMTP-AUTH [9] can be used. Web-based e-mailing services are usually implemented with password-based protection.

In contrast to the CMAA certification process, the registration process is not intended to involve a personal contact. The reason for this is that it would be too cumbersome, as the number of registering organizations is much higher than the number of CMAA applicants.

4. Technological facet

This section describes the technological specification of the framework. This specification consists of the description of the three central data stores, the CDB, the ADB, and the ODB, and of the processes that are related to database administration, to e-mail relaying and bouncing, and to the usage of the abuse system. Regarding the following process descriptions, it is not relevant whether the SO is identical with the CMAA or not. In the former case, the roles "SO" and "CMAA" are both realized by the same organization, and although some process simplifications may then be possible, in principle, the processes are even then intended to run as described.

Further, it should be noted here that all of the technologies required to implement this proposal currently exist. The framework leverages existing technologies and services to reduce spam. The overall infrastructure framework is illustrated in Fig. 3.

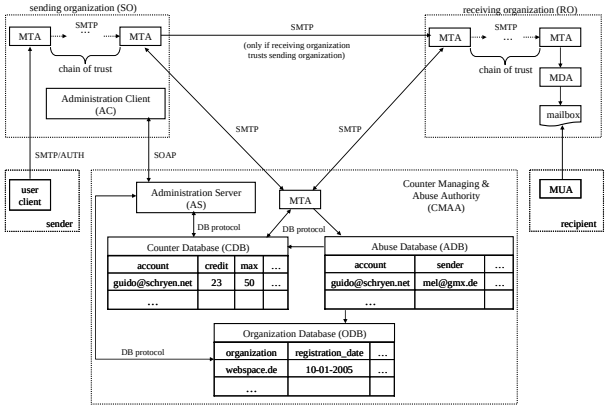

Figure 3. Infrastructure framework

4.1. Databases

Most services offered by a CMAA need to access its CDB. For example, the decision of whether an e-mail is relayed or bounced relies on the data of the particular CDB. We propose any CDB to maintain for every single CMAA-registered e-mail account the following data: *account* is the e-mail address of the database record. E-mails can be sent on its behalf. *credit* contains the current number of e-mails that can be sent on the current day. *max* is the number of e-mails that can be sent per day on behalf of the particular account. *bounce_status* indicates whether a bounce e-mail has already been sent to the account. This would happen when the e-mail limit is first exceeded. Then, *bounce_status* would have to be changed to indicate that no further bounce e-mail is necessary. *setup_org* contains the SLD and the TLD of the organization that set up the record and which offers the e-mail account to the particular user. Only the organization that set up a record is authorized to relay e-mails on behalf of the e-mail address stored in that particular record. *setup_date* gives the date of the set-up procedure and allows for statistical evaluations. *holder* provides the name and the mail address of the holder of the account. This information is mandatory,

if the credit of the account exceeds the default credit, thus offering the option of sending solicited bulk e-mail on behalf of that particular account. If this account is misused for the sending of unsolicited bulk e-mail, then the holder information may be used for prosecution. *idle_days* is the number of days the account has not been used. When certain thresholds are exceeded, the responsible organization – stored in *setup_org* – is informed about the possibly upcoming removal of the account and, finally, about its removal. *blocks* gives the number of times the account has been blocked so far. *status* allows the provision of information about the status of the account. Possible values are "open" and "blocked". The status "blocked" may be reached, when a specific number of complaints have been received.

Regarding the misuse of the abuse system, we propose that each complainant can only submit one abuse complaint per day and account. The ADB would contain the following data: *account* is the e-mail address being complained about. *setup_org* contains the same type of information as the corresponding entry in the CDB. This redundancy serves the purpose of efficiency, when organization-related abuse information is being composed. *sender* provides the e-mail address of the complainant. This information is necessary to ensure compliance with the restriction mentioned above. *date* gives the date of the abuse complaint.

The ODB contains information about the organizations that have successfully registered for the usage of the CMAA services. We propose storing the following information: *organization* contains the same type of information as the corresponding entry in the CDB. *registration_date* gives the date of the registration process. *complaints1,…, complaints30* provide the number of abuse complaints for the last 30 days, whereby 30 is an arbitrary number. *admonishments1,…, admonishments6* allows the storage of the number of admonishments for the last six months. Again, six is an arbitrary number. *status* provides information about whether the organization has been excluded or whether it may still use the CMAA services.

It should be noted that the protection of e-mail addresses that are stored in the databases is very important, because the databases would otherwise provide valuable resources for spammers. Although the usage of hash values or encrypted addresses would seem to be solutions to this problem, they suffer from these drawbacks: If only hash values of addresses are stored, then the addresses cannot be recovered efficiently. However, the addresses are needed for some CMAA administration messages. If the addresses are stored encryptedly, they can be recovered by applying the decrypt function. However, most CDB administration processes and the e-mail delivery process include the sending of an e-mail address that would have to be encrypted or decrypted. Because of the high number of expected queries, the use of cryptographic functions would probably consume too much time. Therefore, the usage of other mechanisms, such as authorization-based ones, should be explored.

This discussion reflects the challenge to many infrastructures and systems in finding an appropriate balance between security, functionality, and (time-related) efficiency.

4.2. Database administration processes

Access to the CDB is granted to SOs that have been approved for the usage of the CMAA-specific CDB and to the CMAA itself. SOs are allowed to set up, modify, and remove records, herein denoted as processes P1, P2, and P3. The CMAA is responsible for the periodical maintenance of the CDB records in many regards. It has to reset values of each record, for example, the credit, by a fixed time of the day (P4), to trace accounts that have not been used for a specific time in order to remove those particular accounts or to inform the responsible SO about the possible upcoming removal (P5), and to block accounts due to spam complaints (P6).

The administration of the ADB and of the ODB is reserved for the CMAA. It is responsible for both the detection of accounts, for which many abuse complaints have arrived, and the detection of SOs that are responsible for such "suspicious" accounts. As a consequence, accounts have to be blocked and SOs have to be admonished or even excluded from all CMAA services (P7). All complaint and admonishment information stored in the ODB has to be updated periodically, because complaints only refer to the last 30 days and admonishments only to the last six months (P8).

As the data that are exchanged between an SO and an CMAA are structured, the usage of e-mails seems to be improper. Rather, the Simple Object Access Protocol (SOAP) [6] provides means for this communication. Any CMAA message is intended to be cryptographically signed by using the "SOAP Security Extensions".

Process P1: setting up a CMAA record P1 is illustrated in Fig. 4, which models the process with a UML 2.0 activity diagram. The process is initiated by a user when he/she wants to set up an e-mail account at an SO. The user usually applies online by using a web form and is intended to have two options regarding

credit: if the user needs more than the default credit, then he or she has to authenticate. This authentication is intended to be submitted offline by mail or fax and must disclose the user's identity and address. For a possible prosecution due to spamming, we propose ensuring that the user underlies an opt-in legislation. If the user applies for an account with default credit, then either the same authentication procedure applies or an effective CAPTCHA procedure has to validate that a human user is applying. If the authentication/validation succeeds, then the SO applies for a CDB record at its CMAA. The CMAA first checks the authenticity. We propose applying a (cryptographic) signature-based procedure for this, because this approach makes it rather difficult, if not practically impossible, to spoof sender data , which would easily lead to the setting up of an arbitrary number of accounts. The SOs' public keys could be stored in the DNS. If the authentication fails for any reason, a rejection message is sent to the SO. Otherwise, the CMAA has to proceed with the authorization of the SO to set up a record for the particular e-mail account. The SO is granted this permission if it is responsible for the e-mail account. This responsibility is defined as follows: either the SLD.TLD domain of the e-mail address is a domain of the SO, for example schryen@winfor.rwth-aachen.de, where rwth-aachen.de is a domain of RWTH Aachen University, or the domain is hosted by the SO, for example, the domain of the e-mail address guido@schryen.de, schryen.de is hosted by the SO. In both cases, each authoritative name server for the given domain belongs to the SO. This verification can be undertaken by using the DNS: let DNSNS(domain) be the operation that requests the DNS for a name server of domain, let RDNS(IP) be the operation that requests the DNS for the host that matches IP, let SLDTLD(address) be the operation that returns the SLD.TLD part of a host or an e-mail address, let setup_org be the SLD.TLD part of the organization that requests the record set-up, and let address be the e-mail address for which a record is requested. Then, the requirement can be verified by using two, possibly cascading, accesses to the DNS:

$$SLDTLD(RDNS(DNSNS(SLDTLD(address)))) = setup\_org \quad (1)$$

If the verification of responsibility succeeds, the CMAA sets up the record and sends a confirmation to the SO, which then sets up the particular e-mail account and sends a confirmation message to the user. If the verification fails, the CMAA sends a rejection to the SO, which then sends a rejection message to the user. The CMAA SOAP server application has to consider that holder data must be provided if the value of max is higher than the default value, which still has to be defined.

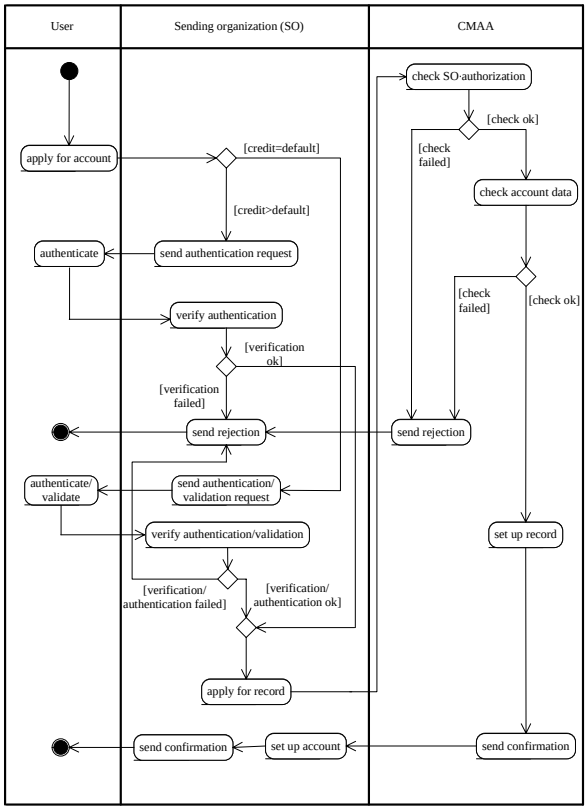

Figure 4. Activity diagram modeling the set-up of a CDB record

Process P2: modifying a CMAA record SO is allowed to modify the max value and/or the holder value only. If the max value is reset to the default value, no holder data must be given. Otherwise the provision of holder data is mandatory. When an SO sends a modification request to the CMAA, the CMAA proceeds analog to its operations in P1.

Process P3: removing a CMAA record The deletion of a CMAA record only requires that the SO provide the account name. Regarding the SO's SOAP message, the notes on P2 apply.

Processes P4 and P5: resetting the credits of CMAA records and tracing for idle CMAA accounts By a fixed time of the day, the CMAA would have to reset the values of each record. The tracing idle CMAA accounts can be shared with this procedure.

Process P6 and P7: blocking CMAA accounts or/and SOs The CMAA should daily consolidate abuse complaints. This consolidation may lead to the

[3] Note that for a successful authorization, each requesting organization is responsible for the provision of adequate DNS entries.

blocking or removal of user accounts. Furthermore, if too many complaints refer to different accounts of one specific SO, then the SO has to be admonished or even excluded from all CMAA services.

When the number of abuse complaints on a specific account exceeds the daily limit, the account is blocked for one day. Each account may be blocked a number of times, which are still to be specified. If the total number of blocking exceeds this value, then the account is removed and the responsible SO is informed about this deletion.

It may happen that SOs ignore, tolerate or even support the abuse of e-mail accounts. Therefore, for each organization, all complaints about those accounts that the organization is responsible for are counted and stored in the ODB, which contains for each SO the number of abuse complaints for each of the last six months. We differentiate between three abuse states that an organization can be assigned: normal, medium, and strong. The status results from the application of the CMAA's policy on the SO's six-month complaint history and the SO's number of past admonishments. The following actions have to be taken by the CMAA, depending on the SO's status: If the history is assessed as "normal", nothing has to be done. If the value is "medium", then the CMAA sends an admonishment to the SO and records this. In the case of a "strong" violation, that particular SO would have to be excluded from all CMAA services. The status would be set to "excluded", all accounts that had been set up by the SO would be removed, and the SO and all other CMAAs would be informed about this exclusion.

Process P8: removing complaint and admonishment information Complaints older than 30 days and admonishments older than 6 months are intended to be removed from the ODB. The removal of complaints has to be executed once a day, the deletion of admonishment information once a month.

4.3. E-mail delivery process

The process of sending an e-mail has to be extended by the integration of a CMAA. Although a CMAA's involvement makes the delivery process more complicated, the modifications are intended to be hidden from the user, who may continue using his/her e-mail client software without any changes. Fig. 5 shows the process by using an activity diagram. The process can be divided into the following components:

User authentication: First, the user has to authenticate, so that his/her account is protected from misuse by an unauthorized person. We propose using the IETF standard SMTP-AUTH [7] with a (user,

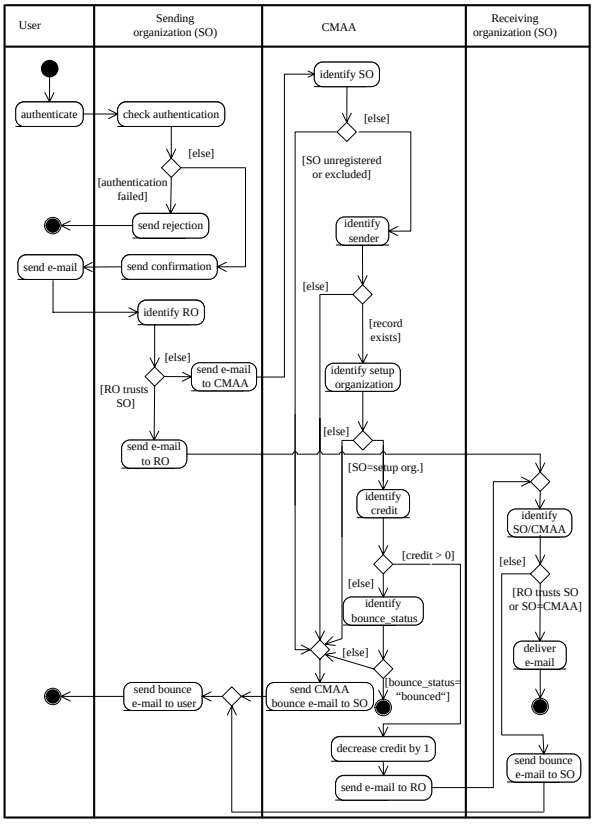

Figure 5. Activity diagram modeling the e-mailing process

password) SASL authentication mechanism [8]. However, for effective protection from misuse, the password must be strong and protected from being read by malicious software. If the authentication fails, the process is terminated, otherwise the user can send the e-mail to his/her SO.

SO's relaying decision: For each recipient of the e-mail, the SO looks for the RO in the internal database that stores the names of those organizations that accept direct e-mail communication with the own organization. If the RO is listed, the e-mail is sent directly to the RO, otherwise it is sent to the SO's CMAA. The case where the SO is identical to the RO is covered implicitly. In such a case, the involvement of a CMAA is not intended. However, it should still be an option for an SO to let a CMAA count those e-mails that are not directed to another SO, in order to protect their users' accounts from being spammed. For the sake of simplicity, this option is omitted in Fig. 5.

CMAA's relaying decision: The CMAA checks whether the SO is registered and not excluded – the SO data can be obtained from the e-mail's FQDN, which has to be successfully validated against the IP of

the sending host by using an LMAP-based procedure – , if the CMAA maintains a record for the sender and if the SO is allowed to send e-mails on behalf of the sender account. If one of these conditions is not fulfilled, the CMAA refuses the relaying and sends a bounce e-mail to the SO, which, then sends a bounce e-mail to the sender. If all tests succeed, the CMAA checks the sender's credit. If no credit is available, the relaying is refused and, provided that no bounce e-mail due to the unavailable credit has been sent, a bounce e-mail is sent to the SO. It is important to send, at most, one bounce e-mail per day and account due to credit unavailability, because it would be otherwise possible to maliciously initiate the sending of an arbitrary number of bounce e-mails to a compromised account: once a password is read or guessed, an attacker could easily send e-mails on behalf of this account thereby causing the CMAA to send a bounce e-mail for each e-mail that exceeds the account's e-mail limit. If the credit is larger than 0, then the credit is decreased by 1 and the e-mail is relayed to the SO.

RO's acceptance decision: When an organization receives an e-mail, it first operates an LMAP-based validation as described above. If the validation fails, the process terminates. Otherwise, the RO checks whether the SO is whitelisted regarding a direct e- mail communication or if the delivering host belongs to a CMAA. If this check is successful, the e-mail is accepted and delivered to the e-mail's recipient. Otherwise, the e-mail acceptance is refused and a bounce e-mail is sent to the SO.

If a CMAA participates in e-mail delivery, its MTA(s) add Received entries to the header as described in [9]. No further modification is necessary.

4.4. Abuse complaint process

The success of the abuse system depends on the user participation in the sending of abuse e-mails to the CMAAs. In order to make a user send a spam complaint to a CMAA, he/she has to know to which CMAA the complaint has to be directed. We envisage at least two options for providing this information: either the CMAA that relays a message adds a new header entry to the e-mail or it adds this information to the e-mail's body as part of a CMAA signature. The first option would be preferable for keeping an e-mail text free from any CMAA (meta) information and for easing the implementation. The reason is that the header entry could be added at the beginning of the message without seeking the appropriate position in the body, which could contain several MIME parts thereby complicating the e-mail's structure. The second option allows the recipient to easily identify the abuse address without having to make the header entries visible. In addition, many users are likely to know little or nothing about the header.

When a user wants to complain about a received e-mail, then the user would have to send an abuse e-mail to the responsible CMAA via his or her organization. The CMAA that receives the complaint e-mail would have to perform three checks: (1) Is the SO registered and not been excluded? (2) Does the CMAA maintain a record for the account being complained about? (3) Does the ADB already contain a complaint tuple (account,sender,date)? The purpose of the third check is to prevent the abuse system from being misused by users sending multiple complaints about the same account in order to discredit it. Only if all checks are positive, is a new complaint record added to the ADB. The setup org data can be obtained by requesting the CDB. As this process is very simple, a graphical representation is omitted here.

The content and format of a complaint e-mail is not specified here, in order to avoid an overstandardization; however, an abuse e-mail must contain the account and the date of the e-mail being complained about. The content and format may vary between different CMAAs, although for the purpose of consistency, it is useful to standardize these issues.

5. Deployment and impact on e-mail communication

A precondition for any deployment of the proposed framework seems to be its adoption by the ISOC, ICANN, and/or large ESPs. This adoption includes both the maintenance of a CO and the propagation of the framework in the Internet e-mail community.

The framework is designed to use both a direct e-mail communication and an indirect one by integrating CMAAs. This flexibility means a scalability of the framework that allows the avoidance of a "big bang" at its introduction, but leaves the (time) schedule for using using CMAAs and its grade up to each organization. An ESP, for example, can decide not to use CMAAs at all, to use CMAAs for incoming e-mails, to register for a CMAA's services, or even to apply for a CMAA certificate. Although no organization is forced to participate in the centralized services, market pressure – assuming that the infra-structure has been widely adopted – will push them to do so, as they are otherwise in danger of being excluded from large parts of the world-wide e-mail communication. This consequence would make the ESP unattractive or even unacceptable from the users' view.

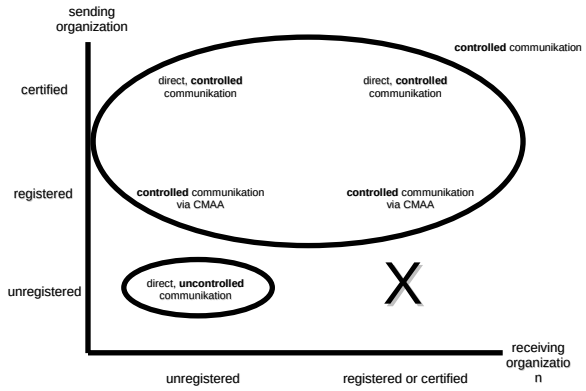

Figure 6. Partitioning of the Internet e-mail communication

If we categorize communication scenarios according to the SO and RO types, we get those categories illustrated in Fig. 6. Organizations that are certified or registered are not limited in their e-mail communication. Other organizations would not be allowed to send e-mails to registered certified organizations, which would usually insist on the registration or certification of the SO. This means that the overall e-mail communication becomes limited. The area of limitation is indicated by the "X". The grade of limitation will depend on the extent to which the CMAAs will be accepted and used. If the proposed infrastructure is either widely accepted or hardly accepted, then the limitation is low, because most e-mail communication belongs to one of the categories, which are displayed as ellipses. A high limitation, i.e. "X" indicates a large subset of e-mail communication, would result from a balanced distribution.

6. Drawbacks and limitations

The implementation of the framework requires both organizational and technological modifications of today's Internet e-mail infrastructure. These modifications have to be propagated by Internet organizations and providers in order to become widely accepted. However, even then, the framework has some drawbacks and limitations.

First, some option for spamming still remain, even if they consume more resources than today. For example, e-mail accounts can be set up manually at registered organizations and then used for spamming, accounts of legitimate users can be compromised by malicious software, organizations that have successfully registered for CMAA services may be corrupt or may tolerate spammers, and an SO that is stored on an RO's whitelist can bypass any CMAA and send an unlimited number of (spam) e-mails.

Second, the approach requires a critical mass of organizations to drive the framework's adoption.

Third, the DNS becomes an even more critical and important resource than it is today for the following reasons: (1) The DNS has to provide entries for public keys of registering organizations. Ideally, the public keys are signed by a trustful organization. (2) LMAP records of SOs and CMAAs have to be added to the DNS. Currently, no single approach has been adopted as a world-wide standard. (3) DNS spoofing would have an impact on the sending of spam e-mails: a CMAA's decision to relay an e-mail depends on the LMAP record. If this entry is spoofed, then a third party could send e-mails on behalf of another registered organization. (4) The availability of DNS servers is closely related to the availability and functionality of the Internet e-mail infrastructure. This attracts attacks on the availability of these servers, such as Distributed Denial of Service (DDoS) attacks.

Fourth, the CMAAs' systems represent a critical resource: (1) The availability of the relays and administration servers is critical with regard to the operational maintenance of large parts of Internet e-mail traffic. Therefore, the consequences of a successful DDoS attack are tremendous. (2) The servers have to handle a huge amount of traffic and requests. This requires a careful implementation of load balancing systems, if e-mail communication is not to become (timely) inefficient. (3) The CMAAs' CDBs contain large numbers of valid e-mail addresses and have to be protected from unauthorized access.

7. References

[1] Anti-Spam Technical Alliance (ASTA), Technology and policy report", Technical report, http://docs.yahoo.com/docs/pr/pdf/asta_soi.pdf, 2004.

[2] G. Schryen, "A Scalable and Flexible Infrastructure Framework for Addressing Spam", Proceedings of IPSI International Conference on Advances in the Internet, Processing, Systems, and Interdisciplinary Research, 2004.

[3] Email Service Provider Coalition, "Project Lumos", Technical report, 2003.

[4] ICANN, "new sTLD RFP Application .mail, Part B. Application Form", Technical report, 2004.

[5] L. von Ahn, M. Blum. and J. Langford, "Telling Humans and Computers Apart Automatically", Communications of the ACM 47(2), pp. 57-60, 2004.

[6] W3C, "SOAP Version 1.2 Part 1: Messaging Framework", http://www.w3.org/TR/soap12-part1/, 2003.

[7] J. Myers, "SMTP Service Extension for Authentication", RFC 2554, IETF Network Working Group., 1999.

[8] J. Myers, "Simple Authentication and Security Layer (SASL)", RFC 2222, IETF Network Working Group, 1997.

[9] J. Klensin, "Simple Mail Transfer Protocol", RFC 2821, IETF Network Working Group, 2001.