University of ???

PhD Thesis

Possible solutions to implement email transfer offering anonymity towards third parties

Author:
Martin Gwerder

Supervisor:
unknown

January 5, 2014

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This document describes a solution, which should offer anonymity against third parties when sending emails based on SMTP and the respective client protocols (e.g. IMAPv4 or POP3). This seemed to bother very few peoples up until information in Echelon became public due to an investigation by a commitee of the european parlament in 2001[4]. Things settled again with local peaks up until a whistle blower named Edward Snowden disclosed 200000 documents proofing activities of the NSA and other secret services. This led to the "2013 mass surveillance disclosures" and damaged the reputation of the american nation in many countries[29].

1.1 Overview over the current situation

SMTP as defined in RFC5321[12] is as of today (2013) state of the art transmission protocol for electronic mail. It is standardized in its current version since 2008 and is one of the few protocols, which is marked as "Standard". While the protocol delivers reliable mail transfer between two endpoint (mail servers) the anonymity of the message content towards any mail server is not given (For a detailed analysis see **??**).

Anonymity against third party is not given due to the following facts.

- There is not always an encryption available between a mail user agent (MUA) and the outgoing mail server.
- There is no way to guarantee that a mail transfer between two SMTP hosts is encrypted.
- There is no always an encryption available between a SMTP host and the MUA of the recipient.

- Encryption based on top level protocols (such as S/MIME or PGP) do hide the message content. The sender, recipient, the subject and some technical information (eg. MIME-Headers) are always in plain available and not protected as such.
- Even if there is a reliable encryption between all endpoints and none of the intermediate servers are compromised sender and recipients might still be identified thru traffic analysis.

Keeping the message content confidential is more and more relevant in these days. The more the importance of mail transfer in today's economy is growing the more is confidentiality and reliability a topic. Unfortunately Secret Services have already discovered the significance of today's mail traffic and start to analyse those. With the presence of Secret Services in the internet, actively investigating data the importance of a reliable data channel for today's messages has become increasingly important.

Quick wins such as the use of "Onion Router Networks" (such as TOR) do not offer any additional security since the message content would be revealed in full to an eventual exit node and any mail server on its way to the recipient.

1.2 Problem statement

This work is an approach to extend the existing mail routing based on SMTP by an intermediate layer, which should offer anonymity against third party.

This work delivers the following results:

- A throughout analysis of current technology and its weaknesses.
  Although the Simple Mail Transfer Protocol (SMTP) is a well-implemented and well proven technology its weaknesses are well known. The SMTP protocol was originally defined in RFC821[23] by Johnathan B. Postel. At this time internet was only available to universities, some mayor companies and governments. The objective of Simple Mail Transfer Protocol (SMTP) is to transfer mail reliably and efficiently[23, p. 1]. Confidentiality or having a tamper proof protocol was no design goal. Over the years many standards arose trying to close some of the gaps. Some of them are being used but most of them are not very common.
- An analysis of possible approaches to improve the current standards.
  Many standards and technologies do exist these days addressing parts of the issues mentioned above. A throughout research should be carried out to identify how can these technologies be combined to achieve the subsequent goals. Furthermore technology advanced. Namely in the field

of cryptology ew possibilities and ideas arose (such as new encryption classes [eg. elliptic curves] or the idea of crypto puzzles). Another field of research which emerged in the analysis of traffic flow is handled under the term "Big Data" where not single events but the sum of events is handled.

- A RFC document describing an approach offering a significant quality improvement of the existing solutions, which could be accepted by the internet community.

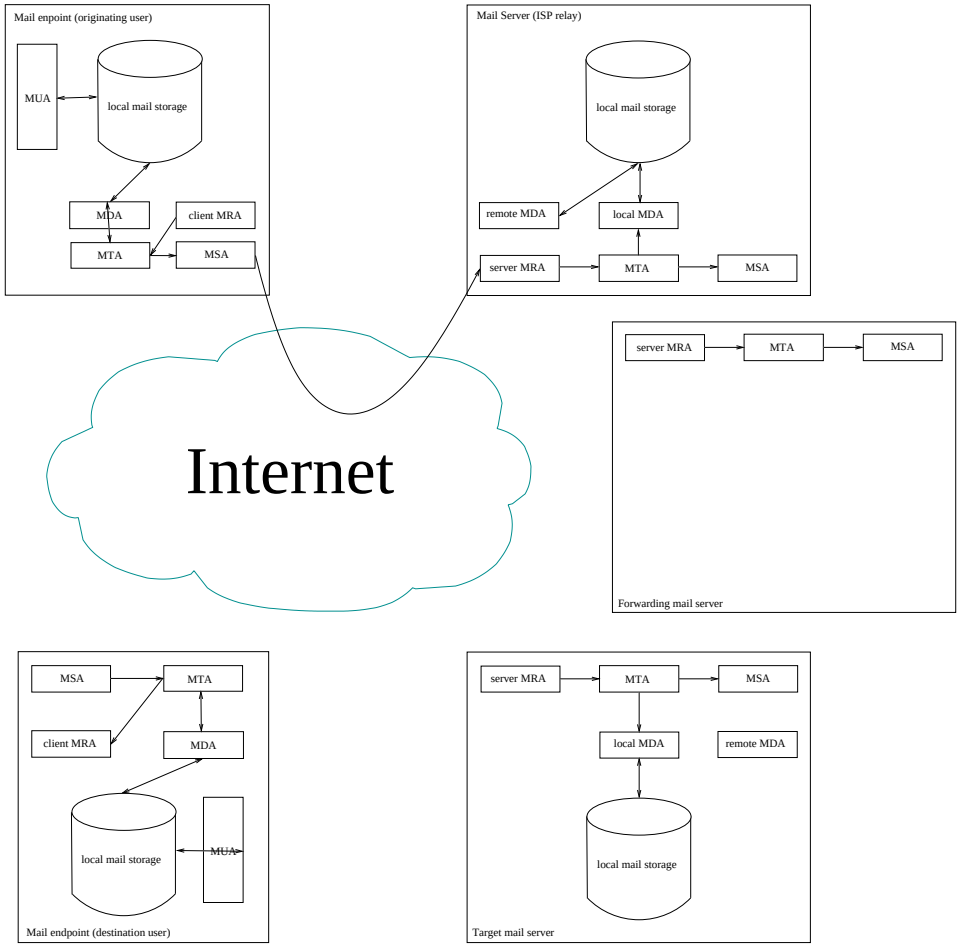
- A prototype reflecting at least the minimum baseline of the RFC document to reflect prove its functionality.
  A prototype should be offered to show the feasibility. The Prototype should be a reference implementation and offer a quick way to use the new technology. It should be distributed under the LGPL license to simplify distribution of the technology.

Chapter 2

Ground theory

2.1 Mail Transport

Todays mail transport is mostly done via SMTP protocol as specified in [12]. This protocol has proven to be stable and reliable. Most of the messages are passed from a MUA to a SMTP relay of a provider. From there the message is directly sent to the SMTP server of the recipient and from there to a server based storage of the recipient. The recipient may at any time connect to his server based storage and may optionally relocate the message to a client based (local) storage. The delivery from the server storage to the MUA of the recipient may happen by message polling or by message push (where as the later is usually implemented by a push-pull mechanism).


Mail enpoint (originating user)
MUA
local mail storage
MDA
client MRA
MTA
MSA
Mail Server (ISP relay)
local mail storage
remote MDA
local MDA
server MRA
MTA
MSA
Internet
server MRA
MTA
MSA
Forwarding mail server
MSA
MTA
client MRA
MDA
local mail storage
MUA
Mail endpoint (destination user)
server MRA
MTA
MSA
local MDA
remote MDA
local mail storage
Target mail server

2.2 Anonymity

2.2.1 k-anonymity

2.2.2 Plausible deniability

2.2.2.1 Deniable encryption

2.3 Identification and data signage

2.4 Encryption

2.4.1 Key exchange

2.4.1.1 Diffie-Hellmann key exchange

2.4.2 Symetric encryption

2.4.2.1 Advanced Encryption Standard

2.4.3 Asymetric encryption

2.4.3.1 RSA

2.4.3.2 El-Gamal

2.4.3.3 ECDSA

2.5 Mix cascades

2.6 Remailers

Agents which do accept Mails from one party and forward it to another party while modifying its content well known under the name of "Remailers". Wikipedia [28] lists four types of Remailers.

Pseudonymous Remailers (or Type-0-Remailers) are remailers that establish a pseudonymity. This means that the senders Email-Address is removed and replaced by a pseudonymous E-Mailadress under the remailers control. This sender address may be used as an ordinary email-Adress to reeach the original sender of the mail. These types of Remailers allow to send mails while one or both recipients do not know their counterpart. The message (or at least parts

of it) might be encrypted but do not have to be. For someone controling the Remailer it will always be possible to make a link between the pseudonymous mail address and a original mailadress. So pseudonymity is only granted towards people not controlling the remailer. Furthermore a person or organisation might be able to discover the Information tuple of Sender and pseudonymous email by analyzing messages and their timely context. So this remailer system is suspectible for traffic analysis.

Cypherpunk-Remailers (or Type-1-Remailers) do function a bit different. They take an encrypted message which was encrypted using the public key of the server, decrypt it and send it to a recipient. The original senders identity gets lost. A reply to a cypherpunk message is not possible. Messages sent to a cypherpunk server might contain messages to other cypherpunk remailers. This daisy-chaining of cypherpunk-nodes allows hiding the original sender-receiver-tuple from a single node. The first node knows only the the originating sender while the last node knows only the final recipient. All intermediate notes do only know the nodes they were linking. However if having traffic information of the entry and exit nodes the tuple might be discovered by traffic analysis.

Mixmaster remailer (or type-2-remailer) is a serie of mailers which split up a message into equally sized chunks and forward them using different paths (via SMTP) to an exit node where the message is reassembled and sent to the final recipient. However if having traffic information of the entry and exit nodes the tuple might be discovered by traffic analysis.

Mixminion remailer (or type-3-remailer) is an enhanced development of Mixmaster remailer. It is currently no longer under active development. It adresses serveral weaknesses of the mixmaster. Namely replies are possible. Forward anonymity is now given. Replay prevention and key rotation is part of the design and there are exit policies allowing ISPs to opt out from receiving remailer traffic. It is based on a proprietary communication network. It furthermore introduces dummy traffic to reduce traceability. This is the most complete approach email anonymity ever given. The aproach has however its weaknesses. To avoid partitioning attacks Miximinon distributes its network information with central redundant directory servers.

2.7 Ethics

2.7.1 Human rights

2.7.1.1 Freedom of speech

Article 19 of the ICCPR states that "everyone shall have the right to hold opinions without interference" and "everyone shall have the right to freedom of

expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice".

2.7.2 Ethics of the Internet

There is an RFC document regarding "Ethics and the Internet"[26, p. 1]. Document states as unethical behaveour:

- An activity that seeks to gain unauthorized access to the resources of the Internet.
- An activity that disrupts the intended use of the Internet.
- An activity that wastes resources (people, capacity, computer) through such actions.
- An activity that destroys the integrity of computer-based information.
- An activity that compromises the privacy of users.

Unfortunately these actions do exist in modern internet and the most powerful players discovered so far are governmental agencies. Using a mixer and cryptographic algorithms definitely wastes resources. But it must be considered the right of every single user of the internet to uphold these points. As a final conclusion the proposed system does not violate the ethics of the internet but it must be designed to be as economically as possible with the existing resources.

2.8 Possible legal issues

One of the first questions I have been asked when working for this topic was: Is this legal? The question is important but not easy at all. The mail system is a global spanning network coming across almost any country of the world. Some of these countries consider almost any kind of secret as illegal as long as the country itself is not able to capture it. Some countries consider it as perfectly legal and some will generally accept its presence as long as the country or establishment is not endangered due to its usage. It is not part of this work to

My personal unscientific point is: I do not care. In my country it is definitely legal as long as I am a well behaved citizen (as long as I do not missuse this system to plan or do illegal actions). There are already proprietary systems available which offer the same functionality. All I do is adding this functionality to the common system instead of reinventing the wheel. There are however many very good reasons to have such a system. Correspondence about my

health, my business relations, my friends or my family (to give just a couple of examples) should be kept private even in an open world. The misuse of information would cause tremendous damage and several events in time (which have been mentioned earlier) showed that there are many secret services and other players using any kind of information to achieve their own goals or the goals of associates. They do this regardless of any country borders or regulations. Since I have no means of controlling the flow of messages in the internet or the hubs where a mail is running thru I consider it as fair to generate an addon to compensate the lack of control in the existing system. Exactly as a car – the system may be legal or illegal and it depends on the users whether he wants to use it or not and in what way.

Chapter 3

Current situation

As of today the de facto standard for asynchronous mail transfer is SMTP as defined in RFC5321[12] and its predecessors. While the transfer protocol SMTP is quite compact, the protocol is enhanced with several standards for encryption, multimedia support and similar. A mail client offers today various support for a lot of sub-protocols. The following list is an excerpt of related sub-protocols which are either related to transport, reliability, identification or encryption.

3.1 Implemented protocols

3.1.1 SMTP

[12]

3.1.1.1 Mail transport

[13]

3.1.1.2 encryption

3.1.2 MIME

[6] [7] [8] [9] [10]

3.1.3 S/MIME

[25]

3.1.4 PGP

[24]

3.1.5 Sender Policy Framework

[31] [11]

3.1.6 Sender ID

[30]

3.1.7 DNS

[3]

3.1.7.1 DNSSEC

[14]

3.1.8 Transport Protocols

3.1.8.1 IPv4

[22] [19] [27] [20] [17] [15] [16] [21, p. 3]

3.1.8.2 IPv6

[2]

3.1.8.3 TCP

3.1.9 POP3

[18]

3.1.10 IMAPv4

[1]

Chapter 4

Analysis of current situation

4.1 Current state of common Technology

4.1.1 Mailrouting

4.1.1.1 SMTP

4.1.1.2 LMTP

4.1.1.3 IMAP

4.1.1.4 POP3

4.1.1.5 MS-OXMAPIHTTP

4.2 Current state of available Technology

4.3 Missing Gap

4.4 Skeleton of Mails and mail transfer

Chapter 5

Designing an approach

5.1 Defining system boundaries

5.1.1 Thread model

As an adverser we assume the following attributes:

- Available founding is huge.
- can monitor all network traffic.
- Can have own mixer infrastructure.
- Is able to read, write or modify network data freely at any point of the net.

His intensions are:

- Discover message flows
- Discover message contents
- Identify users of the system

5.1.2 User model

The assumed user of the system is:

- Does care about privacy.
- Has no special computer knowhow.

- Has the ability to install a program or plugin.
- Has no cryptographic knowhow.
- Is using a device with enough calculation power to solve cryptographic tasks.

His intensions are:

- Send personal or confidential Information securely to another user

His expectations are:

- System should be easy to configure and maintain (in an ideal world: Zero touch).
- System should be fast.
- System should be reliable.
- System should work on any client he is using.
- System should not be a legal problem to him or any of his peers.

5.1.3 Mail server admin model

The assumed mail server admin of the system is:

- Does care about privacy.
- Has cosiderable computer knowhow.
- Has the ability to install a program or plugin.
- Has possibly no cryptographic knowhow.
- Does know his own mail infrastructure
- Is using a device with enough calculation power to solve cryptographic tasks.

His intensions are:

- Support his users in sending personal or confidential information securely to another user

His expectations are:

- System should be easy to configure and maintain (in an ideal world: Zero touch).
- System should be fast.
- System should be reliable.
- System should work on any client he is using.
- System should not be a legal problem to him or his company.

5.2 Basic Requirements of an approach

Chapter 6

Specifying a target solution

Chapter 7

Verification of solution

7.1 User acceptance of the target system

From a perspective of a user Collected requirements to a mail system:

Requirement	cliteria	Weight
The System should transport mails fast under normal conditions	Mails should travel with at least 1MB/min	5
The System should transport mails reliable	Mails should always arrive or their status should be retrievable	9
The System should offer anonymity against spying from third parties	Neither original sender nor final destination or any part of the message content should be determinable by any part of the system except for the original sender and the final recipient.	9
The system must be easy to handle		8
The system must be easy to install	Installation should be almost a “single-click”-Thing. Details should be copied or accessed from the existing configurations.	5
continued on next page		

continued from previous page		
Requirement	cliteria	Weight

Table 7.1: User acceptance requirements

7.2 Admin acceptance of the target system

Collected requirements to a mail system from an admin perspective:

Requirement	Criteria	weight
The System should transport mails fast under normal conditions	Mails should travel with at least 10MB/min	5
The System should transport mails reliable	Mails should always arrive or their status should be retrievable	9

7.3 Possible attacks to the system

7.3.1 Generic DoS attacks

7.3.1.1 Overloading single nodes

7.3.2 Attacks on the users anonymity

7.3.3 Reputaional attacks

7.3.3.1 Misuse for sending spam

Definitions

MTA A Mail Transfer Agent. This transfer agent rroutes mails between other components. Typically an MTA receives mails from an MRA and forwardes them to a MDA or MSA. The main task of a MTA is to provide reliable queues and solid track of all mails as long as they are not forwarded to another MTA or local storage.

MRA A Mail receiving Agent. This agent receives mails from a agent. Depending on the used protocol two subtypes of MRAs are available. Client MRAA client MRA picks up mails in the server mail storage

MSA A local Mail Sending Agent. This agent accepts mails to be sent to a remote MTA.

MUA A Mail User Agent. This user agent reads mails from a local storage and allows a user to read existing mails, create and modify mails.

Privacy From the Oxford English Dictionary: "

1. The state or condition of beeing withdrawn from the society of others, or from the public intrest; seclusion. The state or condition of beeing alone, undisturbed, or free from public attention, as a matter of choice or right; freedom from interference or intrusion.

2. Private or retired place; private apartments; places of retreat.

3. Absence or avoidance of publicity or display; a condition approaching to secrecy or concealment. Keeping of a secret.

4. A private matter, a secret; private or personal matters or relations; The private parts.

5. Intimacy, confidential relations.

6. The state of being privy to some act.

"[5, FIXME]

In this work privacy is related to definition two. Mails should be able to be handled as a virtual private place where no one knows who is talking to whom and about what or how frequent (except for directly involved people).

Bibliography

[1] M. Crispin. *RFC3501 INTERNET MESSAGE ACCESS PROTOCOL - VERSION 4rev1*. IETF, 2003. URL: http://tools.ietf.org/pdf/rfc3501.pdf (cit. on p. 12).

[2] S. Deering and R. Hinden. *RFC2460 Internet Protocol, Version 6 (IPv6) Specification*. IETF, 1983. URL: http://tools.ietf.org/pdf/rfc2460.pdf (cit. on p. 12).

[3] D. Eastlake, E. Brunner-Williams, and B. Manning. *BCP42 Domain Name System (DNS) IANA Considerations*. IETF, 2000. URL: http://tools.ietf.org/pdf/rfc2929.pdf (cit. on p. 12).

[4] Temporary Committee on the ECHELON Interception System. *REPORT on the existence of a global system for the interception of private and commercial communications (ECHELON interception system)*. 2001. URL: http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+REPORT+A5-2001-0264+0+DOC+PDF+V0//EN&language=EN (cit. on p. 1).

[5] FIXME. *Oxford English Dictionary* (cit. on p. 23).

[6] N. Freed and N. Borenstein. *RFC2045 Multipurpose Internet Mail Extensions; (MIME) Part One: Format of Internet Message Bodies*. IETF, 1996. URL: http://tools.ietf.org/pdf/rfc2045.pdf (cit. on p. 11).

[7] N. Freed and N. Borenstein. *RFC2046 Multipurpose Internet Mail Extensions; (MIME) Part Two: Media Types*. IETF, 1996. URL: http://tools.ietf.org/pdf/rfc2046.pdf (cit. on p. 11).

[8] N. Freed and N. Borenstein. *RFC2047 Multipurpose Internet Mail Extensions; (MIME) Part Three: Message Header Extensions for Non-ASCII Text*. IETF, 1996. URL: http://tools.ietf.org/pdf/rfc2046.pdf (cit. on p. 11).

[9] N. Freed, J. Klensin, and J. Postel. *RFC2048 Multipurpose Internet Mail Extensions; (MIME) Part Four: Registration Procedures*. IETF, 1996. URL: http://tools.ietf.org/pdf/rfc2048.pdf (cit. on p. 11).

[10] N. Freed, J. Klensin, and J. Postel. *RFC2049 Multipurpose Internet Mail Extensions; (MIME) Part Five: Conformance Criteria and Examples*. IETF, 1996. URL: http://tools.ietf.org/pdf/rfc2049.pdf (cit. on p. 11).

[11] S. Kitterman. *RFC6652 Sender Policy Framework (SPF) Authentication Failure Reporting Using the Abuse Reporting Format*. IETF, 2012. URL: http://tools.ietf.org/pdf/rfc6652.pdf (cit. on p. 12).

[12] J. Klensin. *RFC5321 Simple Mail Transfer Protocol*. IETF, 2008. URL: http://tools.ietf.org/pdf/rfc5321.pdf (cit. on pp. 1, 5, 11).

[13] J. Klensin, N. Freed, and K. Moore. *RFC1870 SMTP Service Extension for Message Size Declaration*. IETF, 1995. URL: http://tools.ietf.org/pdf/rfc1870.pdf (cit. on p. 11).

[14] B. Laurie, G. Sisson, R. Arends, and D. Blacka. *RFC5155 DNS Security (DNSSEC) Hashed Authenticated Denial of Existence*. IETF, 2008. URL: http://tools.ietf.org/pdf/rfc5155.pdf (cit. on p. 12).

[15] J. Mogul. *RFC922 BROADCASTING INTERNET DATAGRAMS IN THE PRESENCE OF SUBNETS*. IETF, 1984. URL: http://tools.ietf.org/pdf/rfc922.pdf (cit. on p. 12).

[16] J. Mogul and J. Postel. *RFC950 Internet Standard Subnetting Procedure*. IETF, 1985. URL: http://tools.ietf.org/pdf/rfc950.pdf (cit. on p. 12).

[17] Jeffrey Mogul. *RFC919 BROADCASTING INTERNET DATAGRAMS*. IETF, 1984. URL: http://tools.ietf.org/pdf/rfc919.pdf (cit. on p. 12).

[18] J. Myers and M. Rose. *RFC1939 Post Office Protocol - Version 3*. IETF, 1996. URL: http://tools.ietf.org/pdf/rfc1939.pdf (cit. on p. 12).

[19] J. Postel. *RFC791 INTERNET PROTOCOL DARPA INTERNET PROGRAM PROTOCOL SPECIFICATION*. IETF, 1981. URL: http://tools.ietf.org/pdf/rfc791.pdf (cit. on p. 12).

[20] J. Postel. *RFC792 INTERNET CONTROL MESSAGE PROTOCOL DARPA INTERNET PROGRAM PROTOCOL SPECIFICATION*. IETF, 1981. URL: http://tools.ietf.org/pdf/rfc792.pdf (cit. on p. 12).

[21] J. Postel. *RFC793 TRANSMISSION CONTROL PROTOCOL DARPA INTERNET PROGRAM PROTOCOL SPECIFICATION*. IETF, 1981. URL: http://tools.ietf.org/pdf/rfc793.pdf (cit. on p. 12).

[22] Jon Postel. *RFC760 DOD STANDARD INTERNET PROTOCOL*. IETF, 1980. URL: http://tools.ietf.org/pdf/rfc760.pdf (cit. on p. 12).

[23] Jonathan B. Postel. *RFC821 Simple Mail Transfer Protocol*. IETF, 1982. URL: http://tools.ietf.org/pdf/rfc821.pdf (cit. on p. 2).

[24] B. Ramsdell. *RFC2440 Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.1 Message Specification*. IETF, 2004. URL: http://tools.ietf.org/pdf/rfc2440.pdf (cit. on p. 12).

[25] B. Ramsdell. *RFC3851 Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.1 Message Specification*. IETF, 2004. URL: http://tools.ietf.org/pdf/rfc3851.pdf (cit. on p. 11).

[26] *RFC1087 Ethics and the Internet*. IETF, 1989. URL: http://tools.ietf.org/pdf/rfc1087.pdf (cit. on p. 8).

[27] T. Socolofsky and C. Kale. *RFC1180 A TCP/IP Tutorial*. IETF, 1991. URL: http://tools.ietf.org/pdf/rfc1180.pdf (cit. on p. 12).

[28] Wikipedia. *Anonymous remailer — Wikipedia, The Free Encyclopedia*. 2013. URL: http://en.wikipedia.org/w/index.php?title=Anonymous_remailer&oldid=584455506 (cit. on p. 6).

[29] Wikipedia. *Edward Snowden — Wikipedia, The Free Encyclopedia*. 2013. URL: http://en.wikipedia.org/w/index.php?title=Edward_Snowden&oldid=586147644 (cit. on p. 1).

[30] N. Williams. *RFC4401 Sender ID: Authenticating E-Mail*. IETF, 2006. URL: http://tools.ietf.org/pdf/rfc4401.pdf (cit. on p. 12).

[31] N. Williams. *RFC4408 Sender Policy Framework (SPF) for Authorizing Use of Domains in E-Mail, Version 1*. IETF, 2006. URL: http://tools.ietf.org/pdf/rfc4408.pdf (cit. on p. 12).