

# Anti Spam

→ Introduction, Basis and Research

Prof. Dr.

**Norbert Pohlmann**

Institute for Internet Security - if(is)  
University of Applied Sciences Gelsenkirchen  
<http://www.internet-sicherheit.de>



if(is)  
internet security.

# Content

- **Aims and outcomes of this lecture**
- **Terms, Definitions and Damages**
- **E-Mail Infrastructure**
- **Sources of Spam**
- **Anti Spam Techniques**
- **Generalized View (Survey, Estimation)**
- **What can the different Stakeholders do**
- **Summary**

- **Aims and outcomes of this lecture**
- Terms, Definitions and Damages
- E-Mail Infrastructure
- Sources of Spam
- Anti Spam Techniques
- Generalized View (Survey, Estimation)
- What can the different Stakeholders do
- Summary

# Anti Spam

## → Aims and outcomes of this lecture

### Aims

- To introduce in the topic Spam in general and Anti Spam techniques
- To explore the problems and the damages we have with Spams
- To analyze the Anti Spam techniques
- To assess the possibilities we have to solve the Anti Spam problem

### At the end of this lecture you will be able to:

- Understand what the basic problems and damages with Spam are.
- Know something about the different Anti Spam techniques.
- Understand what we can do against Spam.

# Content

- Aims and outcomes of this lecture
- **Terms, Definitions and Damages**
- E-Mail Infrastructure
- Sources of Spam
- Anti Spam Techniques
- Generalized View (Survey, Estimation)
- What can the different Stakeholders do
- Summary

# E-Mail Application

## → Overview

- E-mail is an elastic application with exchanges discrete media, like text, graphic, etc which are time-independent.
- **12% of the bandwidth** of the international IP carrier is **e-mail traffic**.
- **Several billion ( $10^{12}$ ) of e-mail are exchanged worldwide** (120 million/day; 3.6 billion/month; 2007).
- **E-mail is *not* designed as a reliable service**, nonetheless e-mail is used widely for the communication in the business and private field.
- Spam, virus and other **e-mail** related weak points are security problems which **produce high damage and a high security risk**.
- This trend rises the question:  
**How long can we use easy and efficient the e-mail service in the future?**

# Spam

## → Terms

- **SPAM ≠ Spam!**
  - SPAM – Trademark for food (Hormel Foods Inc.)
  - Also a software trademark in USA (correctly SPAM™ / SPAM®)
- **UCE = Unsolicited Commercial E-Mail**
  - Unsolicited (recipient has not given consent)
  - Commercial
- **UBE = Unsolicited Bulk E-Mail**
  - Unsolicited (recipient has not given consent)
  - Bulk (not necessarily commercial)
- **Spam = unwelcome e-mail**

# E-Mail

## → What are the dangers?

---

- **Everyone can send us e-mail!**
  - Those, we want → **OK (Ham)**
  - Those, we do not want (commercial, politic contents, criminal intent, ...) → **Spam**
  - Those, that harm directly → **Virus, worm, Trojan horses and phishing**
- **E-mail is like a post card!**
  - No guarantee of confidentiality (as with BlackBerry problems)!
  - Credit card and bank information is sent as plaintext!
- **No verifiability ! ! ! ! (Problem)**
  - Sender of an e-mail, authenticity of the contents of an e-mail
  - Be sure that an e-mail has actually arrived (Orders, etc.)
  - Commitment/non-repudiability (Hotel rooms, conferences, etc.)



# Definition

## → Spam (1/2)

- Spam is **unsolicited e-mail of no value or sense to the recipient!**
- „**Unsolicited**“ is individual/subjective...
  - 92% consider commercial e-mail as spam (shouldn't this be 100%?)
  - Advertisement of political or civil groups: only 74% consider it as spam
  - ...Nonprofit or charity: only 65%!
- **But: Spam messages have in common:**
  - Spam is sent in bulk
  - There is some kind of commercial or financial background
  - It can be considered a denial of service attack
- Legitimate e-mail is sometimes called **ham** (as opposite to spam)

# Definition

## → Spam (2/2)

- „Spamming is the **abuse of electronic messaging systems** to send unsolicited bulk messages which are generally undesired.“

wikipedia.org

- “An electronic message is "spam" IF:

(1) the recipient's **personal identity and context are irrelevant** because the message is equally applicable to many other potential recipients;

AND

(2) the recipient **has not** verifiably granted deliberate, explicit, and still-revocable **permission** for it to be send.”

spamhaus.org

# Definition

## → False Negative/False Positive (Spam)

- In the context of spam and anti spam
  - **False Negative**  
...if a spam message is not detected as being spam  
(test was **falsely negative**)
  - **Problem:** False Negative is not filtered (is stored in Inbox)
  - Typically < 0,1%, ideally 0%!

*0,1 -> every 1.000 e-mail is wrong (e.g. one per week)*

- **False Positive**  
A legitimate message (**ham**) is falsely considered spam  
(test was **falsely positive**)
- **Problem:** Legitimate e-mail gets filtered (destructive) – unacceptable!
- Typically < 0,001% – 0,0001%, ideally 0%  
(i.e. MessageLabs < 0,0004%)

*0,001 -> every 100.000 e-mail is wrong (e.g. one per year)*

# Losses caused by Spam

## → Overview (1/2)

- **Loss of productivity**
  - Real time notification, detect, filter and delete a spam message
- **Storage**
  - Spam messages take up storage space
- **Bandwidth consumption**
  - Spam messages consume bandwidth
- **Security and threats**
  - Virus, Worms, Trojan horses, ...



# Losses caused by Spam

## → Overview (2/2)

- **Denial of Service on e-mail servers**
  - Backscatter of third-party e-mail servers
- **Reputation**
  - Spammers abuse e-mail servers of companies and institutions (Pornography, violence, illegal activities, ...)
- **Cost of spam**
  - Spam filters need to be put in place, loss of productivity, ...
- **Usability**
  - E-mail might be rendered unusable due to high spam volumes

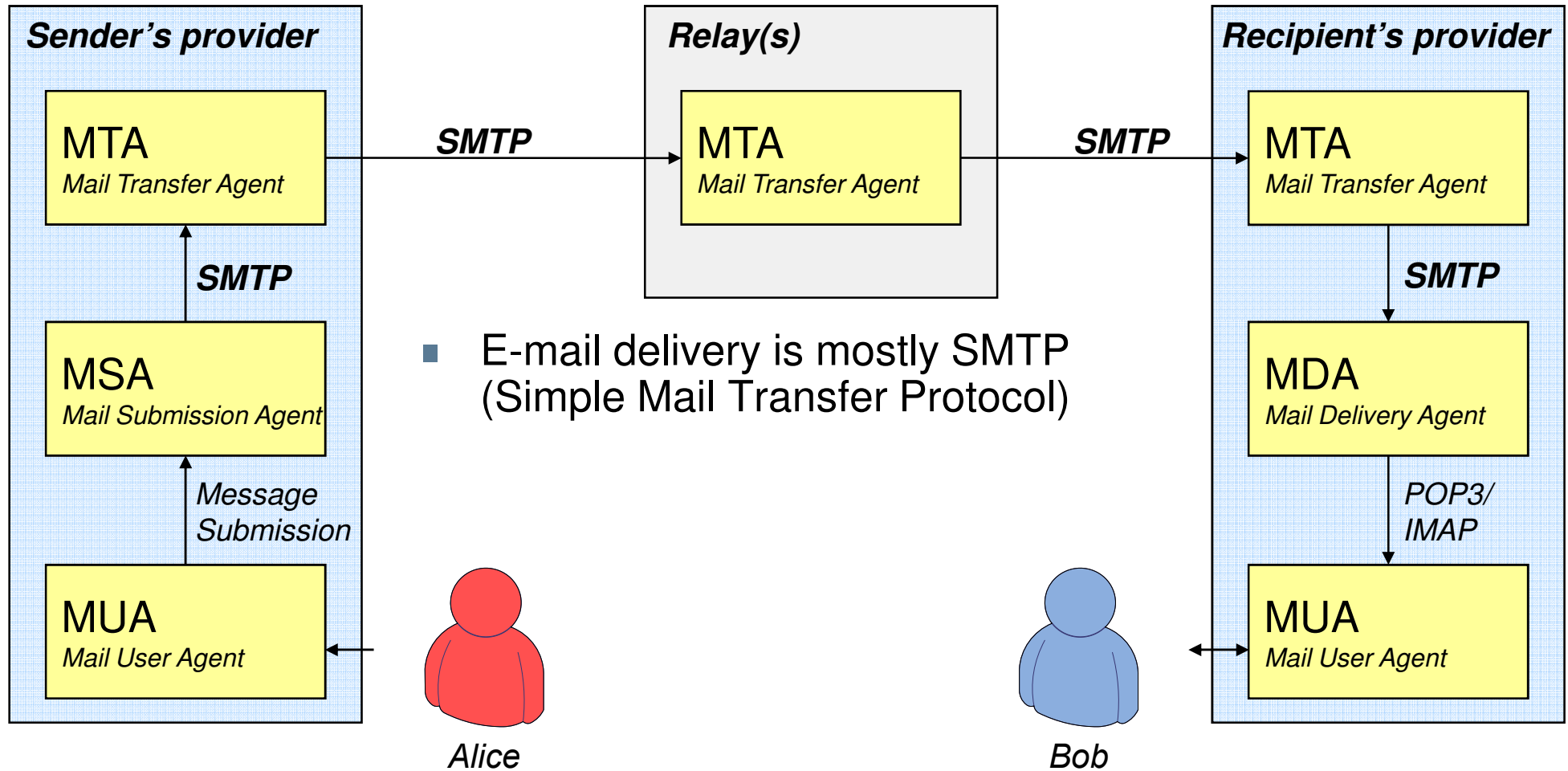


# Content

- Aims and outcomes of this lecture
- Terms, Definitions and Damages
- **E-Mail Infrastructure**
  - Sources of Spam
  - Anti Spam Techniques
  - Generalized View (Survey, Estimation)
  - What can the different Stakeholders do
  - Summary

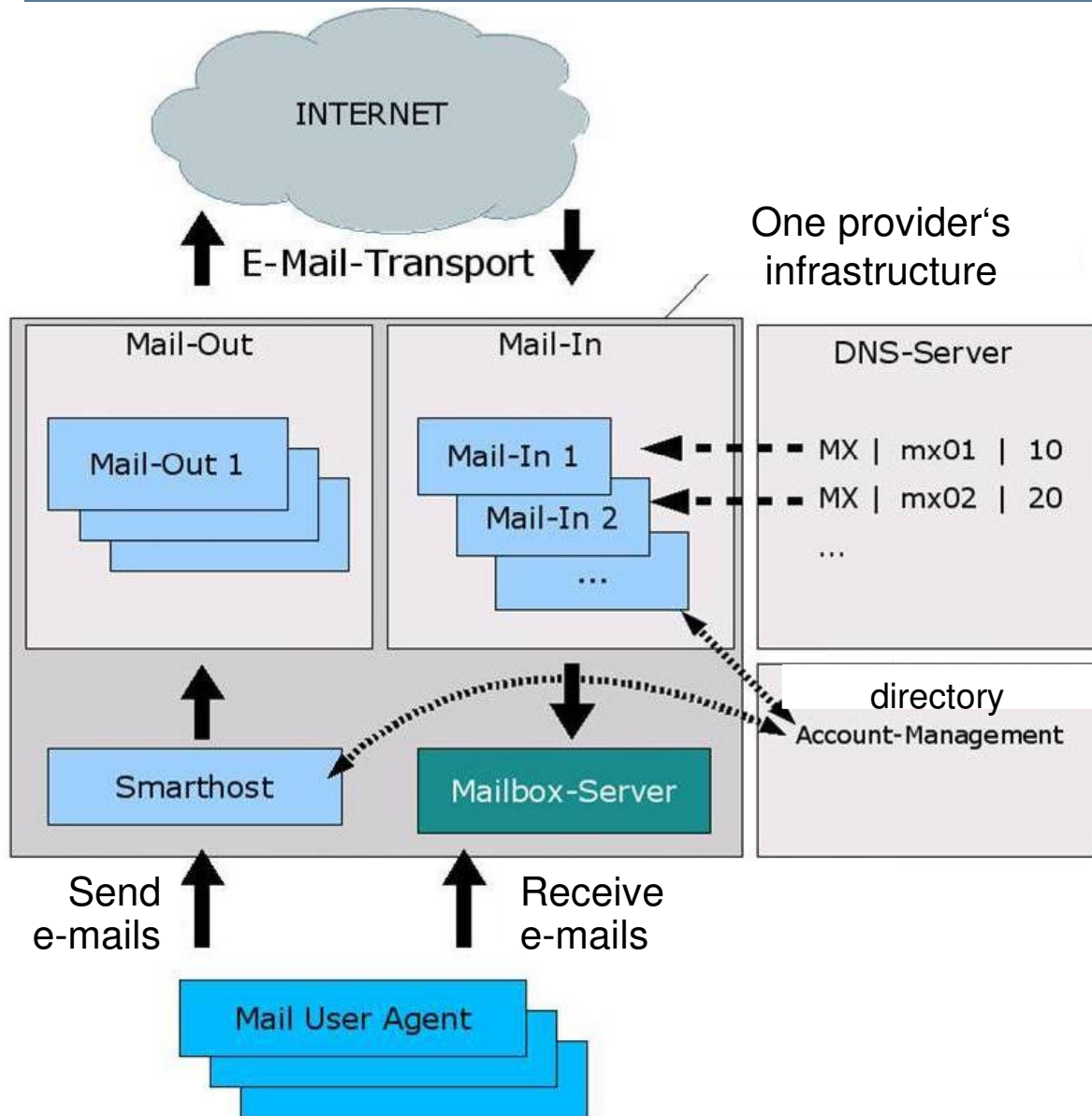
# E-Mail Infrastructure

## → Overview



# E-Mail Infrastructure

## → E-Mail Service Provider



- Mail User Agent interacts with Smarthost und Mailbox-Server (high availability)
- Provider X sends outgoing e-mails via its Mail-Out to Provider Y on its Mail-In
- Account and Identity Management via Directory services
- MX RRs point to several Mail-Ins using different priorities



# Content

- Aims and outcomes of this lecture
- Terms, Definitions and Damages
- E-Mail Infrastructure
- **Sources of Spam**
- Anti Spam Techniques
- Generalized View (Survey, Estimation)
- What can the different Stakeholders do
- Summary

# Sources of Spam

## → Overview

- Spam Servers (1)
- Open Relays (2)
- Open Proxies (3)
- Zombie PCs und Botnets (4)
- Mail Server at the Provider (5)

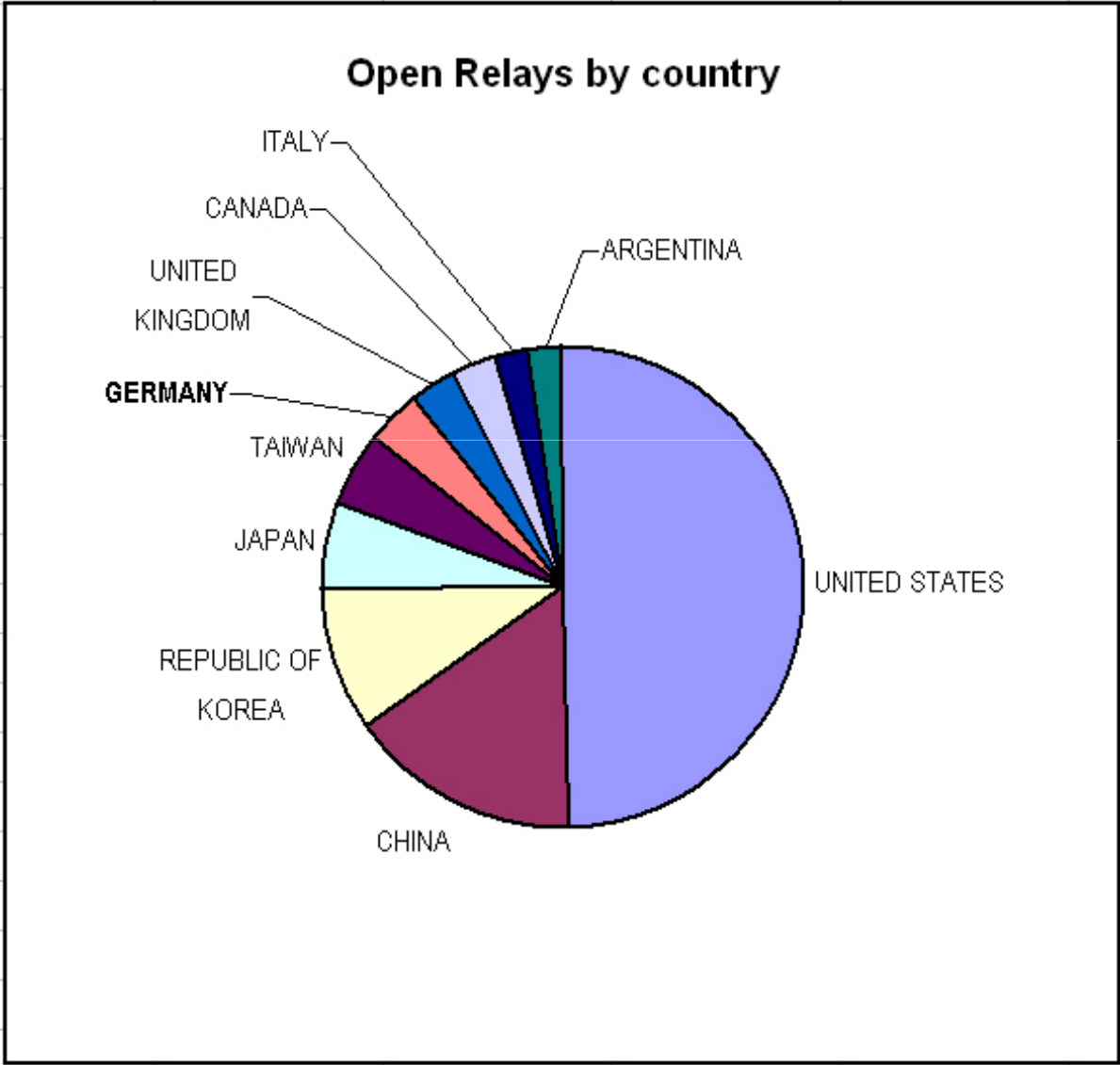
# Spam Servers

- Dedicated e-mail servers, operated by spammers for spamming
- Have static IP addresses
- Still cause a small fraction of spam today
- Usually they can easily be blocked by IP address blacklists

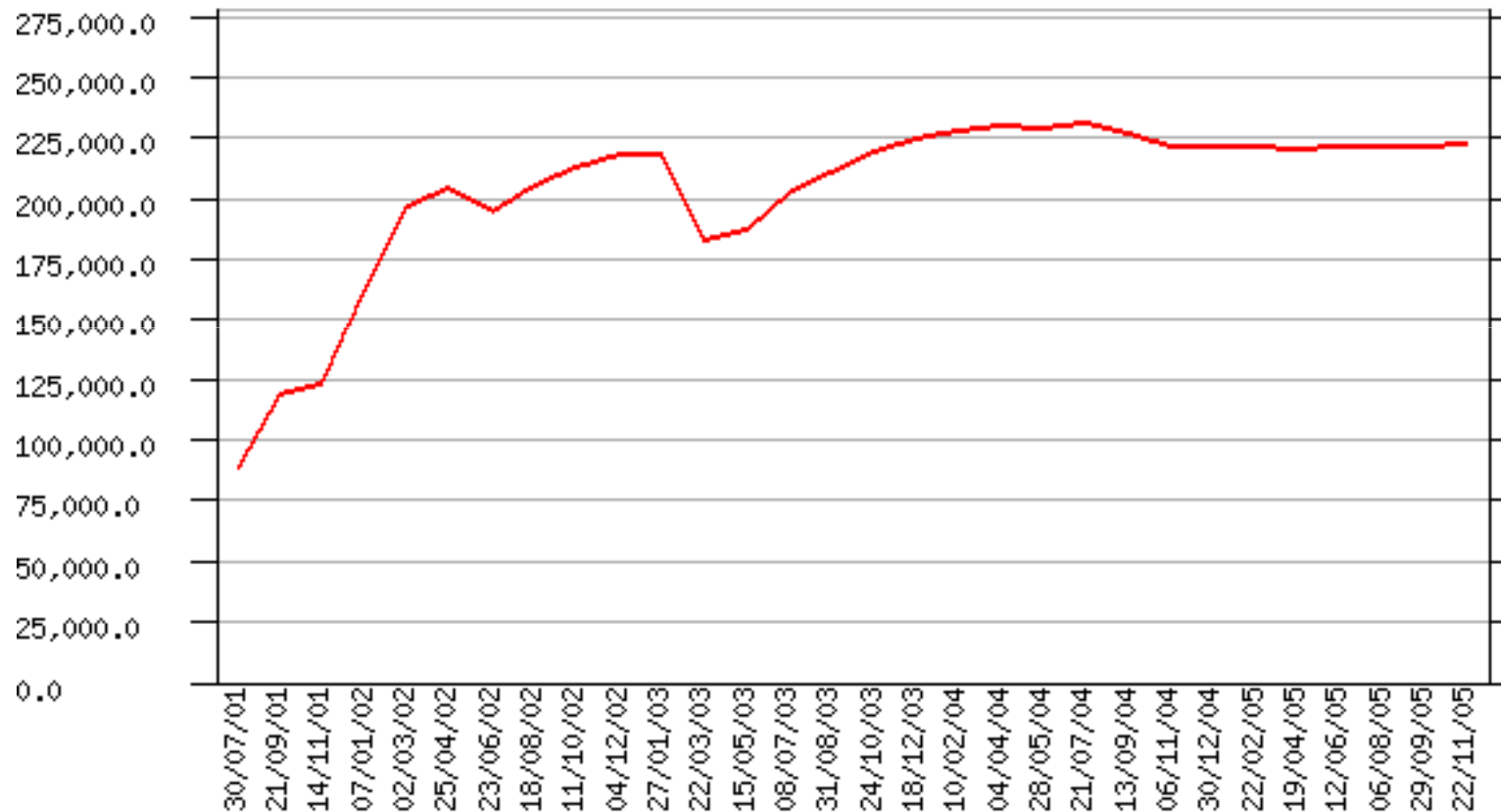
# Open Relays (1/3)

- Open Relay = an **e-mail server** that relays e-mail although there is
  - neither **sender**
  - nor **recipient** belong to the **local** domain
- **Problem:** Spammers abuse Open Relays without being detected
- Today Open Relays exist either
  - due to wrong configuration or
  - On purpose – „free internet“
- Open Relays can usually be blocked once they are known by IP address blacklists

# Open Relays (2/3)



# Open Relays (3/3)



# Open Proxies

- Classic
  - Misconfigured proxies used by spammers to cover the e-mail tracks
  - Anonymizing Proxies are „open“ by definition
    - Z.B. JAP, TU Dresden, <http://anon.inf.tu-dresden.de/>
- Modern
  - Term used for all-purpose bots (Zombie PCs)
  - Relay for all layer 4 protocols
- Usually no additional Received-Header in a spam-mail
- Today: spammers use chains of proxies to obfuscate even more

# Insecure CGI-Scripts / Formmail

---

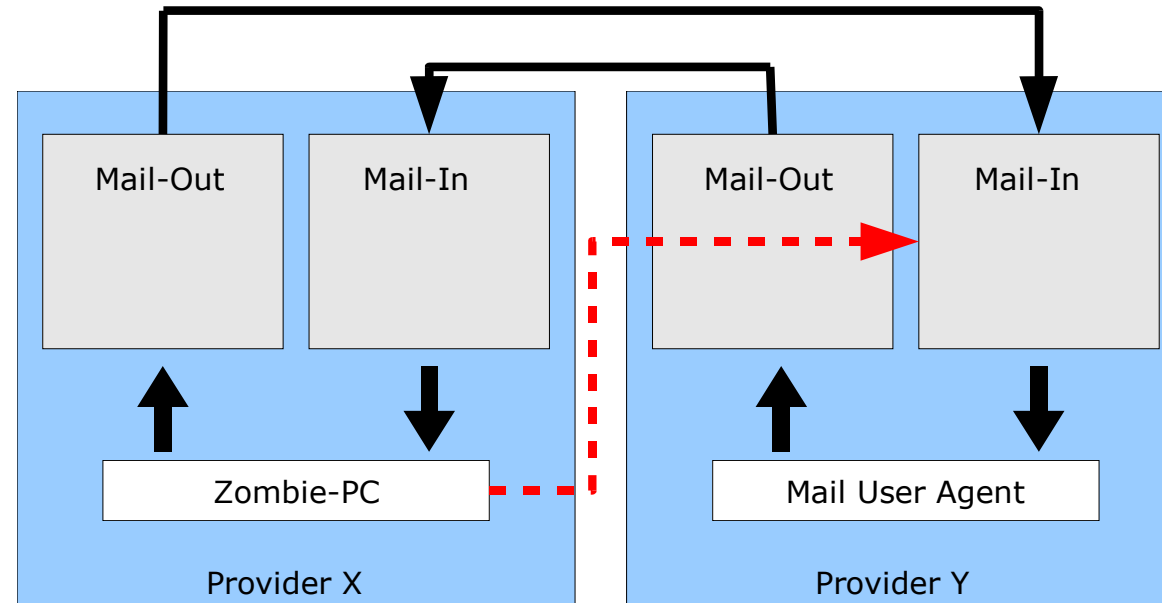
- (Free) CGI scripts in order to send e-mail from HTML forms
- Sometimes not secured
- **Problem:** web server operator is responsible



# Zombie PCs / Botnets

- **Zombie PC = remote controlled, hijacked computer**  
"Nur wegen der ‚Aldi-PC-Besitzer‘ kann sich Sober so stark verbreiten", sagt Prof. Dr. Norbert Pohlmann vom Institut für Internetsicherheit der Fachhochschule Gelsenkirchen."  
[Sueddeutsche]
  - Botnet = Network of Zombie PCs
  - Control channels are e.g. IRC or HTTP(S)
  - **Problem:**
    - Always on (access-connection -> flat rate) = high availability
    - The PC owner is responsible
    - Mostly spam activity is not noticed by the owner/user
  - More than 1.000.000 Zombies (Honeynet Research)
  - Per Botnet up to 400.000 bots
- ➔ **On the rise!**

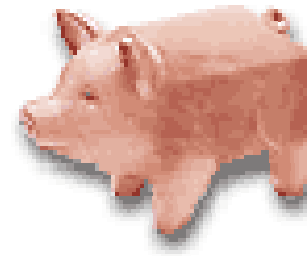
# The Zombie Problem



- Zombie (PC) = hijacked, remote controlled PC (the user does not notice)
- Zombies are aggregated to “botnets”
- Botnets are the main source of spam nowadays
- Zombies fluctuate and change their IP addresses due to dynamic dialup IP addresses – not easy to grasp
- Few ISPs provide their dialup IP space to be used in public blacklists – no e-mail should be sent from a user’s PC to a Mail-In directly!

# E-Mail Servers at Providers

- „Every ISP sooner or later has a spammer as customer “
  - **„pink contracts“**  
contracts that tolerate spamming
  - Examples: PSINet, AT&T, McColo
  - Conflict between Network Service Provider and E-Mail Service Provider (who is responsible?)
- **Nowadays not too much of a problem**



# Content

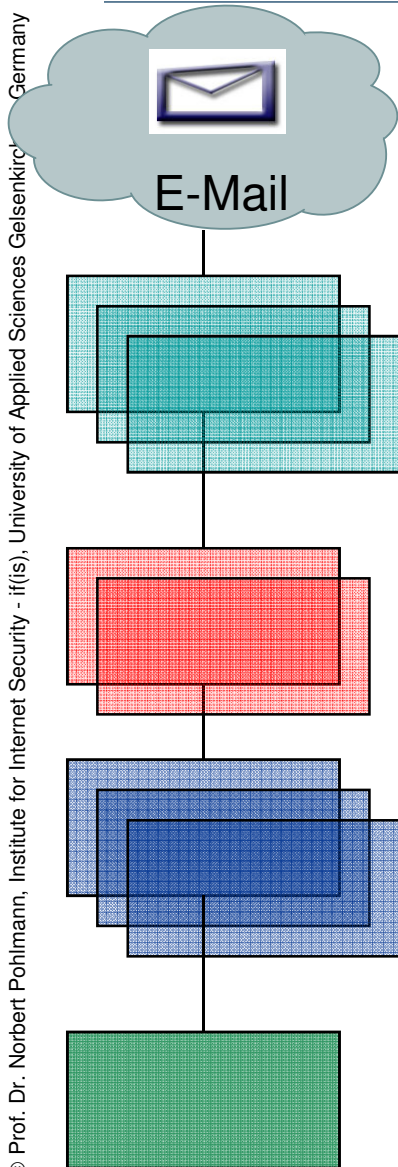
- Aims and outcomes of this lecture
- Terms, Definitions and Damages
- E-Mail Infrastructure
- Sources of Spam
- **Anti Spam Techniques**
- Generalized View (Survey, Estimation)
- What can the different Stakeholders do
- Summary

# Anti Spam Techniques

## → Overview

- **The Layer Model**
- **IP Layer**
  - Blacklisting (DNSBL) and White list
  - Distributed IP Reputation
- **TCP/IP Layer**
- **SMTP Layer**
  - Check HELO
  - Greylisting
- **MARID Methods**
- **Content-based Methods**

# The Layer Model



## External E-Mail-Gateway / E-Mail-Proxy (part of firewall)

- Checks on IP layer (IP address)
  - Blacklists (RBLs, Dynamic/Dial-Up IP, open relay, ...)
  - Reverse MX
  - Frequency analysis
- Checks on SMTP layer
  - Check the HELO argument
  - Check sender address (Black-/White-/Greylist)
  - Check recipient address (DB, directory)

1

## Spam filter

- Check of header and content level
  - Heuristically e-mail header and content analysis
  - Statistical methods, Word lists (Viagra, ...)
  - Checksum comparison, URIs/URLs (Phishing)

2

## Virus filter

- Check message and attachment for viruses

3

## Internal E-Mail-Server

Resource consumption

# IP Layer

## → Blacklisting (1/2)

- **Method:**  
Look up IP address of the connecting party in a blacklist  
(can be applied to white lists as well)
- Method is well-known!
- In practice:
  - Usually DNS-based (so called DNSBLs)
  - Request for 194.94.127.5 is formulated as 5.127.94.194.dns.bl.provider
  - Rewrite the IP address octet-wise and append domain of DNSBL provider

# IP Layer

## → Blacklisting (2/2)

---

- Some blacklists:
  - Spamhaus Block List      manual check (known spammers)
  - XCL/CBL                      combined list
  - RFC-Ignorant.org          put in non-RFC-conformal hosts
  - NiX Spam                      German blacklist
  - See <http://rbls.org>



# DNSBL

## → Detail

```
> web.de
Server: icarus.home
Address: 192.168.28.2

Nicht autorisierte Antwort:
web.de MX preference = 110, mail exchanger = mx-ha02.web.de
web.de MX preference = 100, mail exchanger = mx-ha01.web.de
> exit

chris@leo ~
$ nc mx-ha01.web.de 25
220 WEB.DE
helo leo
554 Transaction failed. For explanation visit http://freemail.web.de/reject/?auswahl4=80.133.22.95

chris@leo ~
$ nslookup 95.22.133.80.dnsbl.sorbs.net
Server: icarus.home
Address: 192.168.28.2

Nicht autorisierte Antwort:
Name: 95.22.133.80.dnsbl.sorbs.net
Address: 127.0.0.10

chris@leo ~
$ nslookup -qt=txt 95.22.133.80.dnsbl.sorbs.net
Server: icarus.home
Address: 192.168.28.2

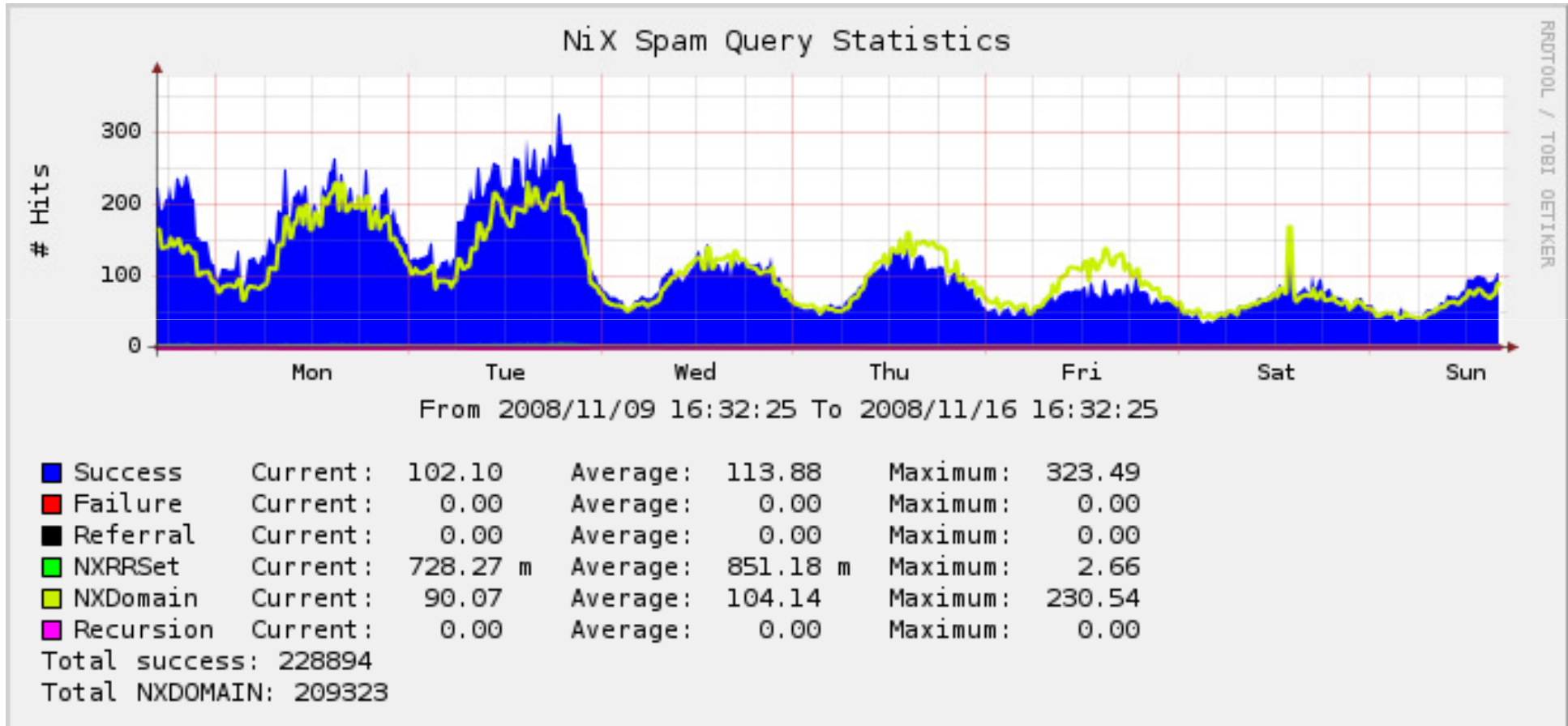
Nicht autorisierte Antwort:
95.22.133.80.dnsbl.sorbs.net text =

"Dynamic IP Addresses See: http://www.sorbs.net/lookup.shtml?80.133.22.95"

chris@leo ~
$
```

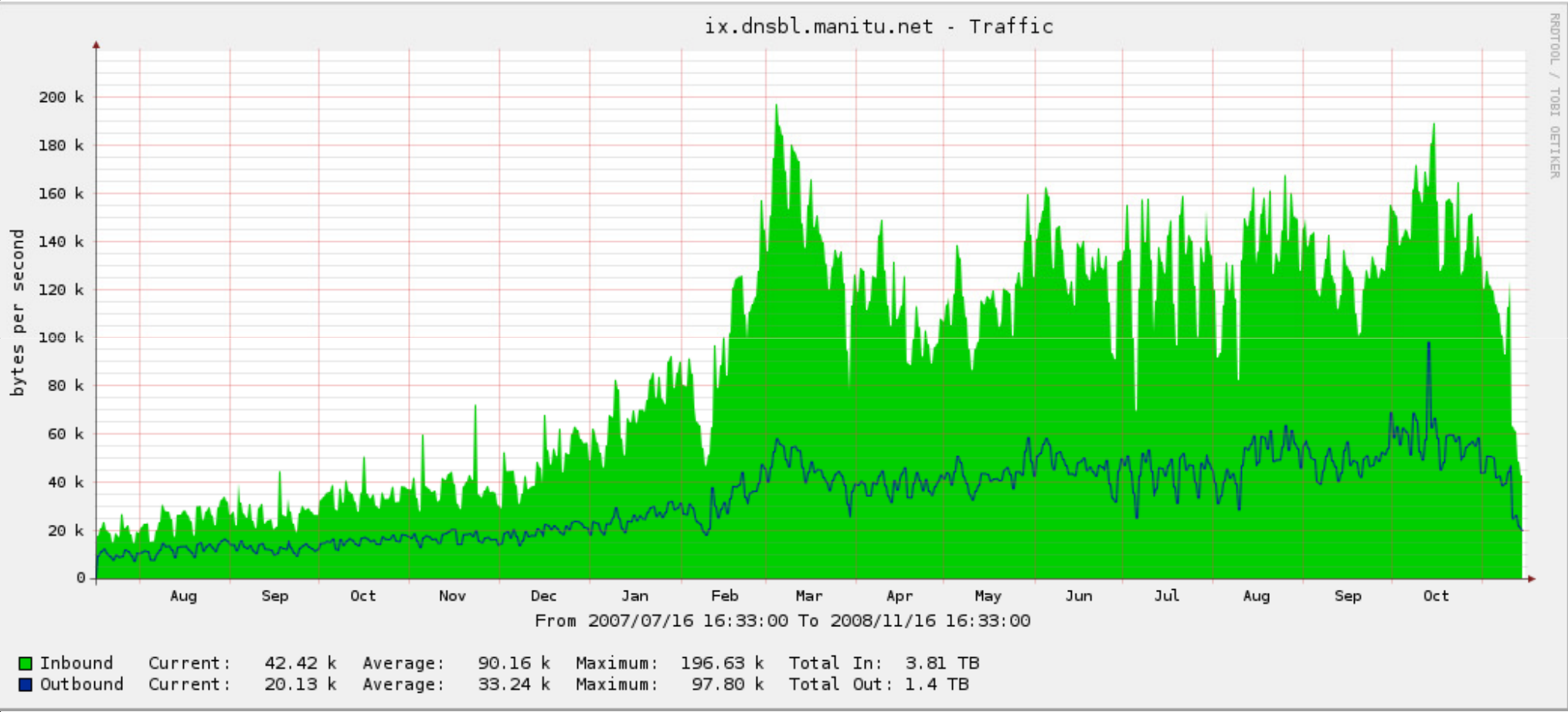
# DNSBL

## → Some Figures (1/2)



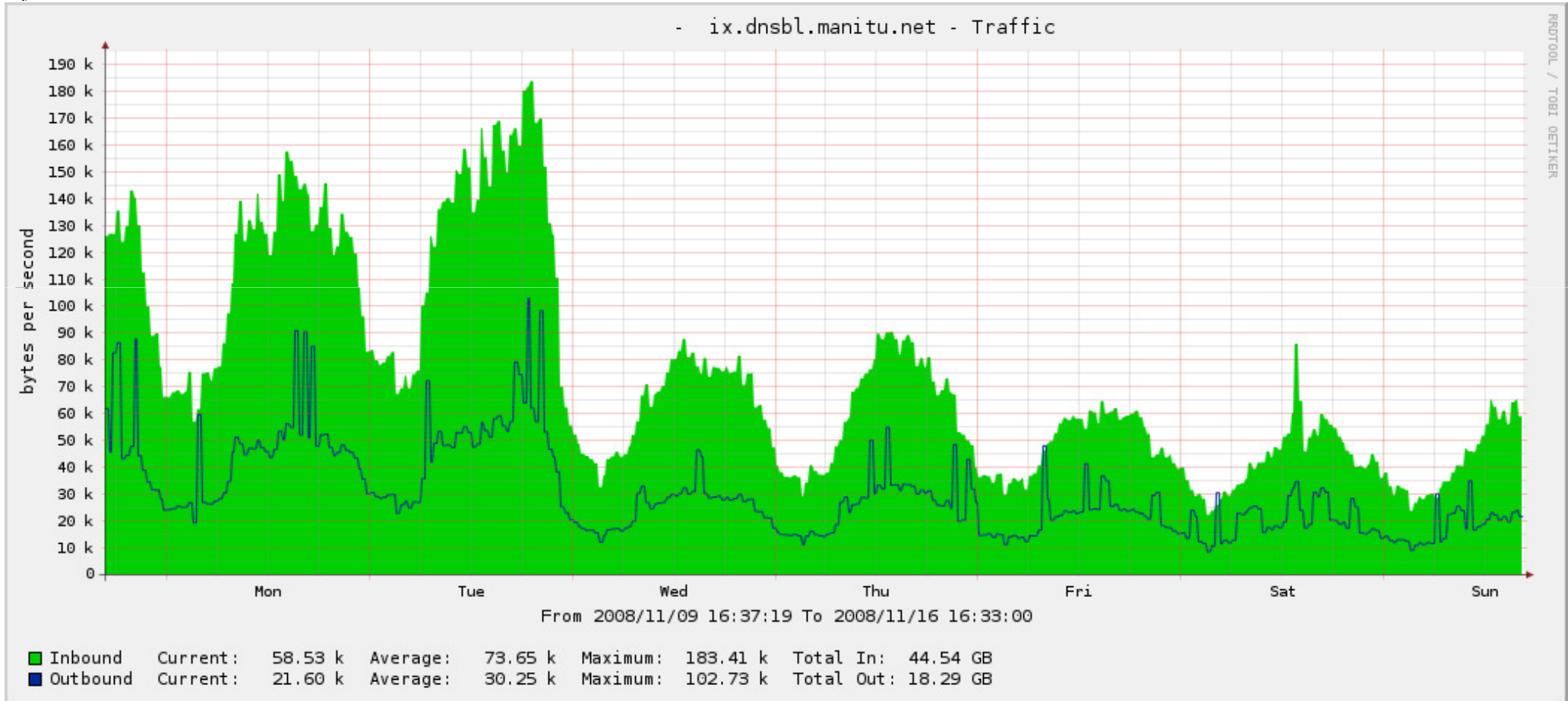
# DNSBL

## → Some Figures (2/2)



# McColo (US ISP) → Taken offline 11/11/2008

, Germany



PROTODOL / TOBI OETIKER

© Prof. Dr. Nor



# Review

- In order to compare several anti spam methods, collect 4 criteria
  - Advantages
  - Disadvantages
  - Indications to danger
  - Uncertainties

- Symbols



= Advantages



= Disadvantages



= Indications to danger



= Uncertainties

# Blacklists

## → Review



Blacklisting needs very few resources and protects against resource misuse, since email delivery is denied beforehand.



Blacklisting is independent from email's content, i.e. no liability to weakness of content filtering.



Adopting this method is very simple, even for multiple blacklists.



Risk assessment of blacklists is a very complex process and should be supported by experienced groups.



Blocking SMTP connections without looking into the emails might be dangerous, because no quarantining and in this way no recovery of false positives is feasible. Using blacklists for a scoring system instead of blocking of connections should be considered.

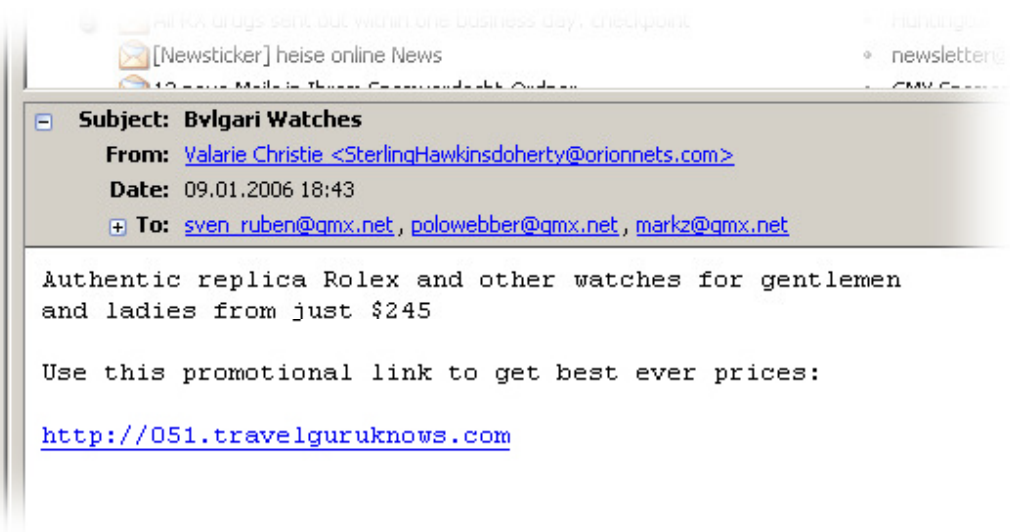


Condition of blocked IP might change more quickly than its reputation.

# URI based blacklists

## → Idea

- Not a „classic“ blacklist on the IP layer
- Based on the contents of a message, only the URIs
- **Method:**  
Look up the URIs of the message contents in a blacklist.
- **Advantages:**
  - Spammers must place a URI somewhere in their messages
  - Found by pattern matching
- **Disadvantages:**
  - URIs can be made individual per message (add changing subparts)
  - Use „disposable“ domains
  - Text in images cannot be detected easily (OCR)
  - Danger of DoS: put a domain (google.com) in the blacklist  
→ users cannot use google-links in emails anymore



# URI based blacklists

## → Review



Very efficient for spam mails with familiar domains.



Public domains (e.g. shortlink services) can be misused by spammers to avoid getting blacklisted with their own domain. Recursive queries could be a work-around, but would take some time.



No benefit for fighting spam without URIs (e.g. stock spam)



# White lists

## → Idea

- White listing is the opposite of blacklisting and prevents from anti-spam mechanisms being applied on the network level (i.e. grey- and blacklisting) for well known communication channels.
- A majority of (ham) e-mails is sent by **well known sources**, which do not need to be checked against grey- and blacklists, such as Mail-Outs of big E-Mail Service Providers
- **Private white list**
  - Usually configured in the Mail Transfer Agent
- **Public white list**
  - Usually also implemented by DNSBL-technique

# White lists

## → Review



Whitelisting disburdens the resources which are needed for requesting blacklists and allows abstaining from interference of legitimate connections.



Whitelisting might correct false entries of remote administrated blacklists.

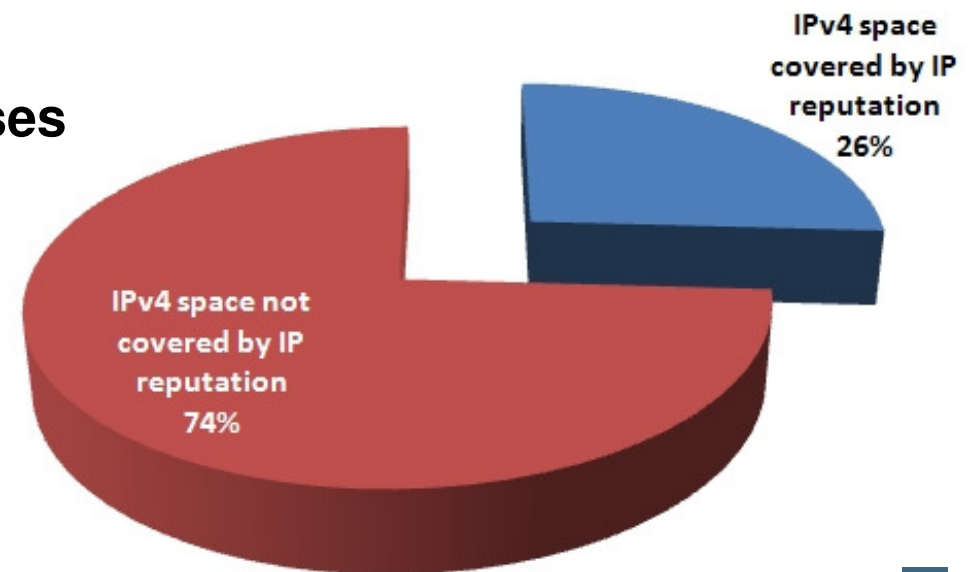


Once a whitelisted server begins sending spam, the trust to this server must be proofed if misuses occur.

# IP Reputation

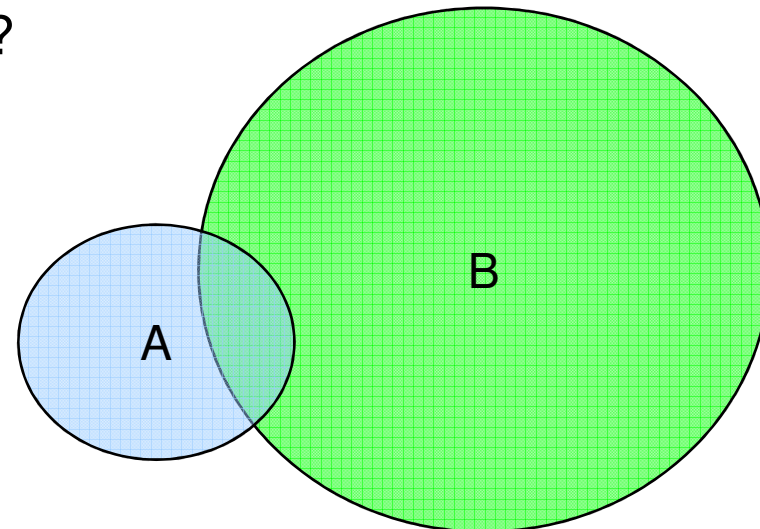
## → Idea

- Combining black and white listing is one of the most efficient anti spam mechanisms
- **However**
  - High dependency on black/white list providers
- **Aggregation of the most important black lists**  
=> **only 26% of advertised IPv4 addresses can be judged concerning e-mail reputation!**  
=> **~74% of advertised IPv4 addresses have no reputation**
- **More IP reputation attributes and IP space are needed!**
- **Less dependency on single Provider is also important.**



# Intersections between blacklists (1/2)

- **Blacklists are similar to each other**
  - Same data sources
  - Data exchange between blacklists
  - Same spammers are detected by many blacklists
  
- **Analysis of intersections**
  - How much does blacklist A cover blacklist B?
  - Which conclusions are possible?



# Intersections between blacklists (2/2)

- Array with intersections:

	all.dnsbl.sorbs.net	UCEPROTECT L1	NiX	dns	sbl.spamhaus.org	dns	CBL	pbl.spamhaus.org	xbl.spamhaus.org	dns.wl.org	Bogus ranges
all.dnsbl.sorbs.net	-	1,83	0,28	10,17	10,67	11,03	8,03	36,92	17,92	0,002	7,73
UCEPROTECT L1	11,61	-	2,34	1,97	0,58	2,93	64,14	69,96	64,79	0,026	0,01
NiX			-							0,064	0,02
dns				-						0,002	0,22
sbl.spamhaus.org					-					0,003	9,68
dns						-				0,003	0,28
CBL							-			0,001	0,00
pbl.spamhaus.org								-		0,000	1,48
xbl.spamhaus.org									-	0,001	0,01
dns.wl.org	0,003	0,007	0,002	0,003	0,002	0,003	0,001	0,002	0,002	-	0,027
Bogus ranges	0,03	0,00	0,00	0,00	0,01	0,00	0,00	0,34	0,00	0,000	-

PBL.spamhaus.org comparison ratio.

A blacklist includes bogus ranges only rarely.

Spamhaus integrates the entire CBL to their XBL.

Intersections between whitelists and blacklists are quite common.

Let's ask the operator of dns.wl.org:  
Seemed to be typos, I disabled the entries that were not in the list. I control the data more strictly.

Huh?!

# Regional views of blacklists

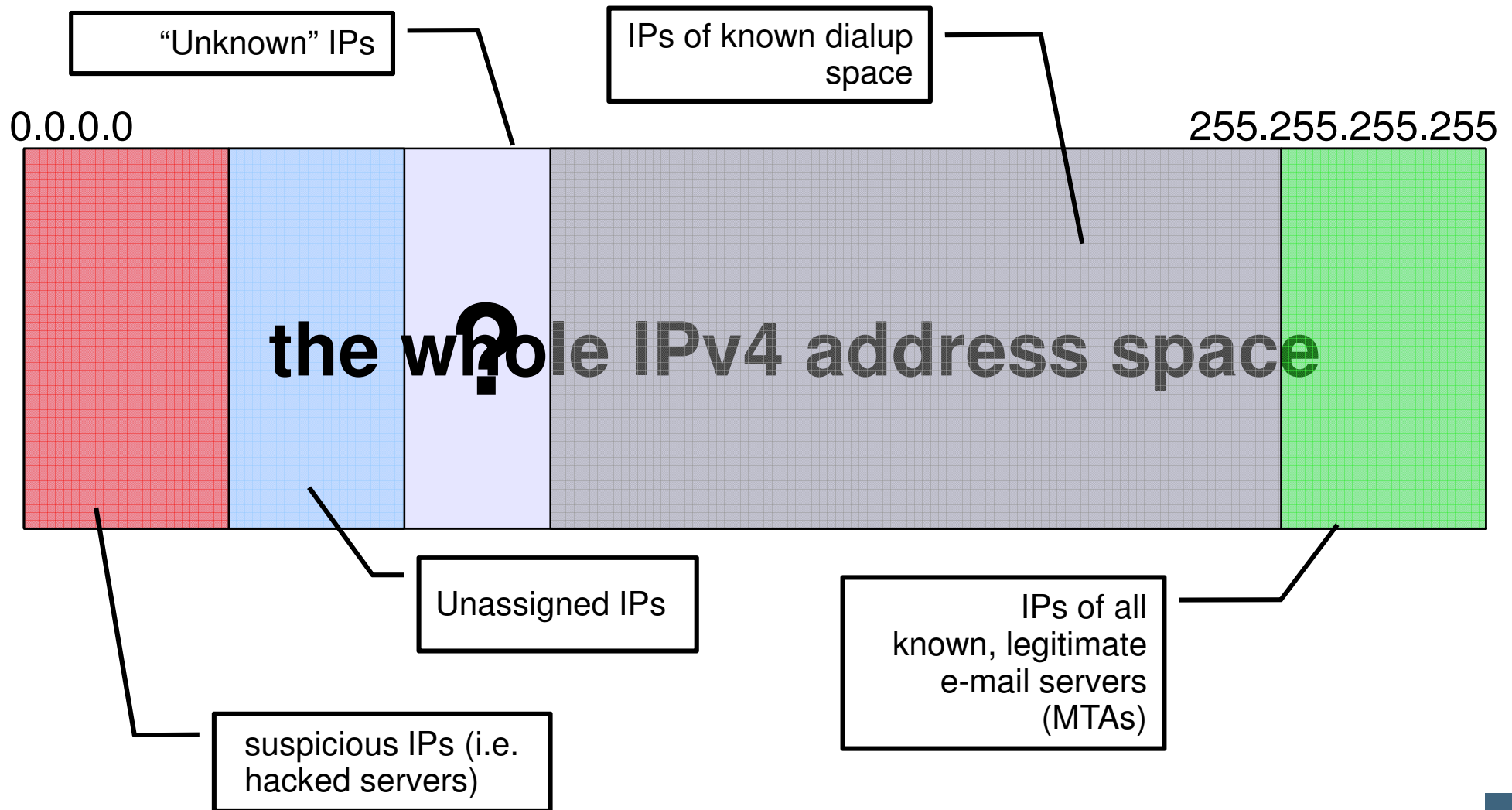
- Assign blacklist entries to regional attributes, e.g.
  - Country
  - RIR
  - Autonomous System
  - ...

rank	country	entries	range	quota
1	United States	49604	120467378	8.78%
2	Japan	5999	28940095	18.74%
3	China	8383	27448962	23.43%
4	Germany	1418	23568477	34.17%
5	(unknown)	3078	16897301	n/a
6	Canada	9233	10427689	14.29%
7	United Kingdom	2458	7778451	12.12%
8	France	1794	6940961	38.63%
9	Taiwan (, Province Of China	1259	6923462	37.01%
10	Mexico	848	6313481	38.83%
11	Spain	925	6247749	31.16%
12	Korea, Republic of (South)	3595	5944359	10.73%
13	Italy	976	5037499	20.90%
14	Brazil	4497	4405759	20.68%
15	Poland	2077	2916732	24.29%
16	Turkey	373	2730352	14.25%
17	Netherlands	1702	2660048	33.09%
18	European Union (can apply to any country in Europe)	2408	2630553	2.18%
19	Sweden	649	2507387	15.32%

Figure: Spamhaus' PBL by country

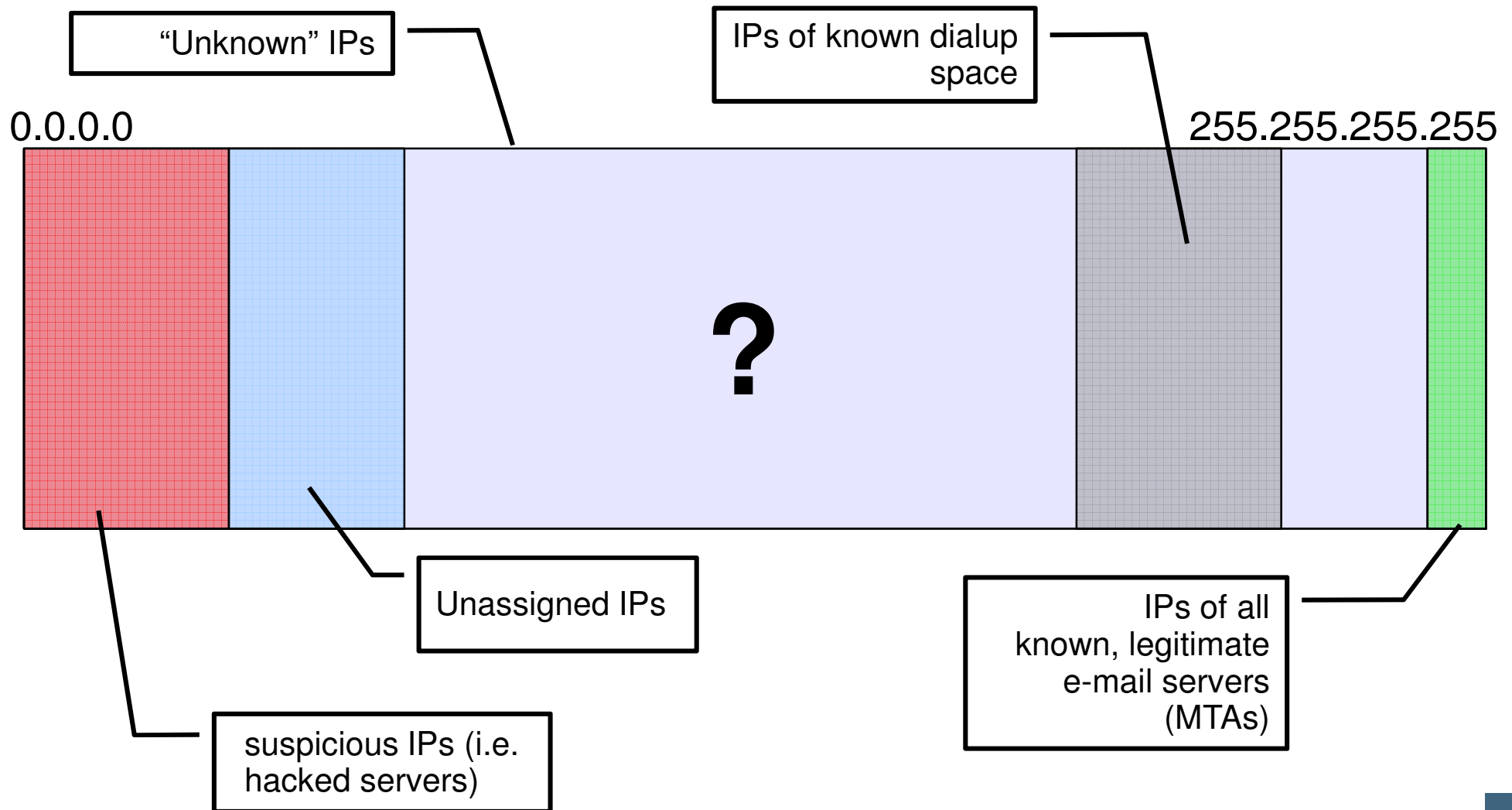
# IP Reputation

→ The IP map (possible attributes/space)



# IP Reputation

## → The IP map (attributes/space today)





# IP Reputation

## → Some Figures (E-Mail Infrastructure )

- Port 25 state analysis (incoming)
- German research „ISP“ - DFN (Deutsches Forschungsnetz)
- 0.08% of all IP addresses respond to a connection attempt on port 25
  - Every 1250<sup>th</sup> IP address „is“ a mail server (open port 25)
- **Challenges**
  - (every) open port 25 = SMTP
  - How many different IP addresses belong to one host?  
(timing measurement problems, honeypots/honeynets, ...)

# IP Reputation

## → Some thoughts (E-Mail Infrastructure )

- **Fact:**

- Theory:  $2^{32} = 4.2$  billion IPv4-Adressen
- 1.872 billion IPv4 addresses advertised (as of 2008-08)

- **Assumption:**

- 0.08% of all IPv4 addresses speak SMTP (incoming)

- **Result:**

- ~1.4 million IPv4 addresses that speak SMTP (incoming)

**→ ~1.4 million legitimate e-mail servers**

# IP Reputation

## → Dialup IPs (The main source of spam)

- Over 90% of all spam originates come from dialup IP addresses (bots)
- Bots send spam on a large scale to the inbound e-mail servers (MX) of e-mail service providers
- Smart hosts are omitted
- PCs get a (dynamic) dialup IP address when connecting to the internet

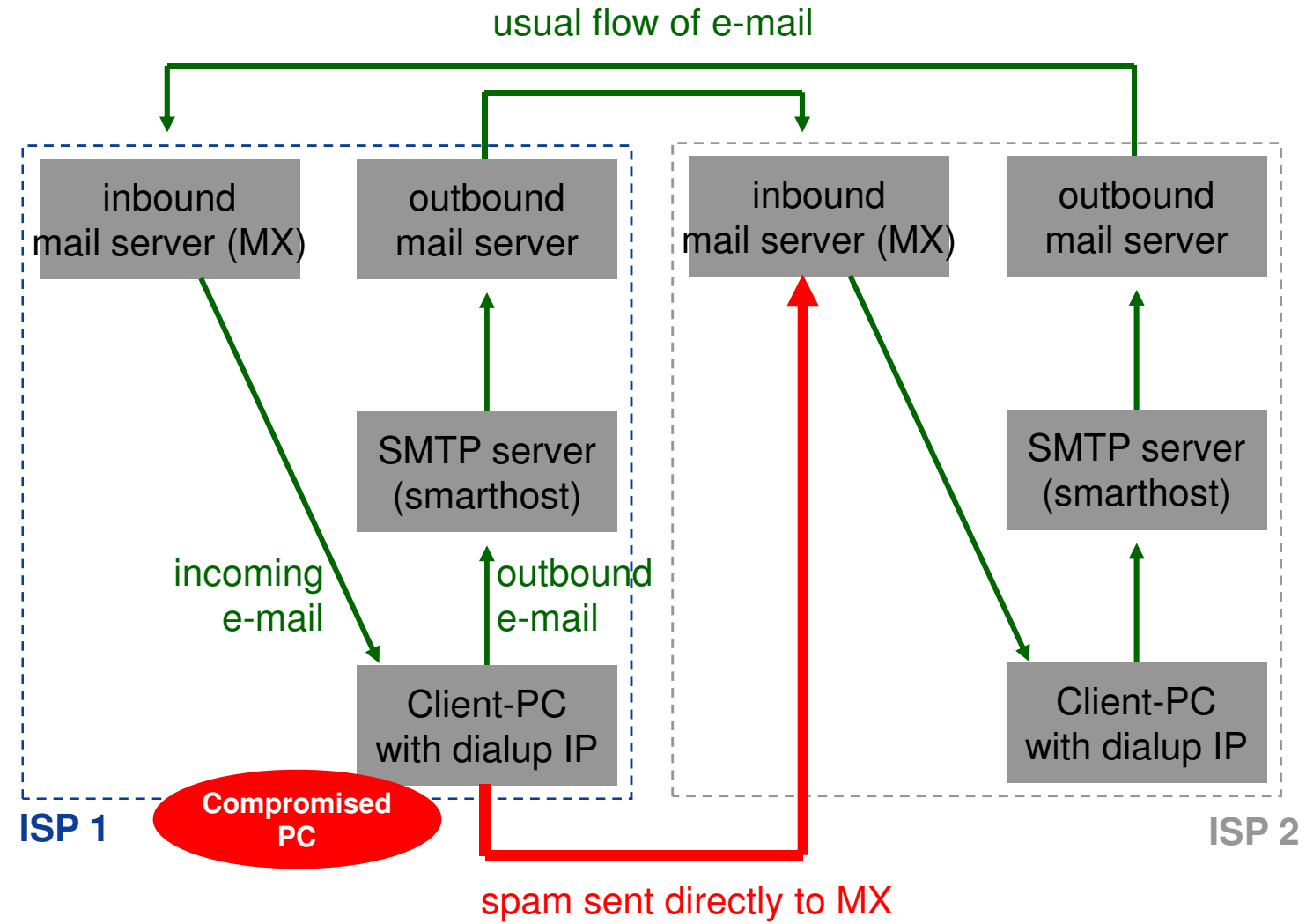
### ■ Conclusion

- **Dialup-IPs never send legitimate e-mail to inbound e-mail servers (MX)**
- **Blocking dialup IP addresses has no major drawback!**

# IP Reputation

## → Spam sent directly to MX

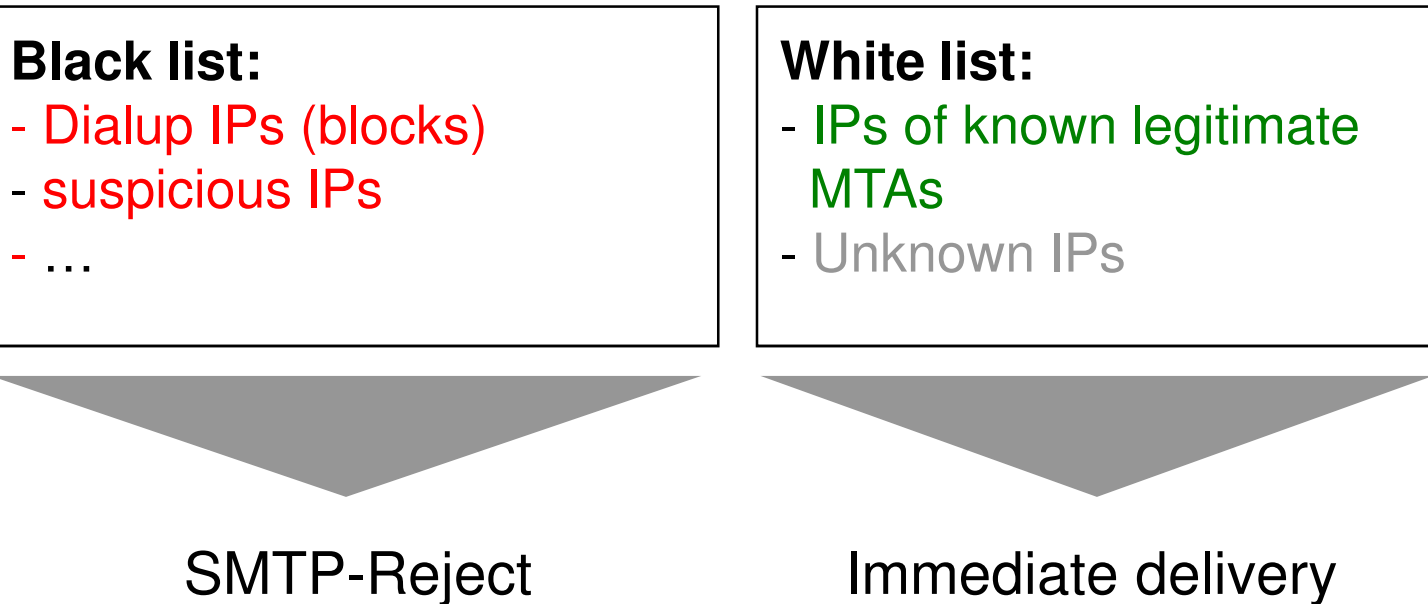
### Spam from dialup IPs



# Current best practice (1/4)

## → Black & white list

- E-mail service providers categorize IP addresses (Mail-Gateways) into black and white lists



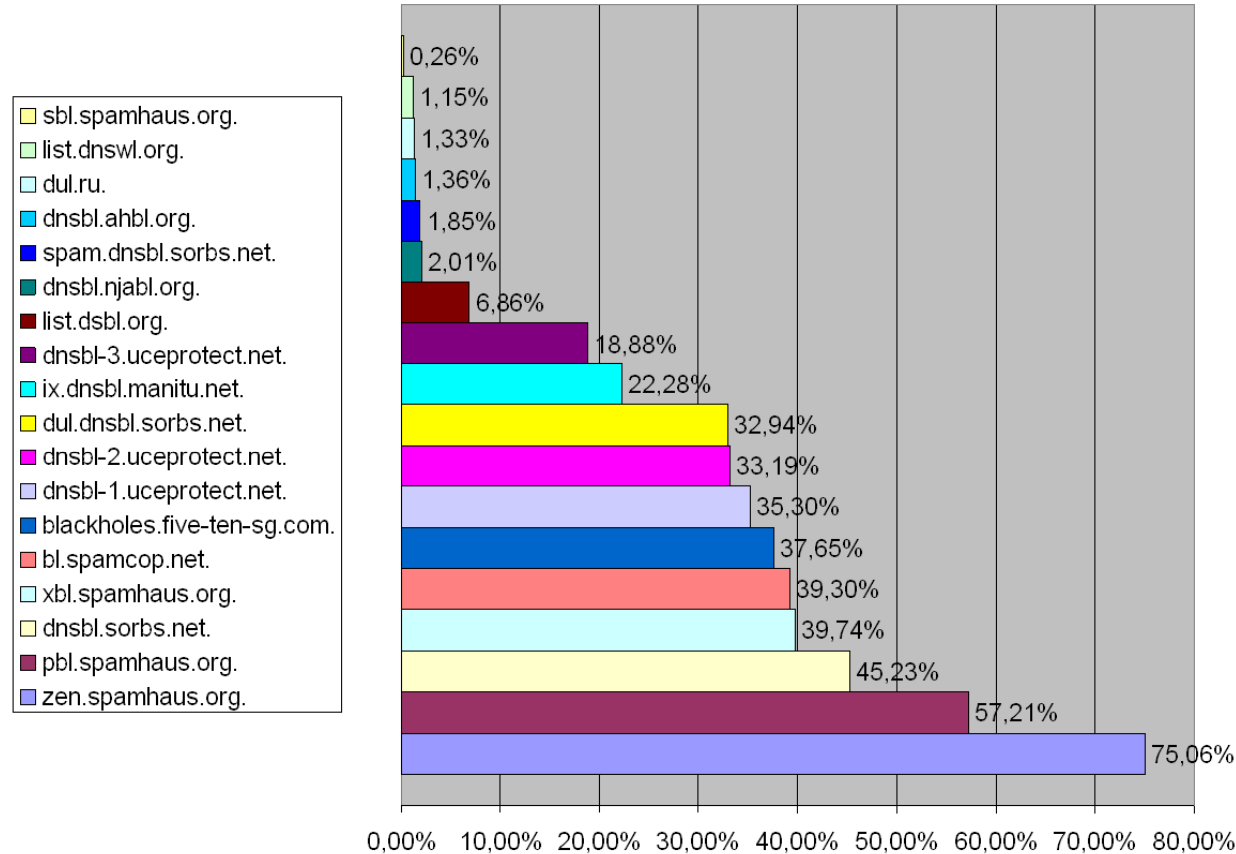
# Current best practice (2/4)

## → IP maps of Mail-Gateways

- Known e-mail gateways IP addresses are based on experience or observation of the ISP landscape
- Connections from unknown or non-suspicious IP addresses are allowed (e-mails are accepted)
- An entry on the black list depends on monitored spam activity (user complaints, amount of e-mails, frequency analysis, valid e-mail addresses, etc.)

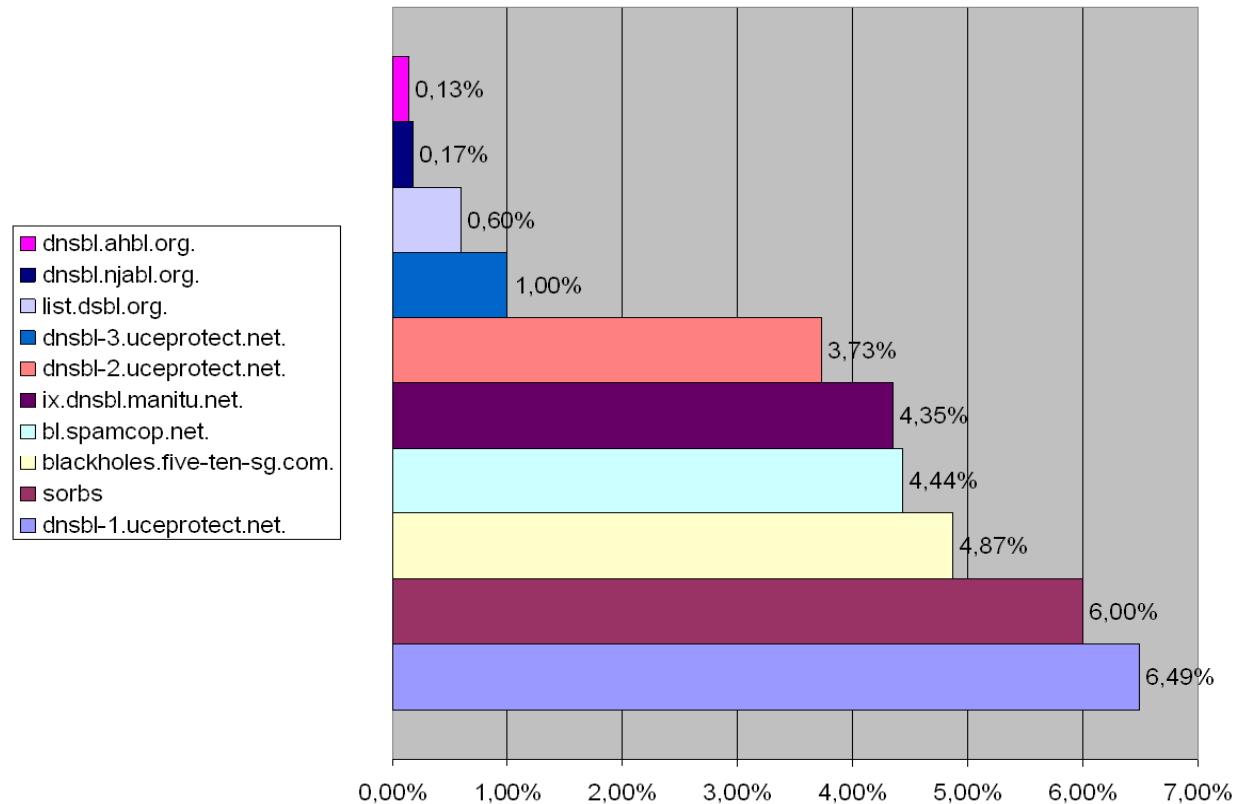
# Current best practice (3/4)

## → Combining multiple (black) lists



- One black list alone blocked 75%
- Others block less than 75%
- What is the gain in terms of blocking spam by using multiple lists?

# Current best practice (4/4) → Combining multiple (black) lists



- Gain in blocking spam using multiple black lists
- Using a second list blocks +6.5%, a third +6%, a fourth +4.8%
- **Result: Combining multiple black lists makes sense!**



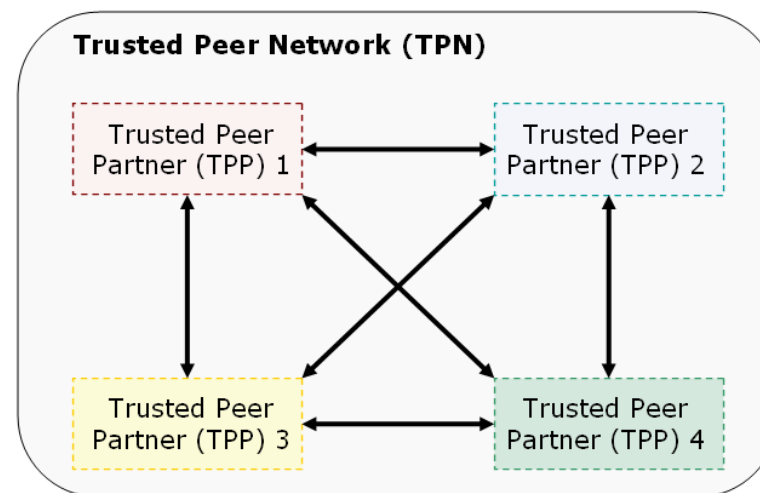
# Potential of IP maps

- The rate of detected spam mails must be enhanced
- Spamming IP addresses should be identified much quicker than today
- An international IP map should be established in order to fight spam in general and for long term
- **Action**
  - **Exchange of IP maps between ISPs world-wide**

# Distributed IP reputation system

## → Idea

- The distributed IP reputation system consists of a **network of participants (Trusted Peer Network)**, which helps to get **describing attributes of IP addresses**.
- Therefore it is based on participants **sharing their view on the whole internet**, expressed in categorizing IP addresses.
- The **idea** of the distributed IP reputation system is to **share information held by many different providers (Trusted Peer Partner – TPP)**.
- Sharing in detail means to pool single IP lists of the participants between other participants and to **achieve a most detailed and complete IP list** by aggregating this information.
- The working time of providers would be shared and **suspicions about spam sources** could be **amplified** or **weakened** by other parties.



# Distributed IP reputation system

## → Self-declaration data / observation

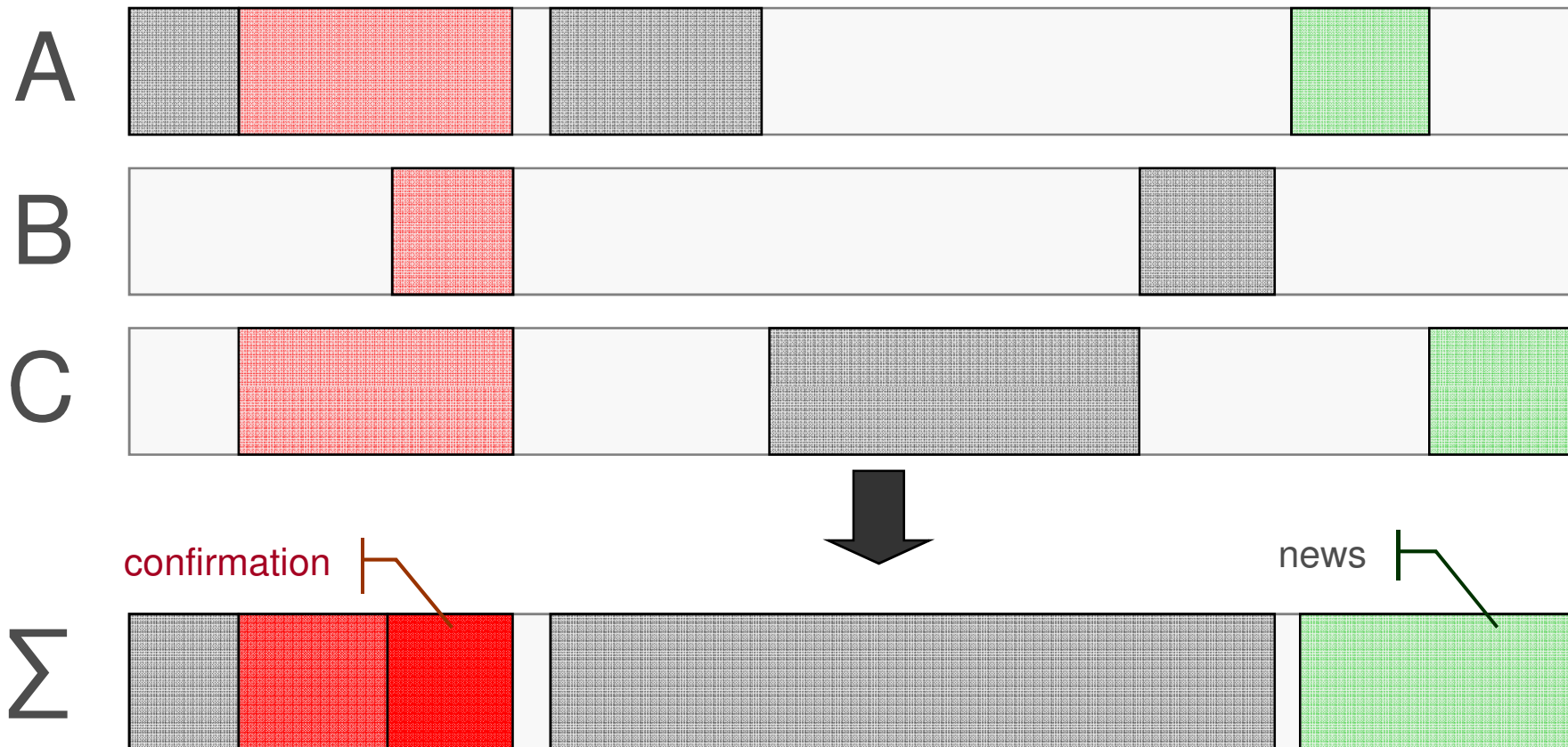
- ISPs exchange self-declaration information as well as observation regularly.
- In fact, these are lists of IP addresses with describing attributes (IP reputation).

Self-declaration data	Observation
<ul style="list-style-type: none"><li>- IPs of outbound mail servers</li><li>- Dialup-IPs (blocks)</li></ul> in addition maybe: <ul style="list-style-type: none"><li>- Static IPs, ...</li></ul>	<ul style="list-style-type: none"><li>- suspicious IPs</li><li>- IPs of non-maintained mail servers</li><li>- ...</li></ul>

- The self-declaration data may be checked against routing information.
- What are the advantages to exchange this kind of information?

# Distributed IP reputation system

## → Interpretation of „IP maps“ (1/2)

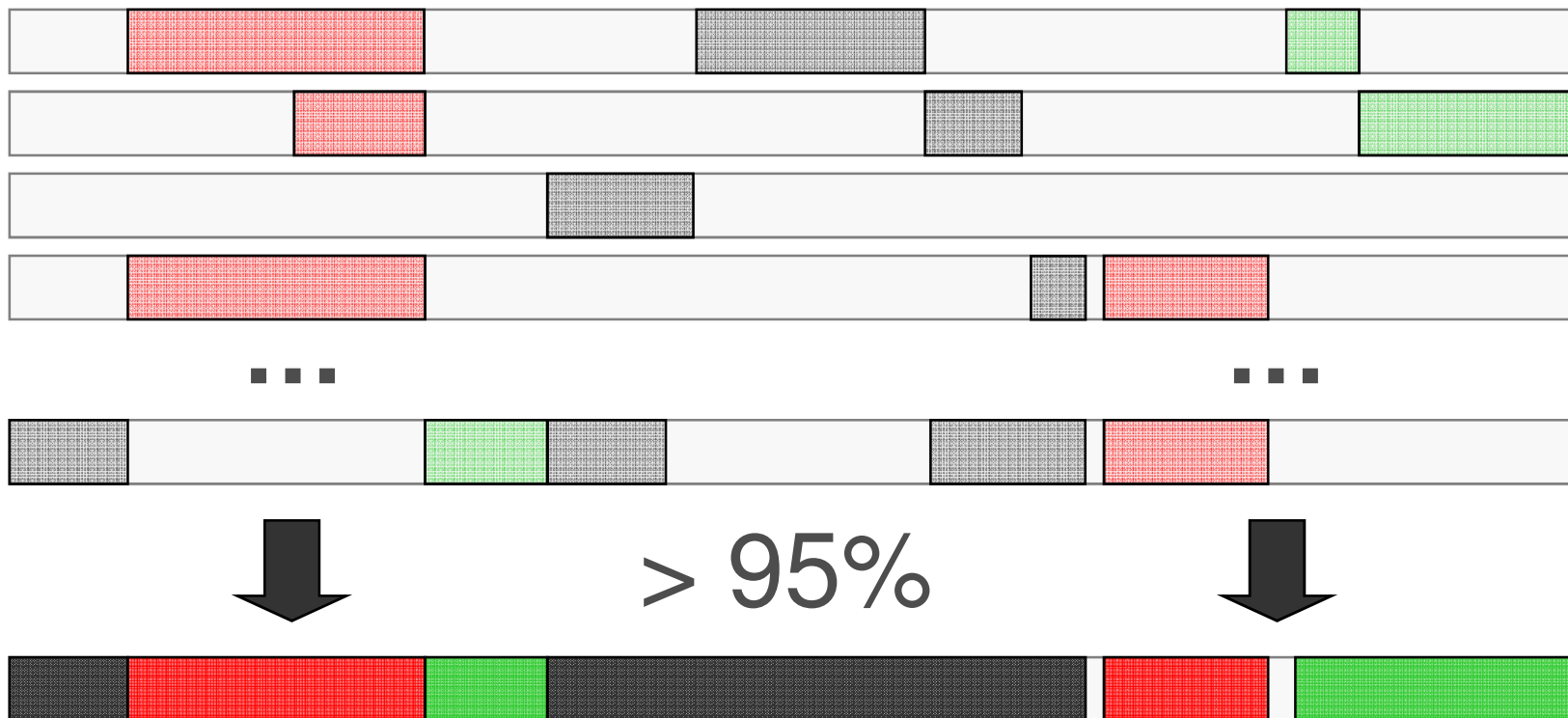


- Assessing the IP maps results in a confirmation of possible spam sources or reduce own observations (false positives)
- Shows new potential spam sources

# Distributed IP reputation system

## → Interpretation of „IP maps“ (2/2)

- Incoming self-declaration information as well as observations are aggregated into a composite IP map, with a better reputation of IP addresses.
- The higher the participation, the better for all users!
- **Aim: (nearly) complete IP map for the whole IP address space.**



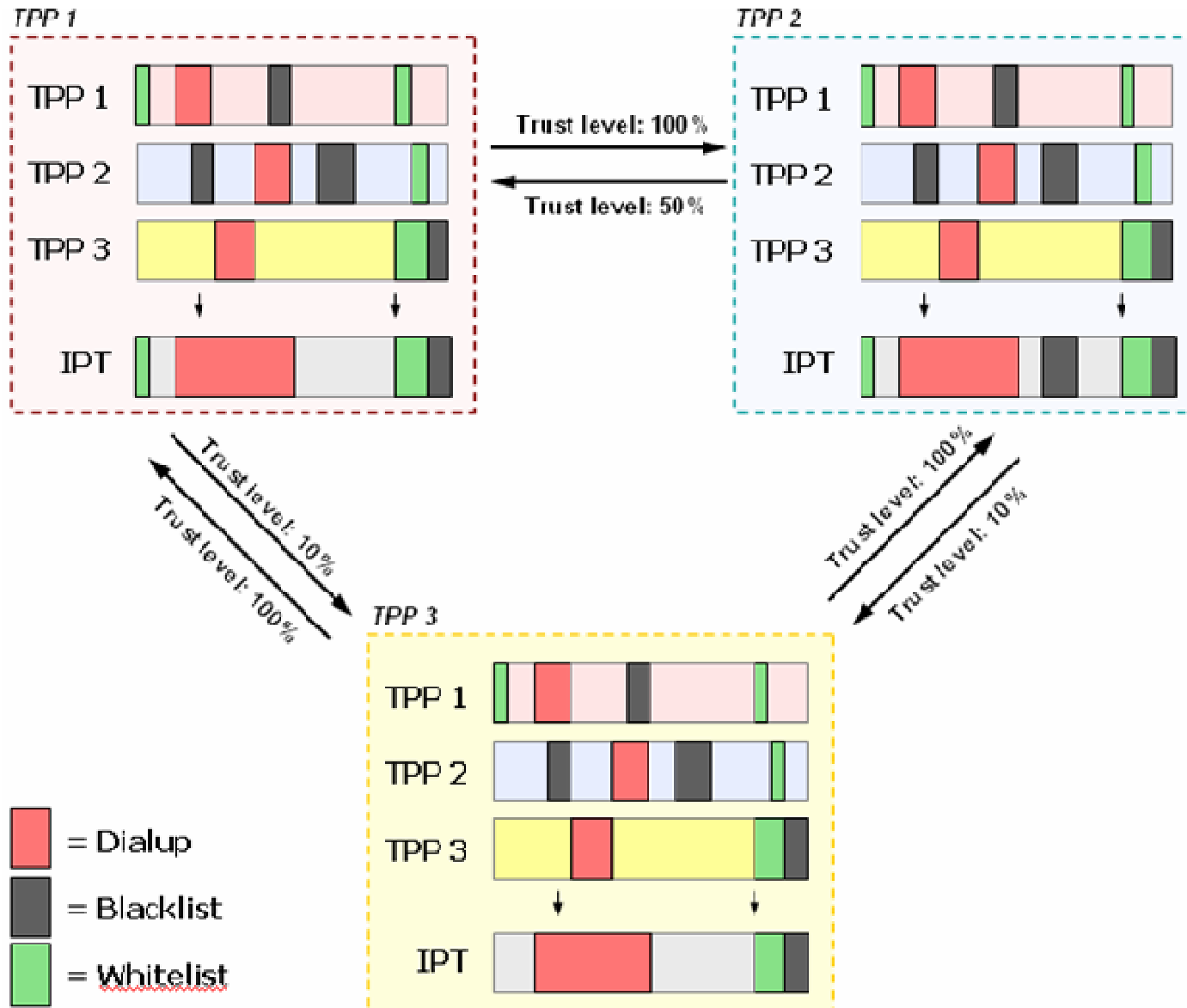
# Types of information

## → Describing attributes

- **Black list - attribute**
  - Describes IP addresses to be blocked permanently (i.e. > 1 week)
- **White list - attribute**
  - Describes legitimate outbound e-mail servers
  - Valid until withdrawn ...
- **Spam activity - attribute**
  - Short-dated (i.e. less than 1 week)
  - Reasons: unexpected high traffic, high bounce rate
  - Many spam activity statements from several nodes of the network may lead to blocking the IP address in question
- **Dialup - attribute**
  - IP address blocks used for dialup
- **Neutralization / Withdrawal**
  - Statements can be corrected/withdrawn

# Example: Distributed IP reputation

## → 3 partner with different trust level



# Distributed IP reputation system

## → Benefits

- As a **Semi-Closed User Group** the IP map is open to all ISPs
- Even with a small number of participants, a high amount of suspicious IP addresses are detected and thus a quick and effective identification of spam sources is in place.
- Self-declaration information mitigates the risk of false positives.
- Every participant is free concerning the use of data (not only to block spam, but also spam over internet telephony (spit)?)
- **Distributed IP reputation system:**
  - **no central point of failure** prevent from misuse by single participants
  - **enhances availability** of the distributed IP reputation system



# TCP/IP Layer

## → Frequency Analysis

- **Method:**  
Measure the number of connection attempts (TCP SYN packets) over time
- When a threshold is reached, let connection establish on TCP layer but greet with an SMTP error of category „non-fatal / temporary“ immediately
- **Examples:**  

```
Dec 6 19:23:12 pluto postfix/smtp[24897]: 52948354079:  
host mx-ha01.web.de[217.72.192.149] refused to talk to  
me: 421 mx19.web.de: Too many concurrent SMTP  
connections; please try again later  
  
...  
Dec 6 19:23:18 pluto postfix/smtp[24897]: 52948354079:  
to=[xxxx@web.de], relay=mx-ha02.web.de[217.72.192.188],  
delay=7, status=deferred (host mx-  
ha02.web.de[217.72.192.188] refused to talk to me: 421  
mx05.web.de: Too much load; please try again later)
```

# Checks on the SMTP Layer

---

- SMTP has lots of details
- SMTP could be picky
- Spam software does not always follow the RFCs

# SMTP Dialog

## → Example 1 (1/2)

{The client established the TCP connection}

**S: Server (Recipient)**

**C: Client (Sender)**

S: 220 smtp.example.com ESMTP Postfix →

← C: HELO relay.example.org

S: 250 Hello relay.example.org, I am glad to meet you →

← C: MAIL FROM:<bob@example.org>

S: 250 Ok →

← C: RCPT TO:<alice@example.com>

S: 250 Ok →

← C: RCPT TO:<theboss@example.com>

S: 250 Ok →

← C: DATA (*Command*)

# SMTP Dialog

## → Example 1 (2/2)

**S: Server (Recipient)**

**C: Client (Sender)**

**S:** 354 End data with <CR><LF>.<CR><LF> →

← **C:** From: "Bob Example" <bob@example.org>

**C:** To: Alice Example <alice@example.com>

**C:** Cc: theboss@example.com

**C:** Date: Tue, 15 Jan 2008 16:02:43 -0500

**C:** Subject: Test message

**C:**

**C:** Hello Alice.

**C:** This is a test message with 5 headers and 4  
lines in the body.

**C:** Your friend,

**C:** Bob

**C:** .

**S:** 250 Ok: queued as 12345 →

← **C:** QUIT

**S:** 221 Bye →

{The server (and client) closes the TCP connection}

# SMTP Dialog

## → Example 2

Sending MTA	Receiving MTA	
	220 foo.com Ready	} Email envelope
HELO bar.com		
MAIL FROM:<xx@bar.com>	250 foo.com Hello...	
	250 OK	
RCPT TO:<yy@foo.com>		} Email DATA
	250 OK	
DATA		
From: <xx@bar.com> To: <yy@foo.com> Subject: my email Date: Fri, 13 May 2007 13:33:37 +01:00 This is a text! .	354 Start mail input	
	250 OK	
QUIT		
	221 Good bye.	

# Checks on SMTP Level

## → HELO

- The first client command in an SMTP session is HELO (or EHLO)
- The argument should be the hostname of the client (RFC)
- If the client specifies a wrong hostname, it violates RFC2821
- > 60% of client software transmits wrong HELO argument!
- **→ Very few providers filter on HELO argument**  
Problem: NAT-masqueraded hosts (hostname must be resolved from the outside)

# SMTP Details

- ESMTP Pipelining is an extension to pipeline commands
- To conform to the RFC an implementation **MUST** wait after HELO/EHLO and DATA (spam-software – if using Pipelining at all – generally does not do so)
- Pre-Greeting Traffic
- Transport Layer Security (TLS) – usually not used by spammers
- Empty Mail-From (according to RFC only for bounces allowed)

```
Nicht autorisierte Antwort:
web.de MX preference = 110, mail exchanger = mx-ha02.web.de
web.de MX preference = 100, mail exchanger = mx-ha01.web.de
```

- Special filtering on Secondary MX
- Very short timeouts! (< 10 seconds)
- Greylisting (following slides)

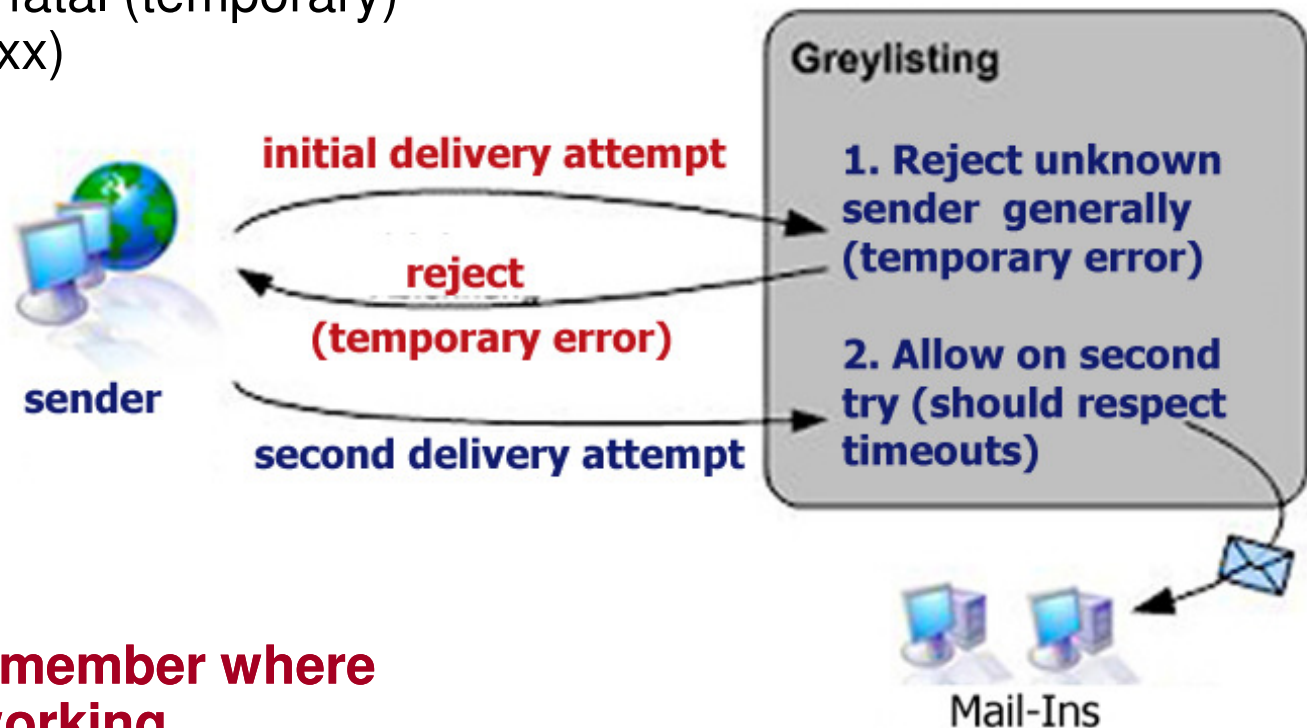
```
chris@leo ~
$ nc mx-ha01.web.de 25
220 WEB.DE
550 Connection timed out.

chris@leo ~
$
```

# Greylisting

## → Idea

- SMTP defines a non-fatal (temporary) error (status codes 4xx)
- Originally used for short-term problems on the receiving side
- Greylisting uses this „error“ in order to force a second delivery attempt
- **Core idea:**  
**The sender must remember where e-mail (i.e. have a working queueing mechanism) was sent and respect timeouts (according to RFC)**
- **A legitimate e-mail sender does more effort to send e-mail than a spammer**

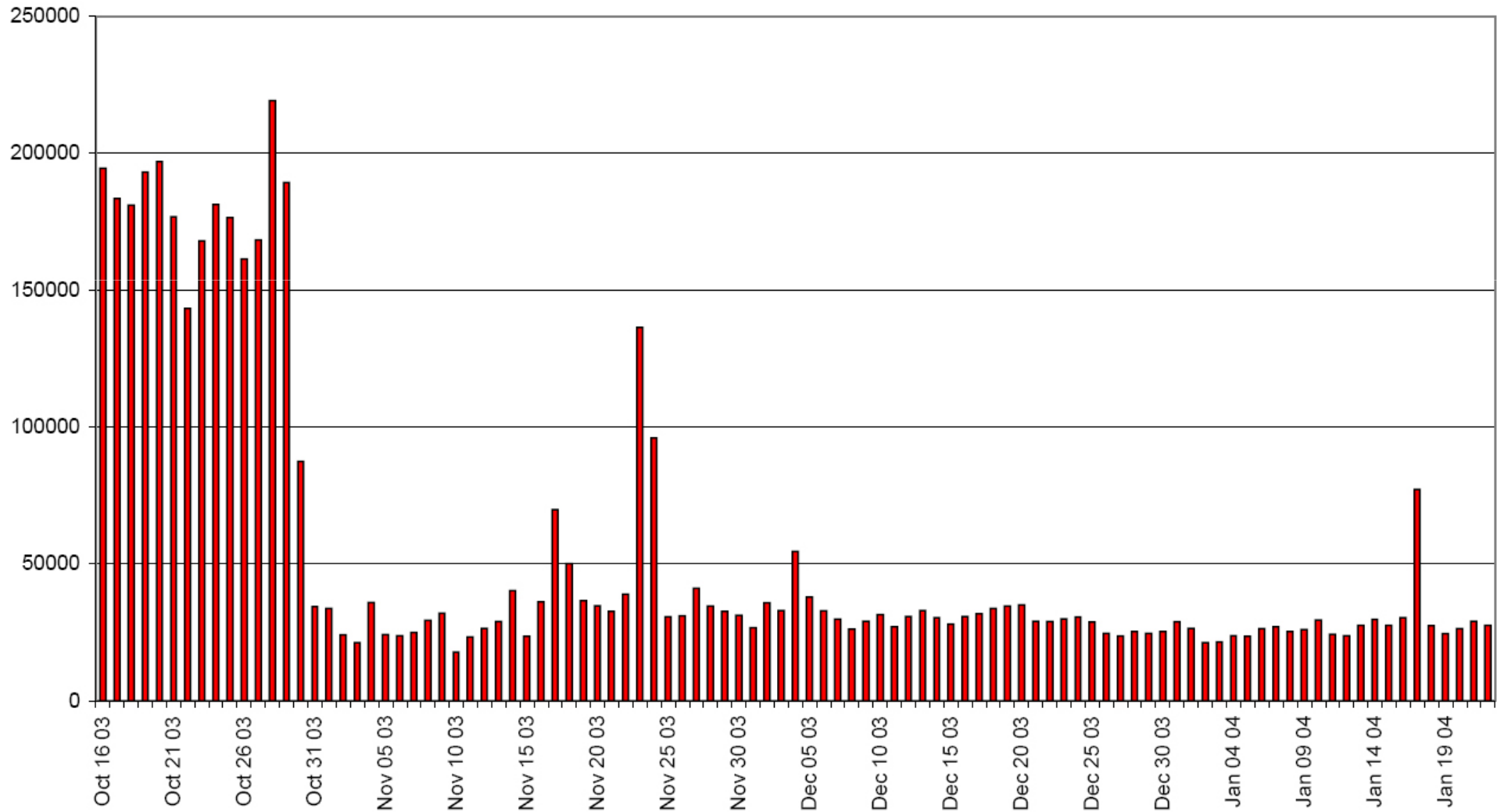




# Greylisting

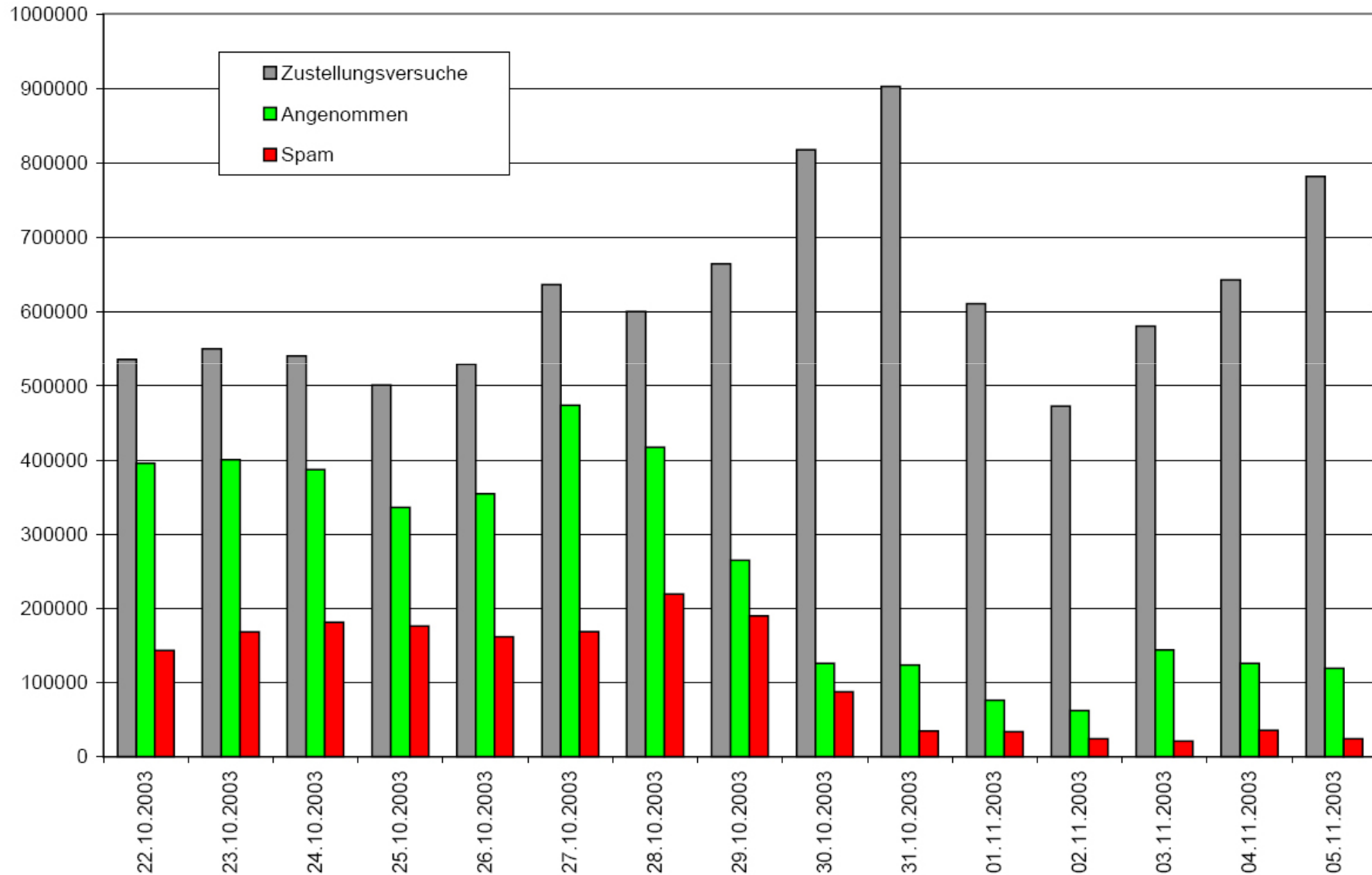
## → Some Figures (1/2)

### Daily spam volume



# Greylisting

## → Some Figures (2/2)



# Greylisting

## → Review



High benefit with very small effort, because many spammers often do not try a second time.



Spammers might adopt their methods if greylisting becomes more regular.



Might interfere legitimate email traffic, if cookie does not exist before sending an email to a server using greylisting.

# Sender Authentication

## → Idea

- E-mail was originally based on a network of confident (and trustworthy) participants.
  - **Lack of (sender) authentication**
- Dedicated MARID (MTA Authorization Records in DNS) working group formed in 2004
- 2 kinds of sender authentication proposals exist
  - **Path-based sender authentication (SPF, Sender ID)**
  - **Signature-based sender authentication (DKIM)**

# Path-based Sender Authentication

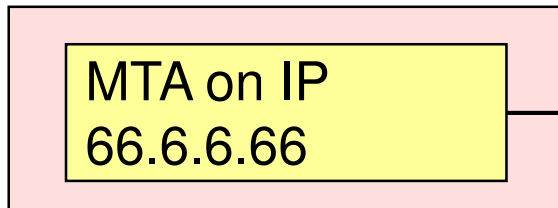
## → Idea

- Path-based authentication methods (like Sender ID and SPF) can be used to test if an e-mail server is authorized to send on behalf of a given domain.
- By using DNS, a receiver can check if the sending MTA is allowed to send in the name of a given domain.
- **SPF (Sender Policy Framework)**
  - Checks envelope's „**MAIL FROM**“
- **Sender ID**
  - Checks e-mail headers **From**, **Sender**, **Resent-From** and **Resent-Sender** (tricky!)

# Path-based Sender Authentication

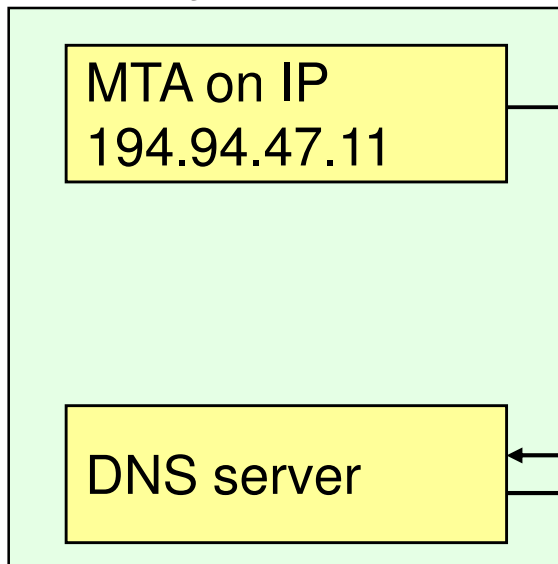
## → Example

### Disguising sender



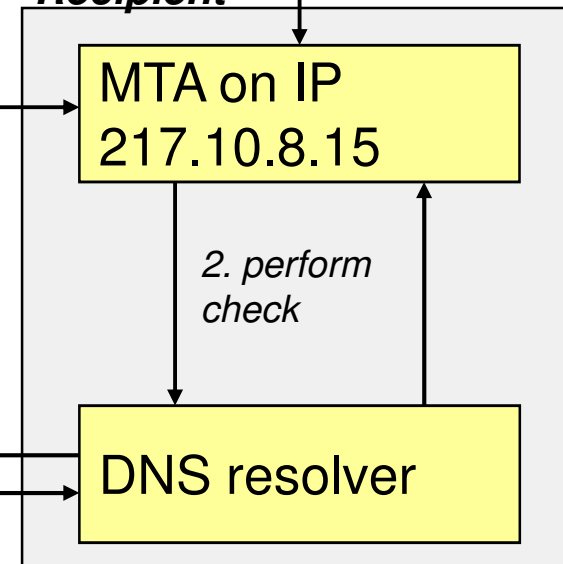
1b. MAIL FROM:  
paul@mydomain.com

### Authenticated sender (owns mydomain.com)



1. MAIL FROM:  
john@mydomain.com

### Recipient



2. perform  
check

3. DNS query for  
mydomain.com

4. DNS response:  
194.94.47.11  
203.10.80.3

**66.6.6.66** not in (194.94.47.11, 203.10.80.3)

# Path-based Sender Authentication

## → Real World Example (SPF)

- Example
  - E-mail sent from the domain **gmx.net**

```
"v=spf1 ip4:213.165.64.0/23 -all"
```

- SPF version 1 (v=spf1)
- Hosts with IP addresses of the range **213.165.64.0/23** are allowed to send e-mail on behalf of **gmx.net**...
- ...and nobody else! (-all)
- Mail.gmx.net with its IPs **213.165.64.20** and **213.165.64.21** (the Mail-Outs for gmx.net) actually fall in **213.165.64.0/23**

```
chris@leo ~
$ nslookup -q=txt gmx.net 217.237.151.225
Server: www-proxy.D01.srv.t-online.de
Address: 217.237.151.225

Nicht autorisierte Antwort:
gmx.net text =

        "v=spf1 ip4:213.165.64.0/23 -all"

chris@leo ~
$ nslookup mail.gmx.net
Server: icarus.home
Address: 192.168.28.2

Nicht autorisierte Antwort:
Name: mail.gmx.net
Addresses: 213.165.64.21, 213.165.64.20
```

# Path-based Sender Authentication

## → Review



Easily manageable for senders using existing technology, since only a DNS record has to be published.



Receivers have to adopt new software in order to check the DNS records.



Email forwarding services are generally prohibited when using this method. Two possibilities to manage these exist: Whitelisting of specific senders or rewriting of the senders email addresses (like Sender Rewriting Scheme<sup>32</sup> for SPF).



If two or more domains are running on the same IP address, an email sender from this IP can use every of those domains to send authenticated emails.



Prone to domain tasting, because spammers could use domains with legitimate DNS records.



# Signature-based Sender Authentication

## → Idea

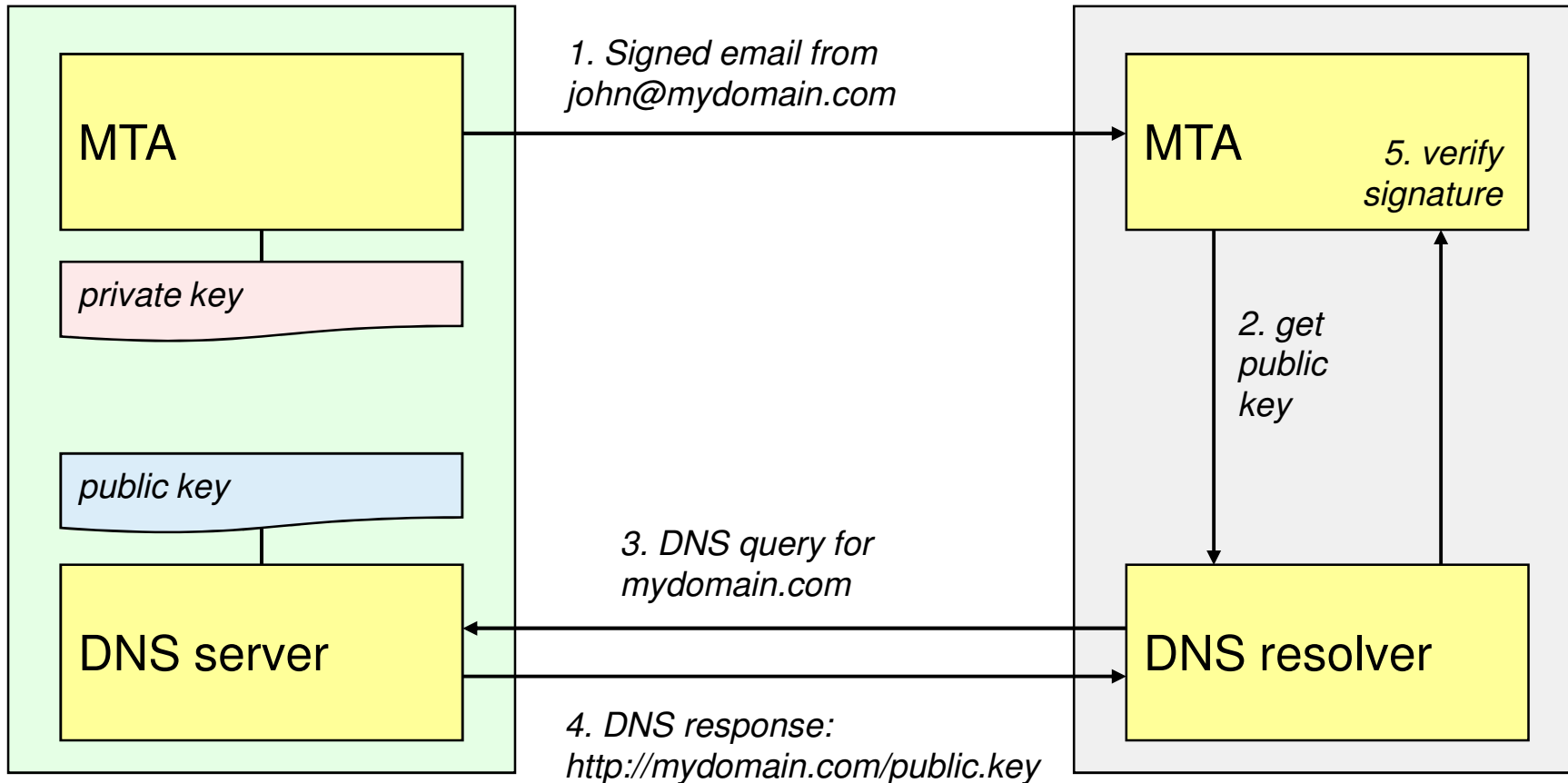
- Signature-based sender authentication methods (like DKIM) make it possible for a recipient to verify if a message actually comes from the domain it claims by the help of a digital signature / public key cryptography.
  - The sender's MTA signs the e-mail
  - The recipient MTA verify the signature
- **Notice:**  
PGP/SMIME are not original signature-based sender authentication methods, but could be considered as such (because they as well authenticate the sender)

# Signature-based Sender Authentication

## → Example

**Authenticated sender**  
(owns mydomain.com)

**Recipient**



# Signature-based Sender Authentication

## → Review



Email forwarding is possible, without changing the sender's address.



Differentiation between several domains hosted with a single IP address is possible.



The modification of emails is no longer possible, i.e. some software implementation must prevent this (e.g. mailing lists that put unsubscribe information to the end of every email).



Senders as well as receivers have to implement new technologies to sign emails and/or to check these signatures



Prone to domain tasting, because spammers could use domains with legitimate DNS records.

# Sender Address Verification (SAV)

## → Idea

- SAV (*sender callouts* or *callout verification*) can be used to check if an e-mail address exists or not
  - Spammers often use invented and/or automatically generated e-mail addresses that usually do not exist
  - SAV helps only to verify if the sender's given e-mail address exists.
  - On the other hand it **does not help to verify** if the sender is **authorized to use** this specific e-mail address or domain.
- Technically the
  - **receiving MTA** performs SAV with the
    - given sender address
    - during the SMTP dialog
    - with the sending MTA.
- In order to do so the receiving MTA establishes an SMTP dialog to the MTA accepting e-mails for the domain stated in the sender address and tries to deliver a **bounce message** to this address.

# Sender Address Verification (SAV)

## → Example

Sender MTA	Receiving MTA	Sender domain MTA	
	220 foo.com Ready	} <b>Main dialog</b>	
HELO bar.com			
MAIL FROM:<xx@bar.com>	250 foo.com Hello...		
	250 OK		
RCPT TO:<yy@foo.com>	<b>451 not yet verified</b>		
} <b>Sender address verification dialog</b>	HELO foo.com	220 bar.com Ready	
	MAIL FROM:<>	250 bar.com Hello...	
	RCPT TO:<xx@bar.com>	250 OK	
	QUIT	250 OK	
			} <b>Continuation of main dialog</b>
	RCPT TO:<yy@foo.com>	250 OK	
	(...)		

# Sender Address Verification (SAV)

## → Review



Spam with wrongly spelled sender's email addresses can be filtered.



Will every MTA reject emails during the SMTP dialog if an email address does not exist or do some MTAs accept emails and create bounce messages?



Loops may occur when not using "<>" as sender address within SAV dialog and both MTAs performing SAV.



Spammers will probably learn from SAV and adopt their sender addresses to existing addresses, undermining the full benefit of SAV.



SAV might lead to DoS attacks against the owner to the used domain for spamming. High spam volumes with a specific sender domain will lead to many SAV checks performed by multiple parties, breaking down this specific MTA.



High resource consumption due to an additional heavy SMTP dialog.

# Content-based Antispam Methods

---

- Heuristical / rule / header checks
- Checksum methods (DCC, Pyzor, Razor)
- Statistical filters (Bayes)

# Heuristic / Rule / Header checks

## → Idea

- Properties of spam messages are used as filter criteria
- **Examples:**
  - A message carries HTML content
  - Spoofed UserAgent String (MS Outlook Ver 6.10.51214.1241 – was never shipped by Microsoft!)
  - A message consists only of images
  - Puns (BIG, u\_n\_d\_e\_r\_s\_c\_o\_r\_e, gappy)
  - Random strings (sidufwbw skfjhawer), in order to evade checksum methods
  - Date and time checks
- Heuristic methods usually aim at finding specific words, regular expressions or misuse related styles in e-mails to classify them as either spam or ham



# Heuristic / Rule / Header checks → Review



The method does not need a training phase.



Heuristical analyses are very efficient with a well-managed policy database.



A leak of performance might occur with at high mail volume or huge policy databases.



Managing the policy list is very time-consuming.

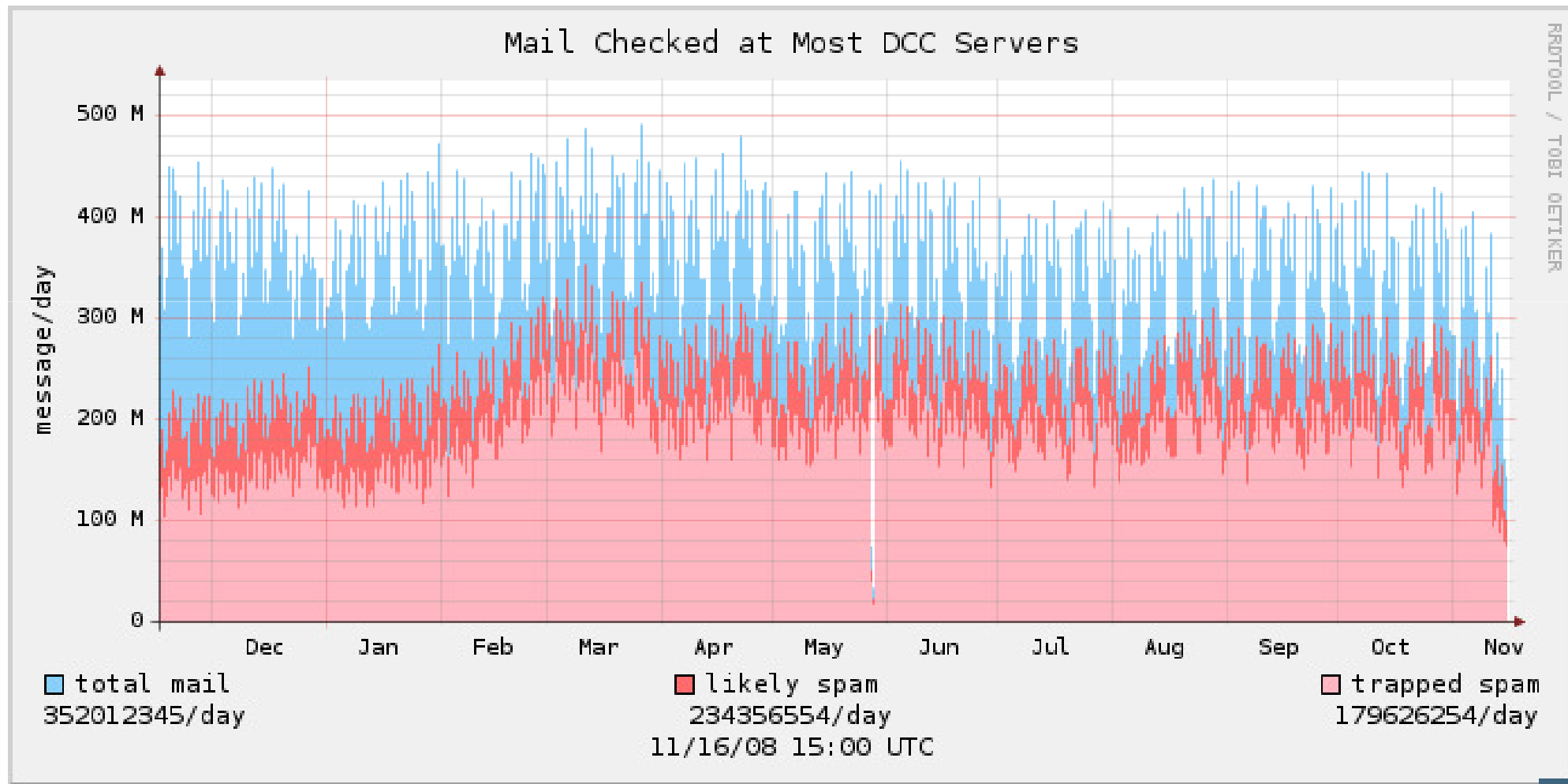
# Checksum Methods

## → Idea

- The same spam e-mail occurs in multiple different mailboxes, but a single recipient cannot decide if this e-mail is being received by many other persons.
- The main idea of checksum comparison is to share **unique fingerprints about received e-mails**.
- This allows each member of this community to know **if and how often a certain message was received** by other users.
- The unique fingerprints are not the e-mail itself (because of data privacy and the high amount of data) but are generated by a checksum algorithm.
- Mostly these algorithms are **fuzzy checksums** (or locality sensitive hash functions) to avoid a modification of the hash value if only little modifications on the text have been made.
- Most checksum functions are proprietary, only very few open (such as nilsimsa)
- Distributed Checksum Clearinghouse, Razor, Pyzor

# Distributed Checksum Clearinghouse **if(is)** internet security.

© Prof. Dr. Norbert Pohlmann, Institute for Internet Security - if(is), University of Applied Sciences Gelsenkirchen, Germany



# Checksum Methods

## → Review



Automated database filling possible via spamtraps possible.



Users can help to apply this anti-spam method, but classifying legal bulk email as spam might lead to high false positive rates



Algorithms are either proprietary (not comparable to public databases) or public (spammers can test them before modifying the mails).

# Statistical Methods

## → Idea

- A statistical filter automatically splits e-mails into several **tokens** (e.g. words) and looks these tokens up in a database.
- The database contains common tokens with a **classification** whether or not it is a common token in spam e-mails.
- This requires a **training phase** of statistical methods, where lots of messages must be classified as spam/ham in order to build up the database.







- **Example:**

- Bayes-filter in Mozilla Thunderbird  
(free e-mail client)

Subject*FREE	0.9999
free!!	0.9999
To*free	0.9998
Subject*free	0.9782
free!	0.9199
Free	0.9198
Url*free	0.9091
FREE	0.8747
From*free	0.7636
free	0.6546

# Statistical Methods

## → Review

-  Automated method, i.e. after training phase usually no manually work needed.
-  Low false positive rate for text spam.
-  Training phase required, but at the level of providers almost negligible.
-  Bayesian poisoning might mislead the Bayesian filters.
-  Image spam has to be translated into text before analysis is possible.
-  Leak of performance at high mail volume.

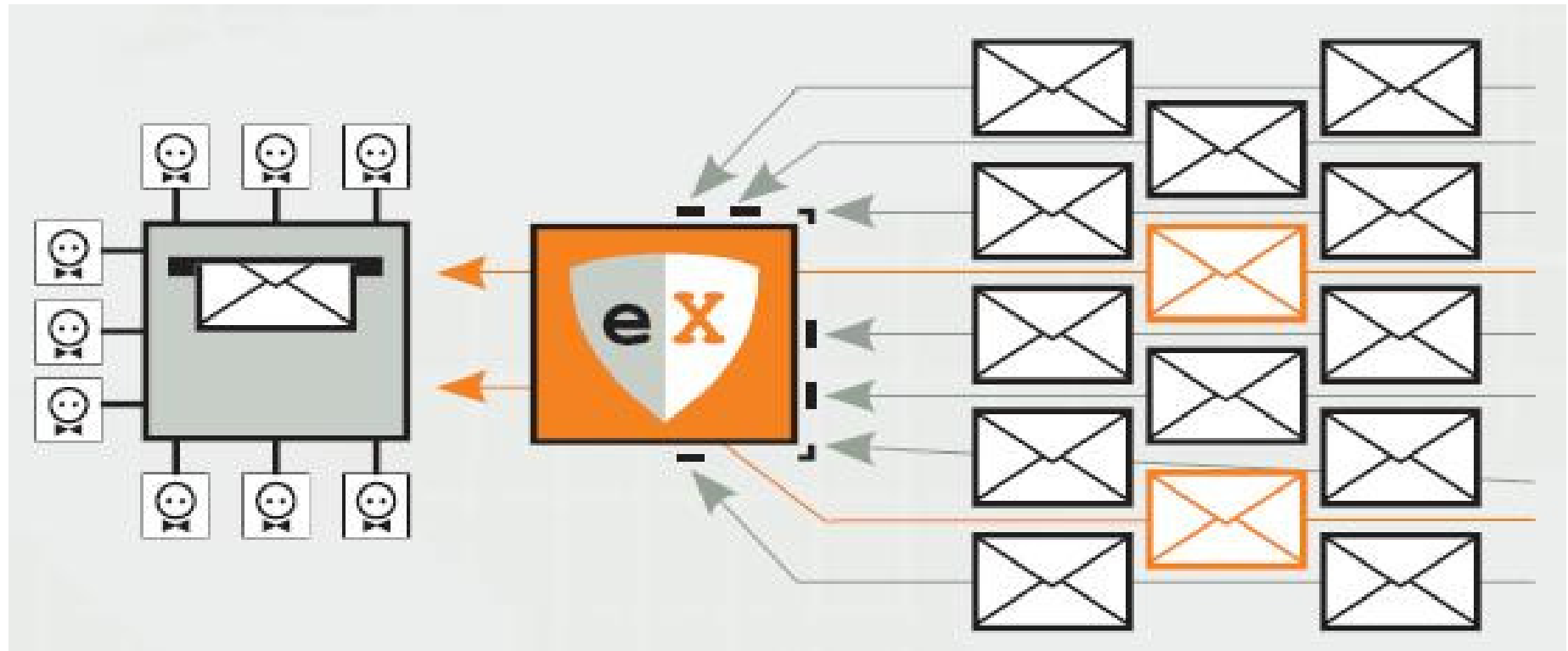
# E-Mail Firewall

## → Idea

- Offer protection against:
  - Great spam peaks
  - Target spam attacks
  - Denial-of-Service attacks
  - Mail bombs
  - Mail loops
  - ...
- Identifying the relevant business partner

# E-Mail Firewall

## → Idea



- Preferential of know sender's
- Rejection of unknown E-Mail Server in the peak phase
- ...

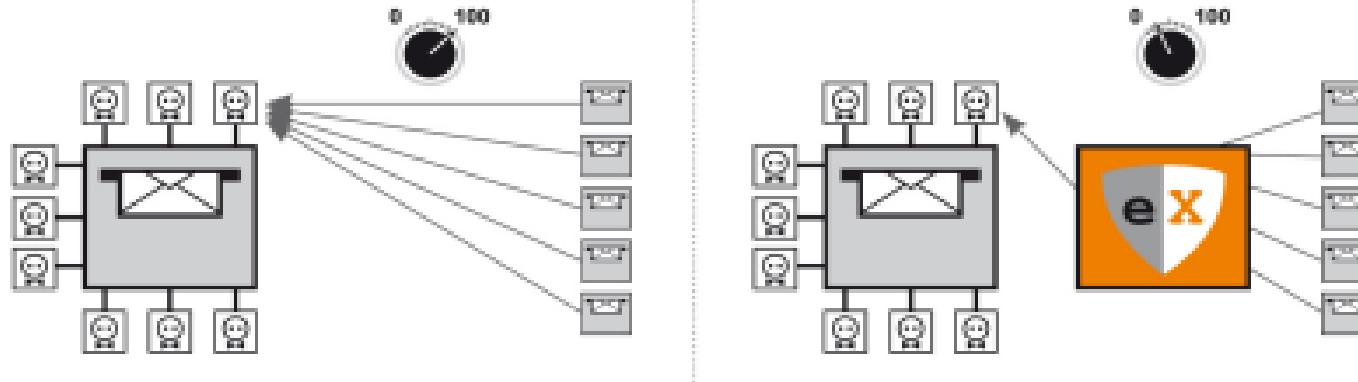


# E-Mail Firewall

## → Mail Bomb Protection

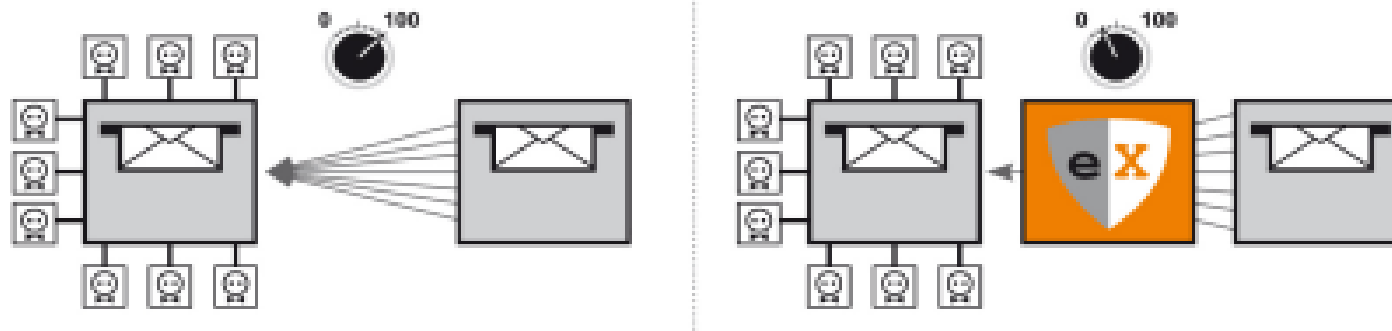
- **Receiver orientation**

- Limitation of the number of e-mails for one receiver.



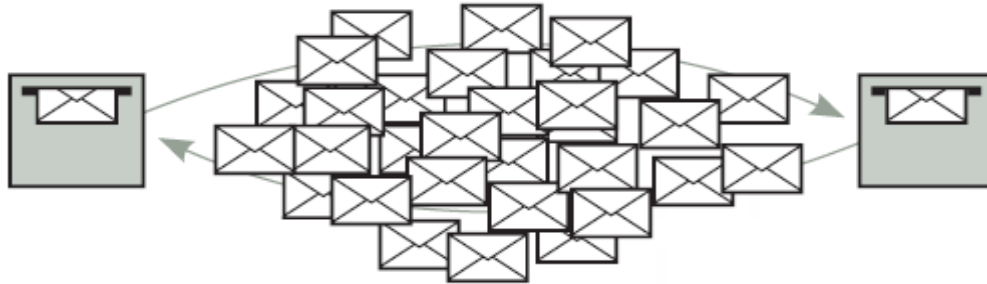
- **Sender orientation**

- Limitation of the number of e-mails which comes from one sender.

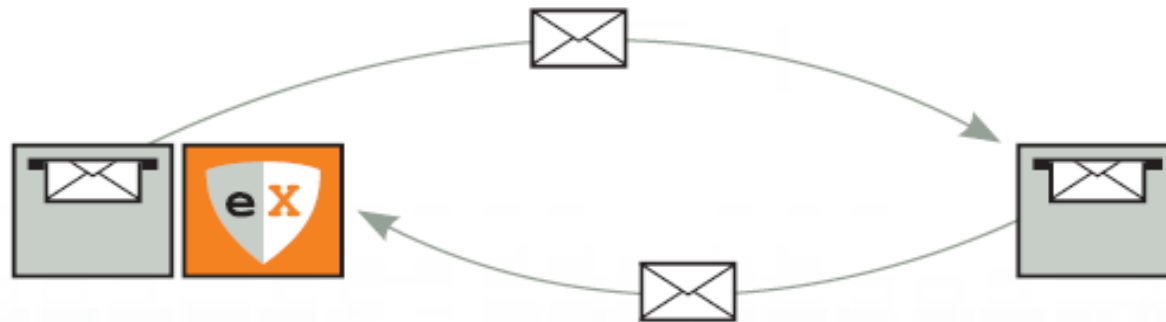


# E-Mail Firewall

## → Mail Loop Protection



- Adjust the e-mail traffic between one internal e-mail account and one sender.



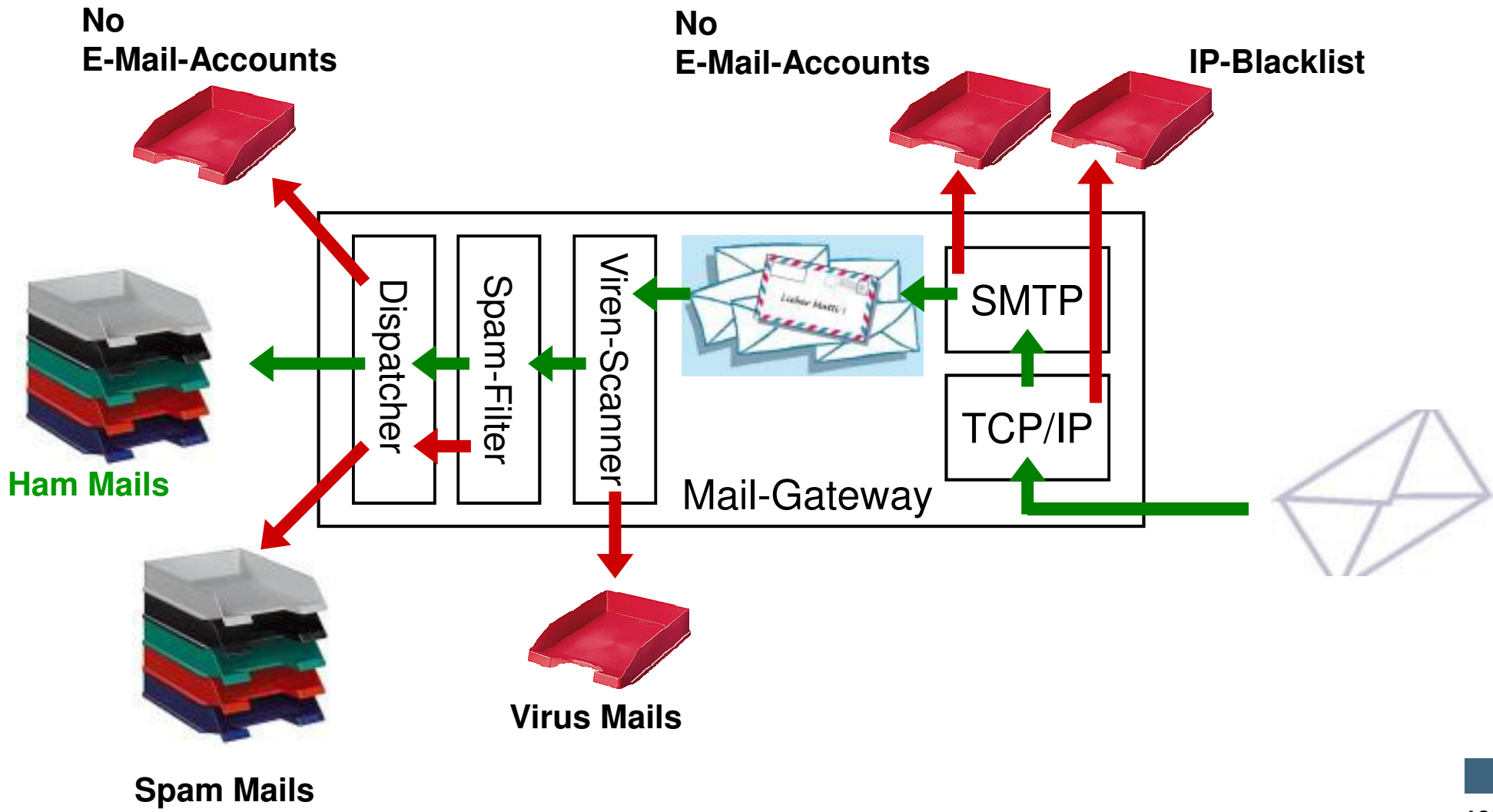
- Therefore it is possible to identify and to slow down mail loops.

# Content

- Aims and outcomes of this lecture
- Terms, Definitions and Damages
- E-Mail Infrastructure
- Sources of Spam
- Anti Spam Techniques
- **Generalized View  
(Survey, Estimation)**
- What can the different Stakeholders do
- Summary

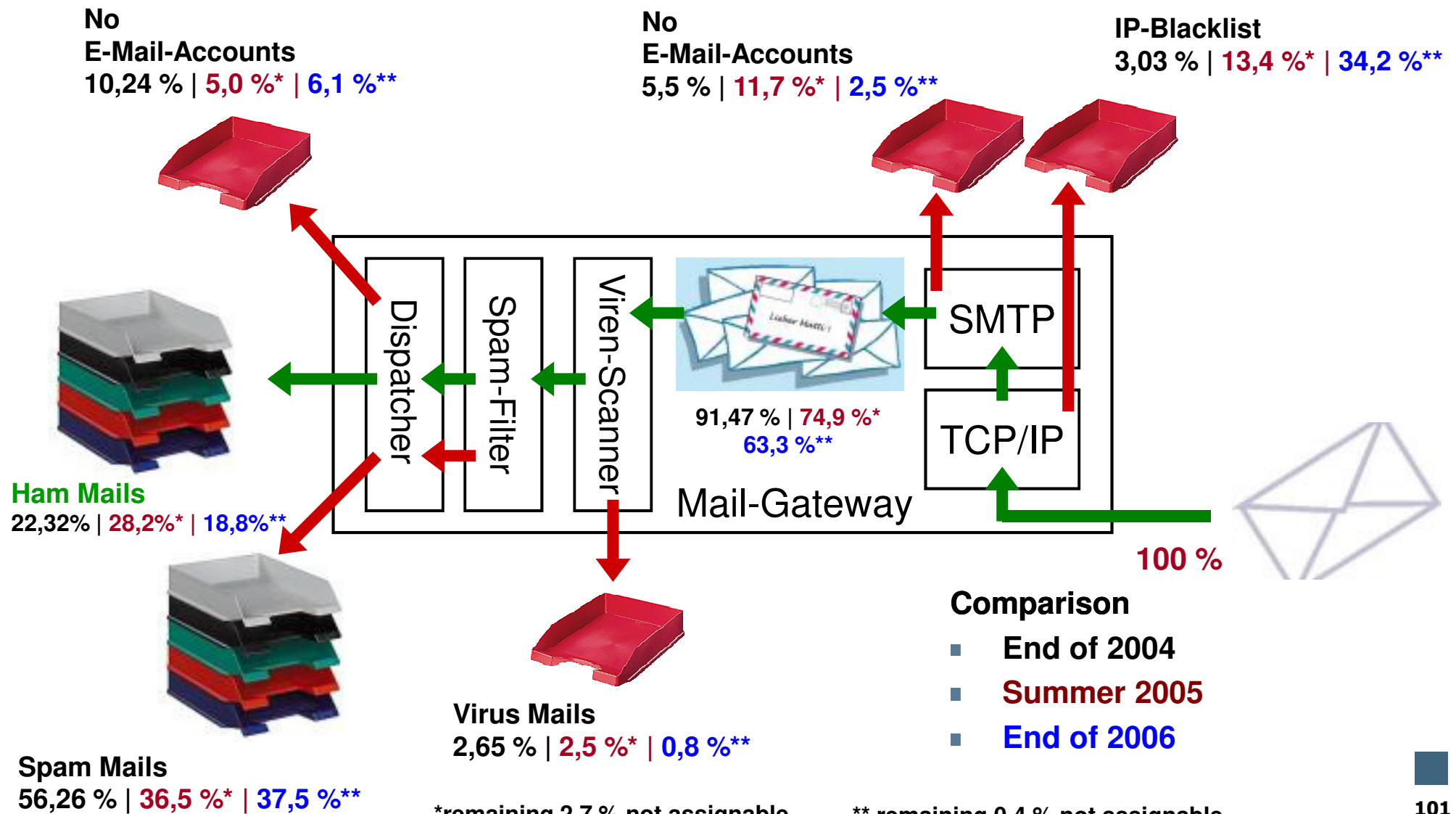
# Generalized View

## → Overall Measurements



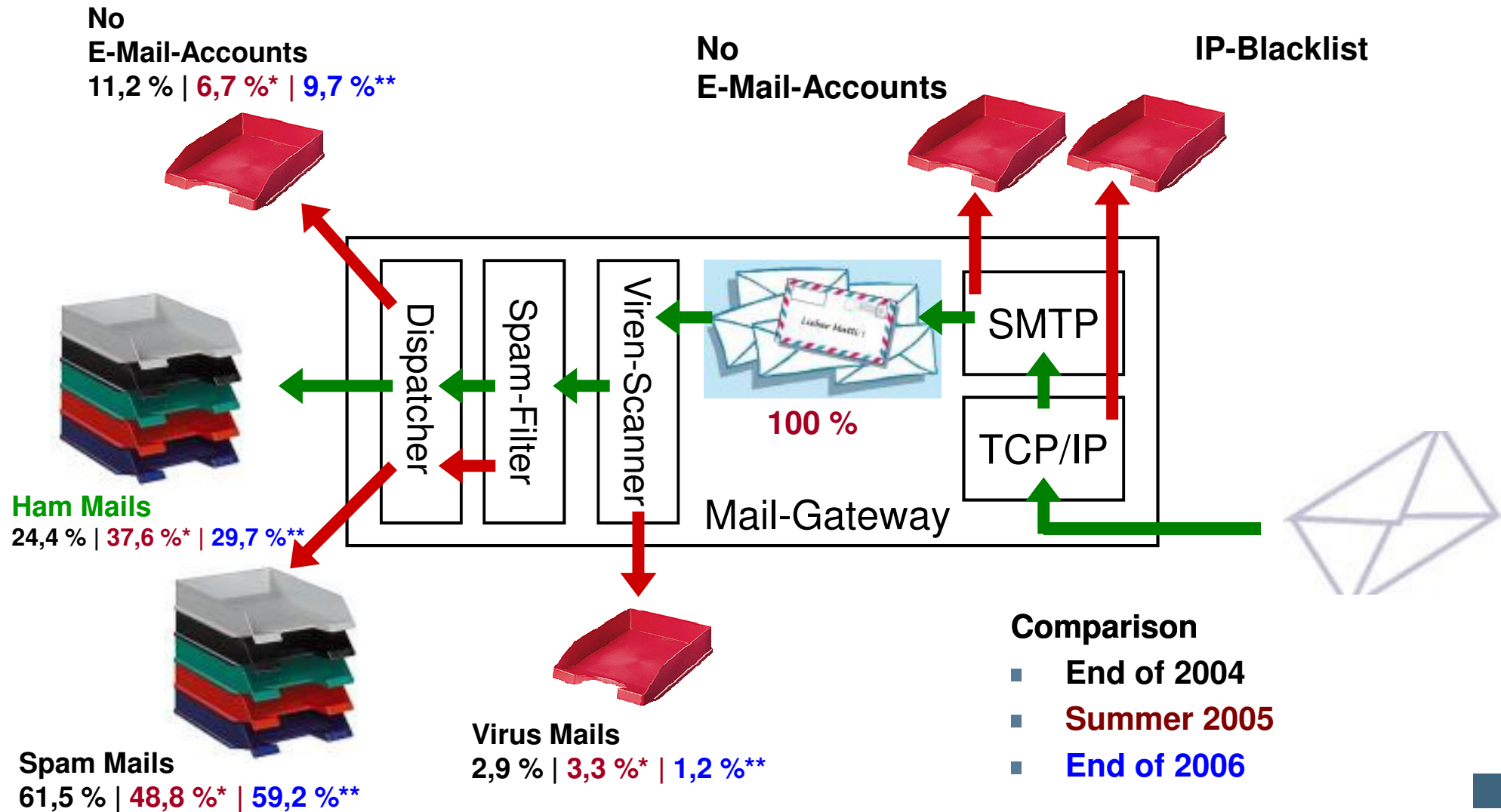
# Generalized View

## → Results: System, Delivery



# Generalized View

## → Results: System, Accepted



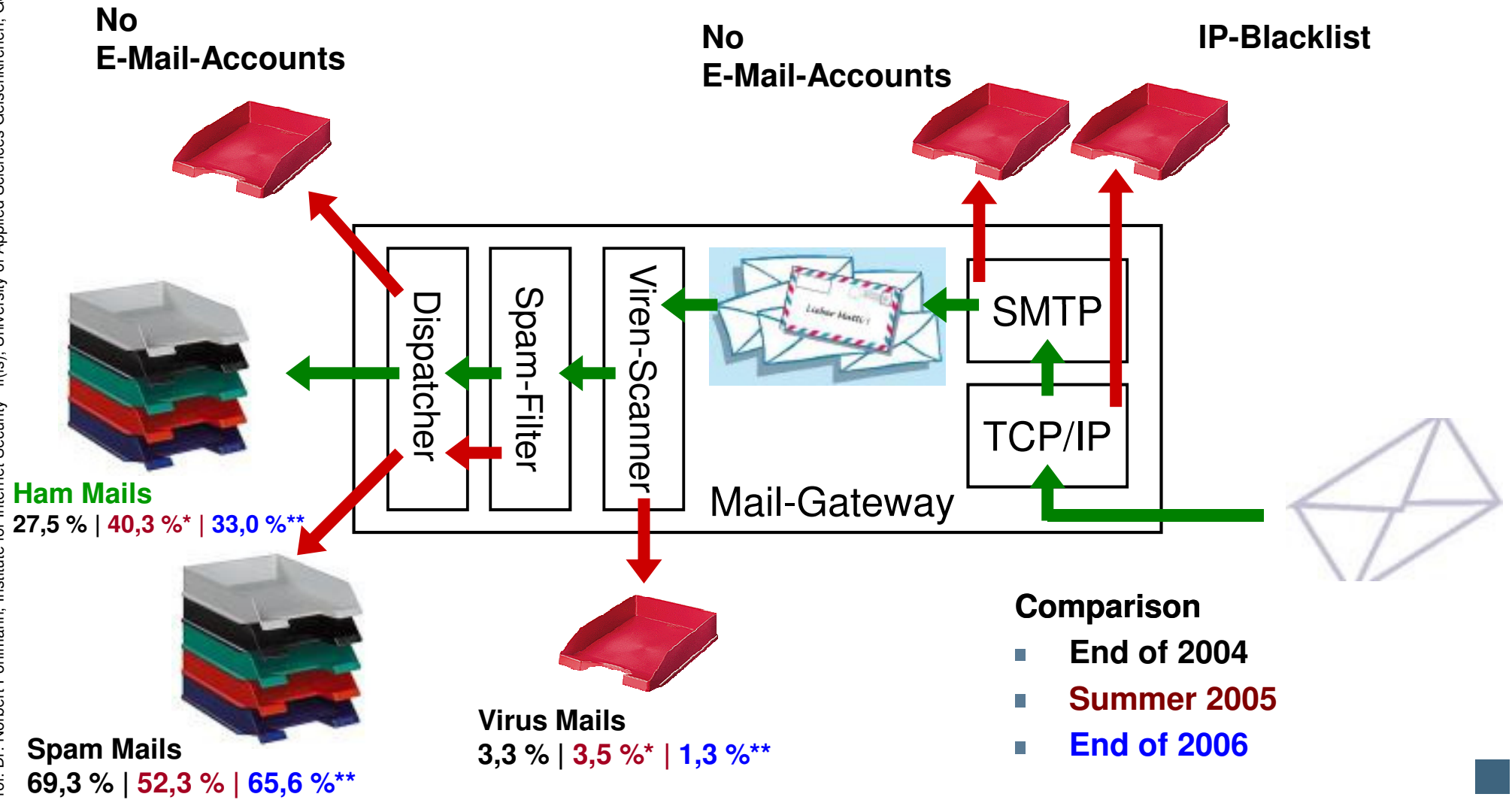
\*remaining 3,6 % not assignable

\*\* remaining 0,2 % not assignable

# Generalized View

## → Results: User Point of View

© Prof. Dr. Norbert Pohlmann, Institute for Internet Security - if(is), University of Applied Sciences Gelsenkirchen, Germany

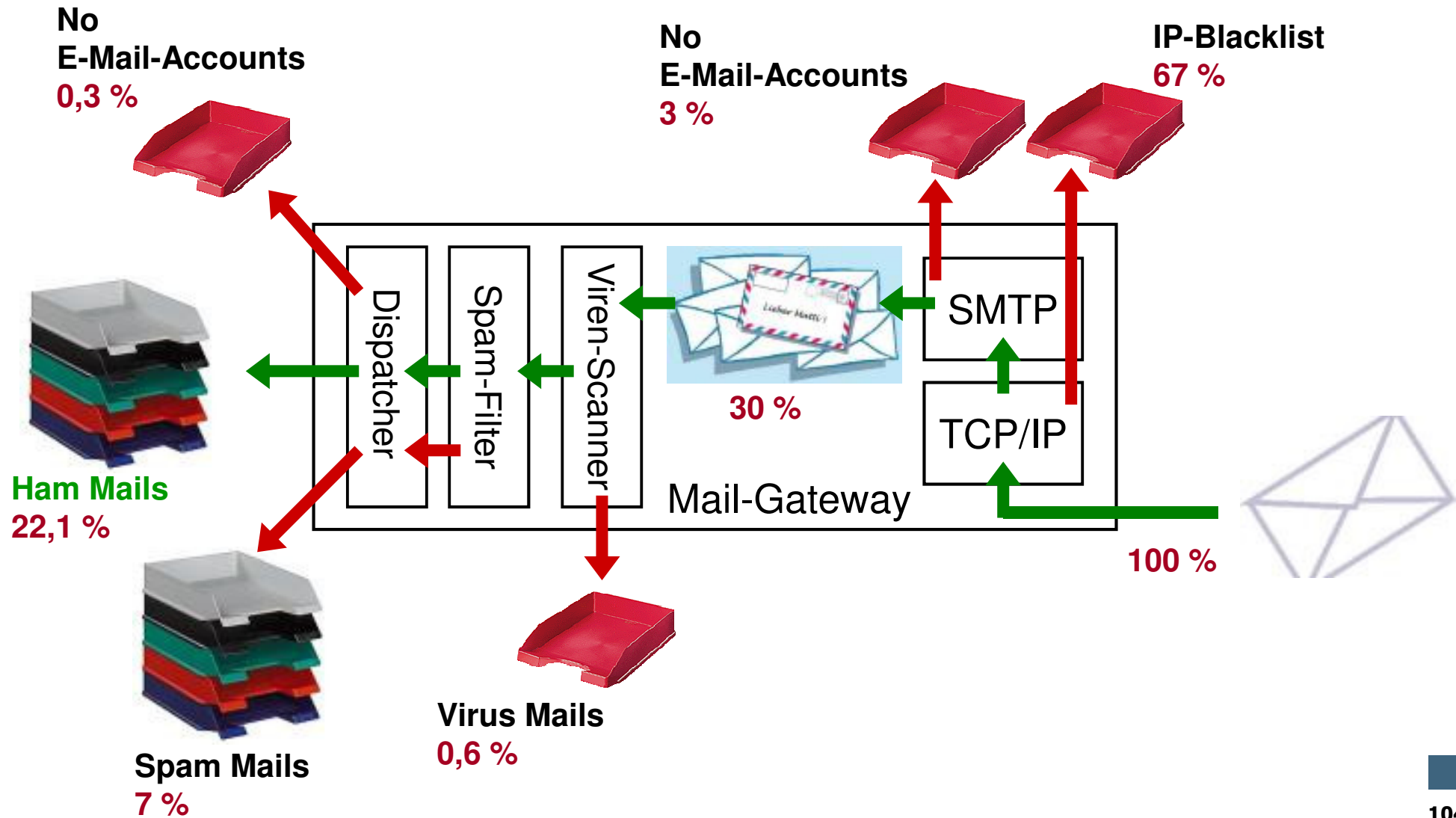


\*remaining 3,9 % not assignable

\*\* remaining 0,1 % not assignable

# Generalized View

## → Estimation: System, Delivery

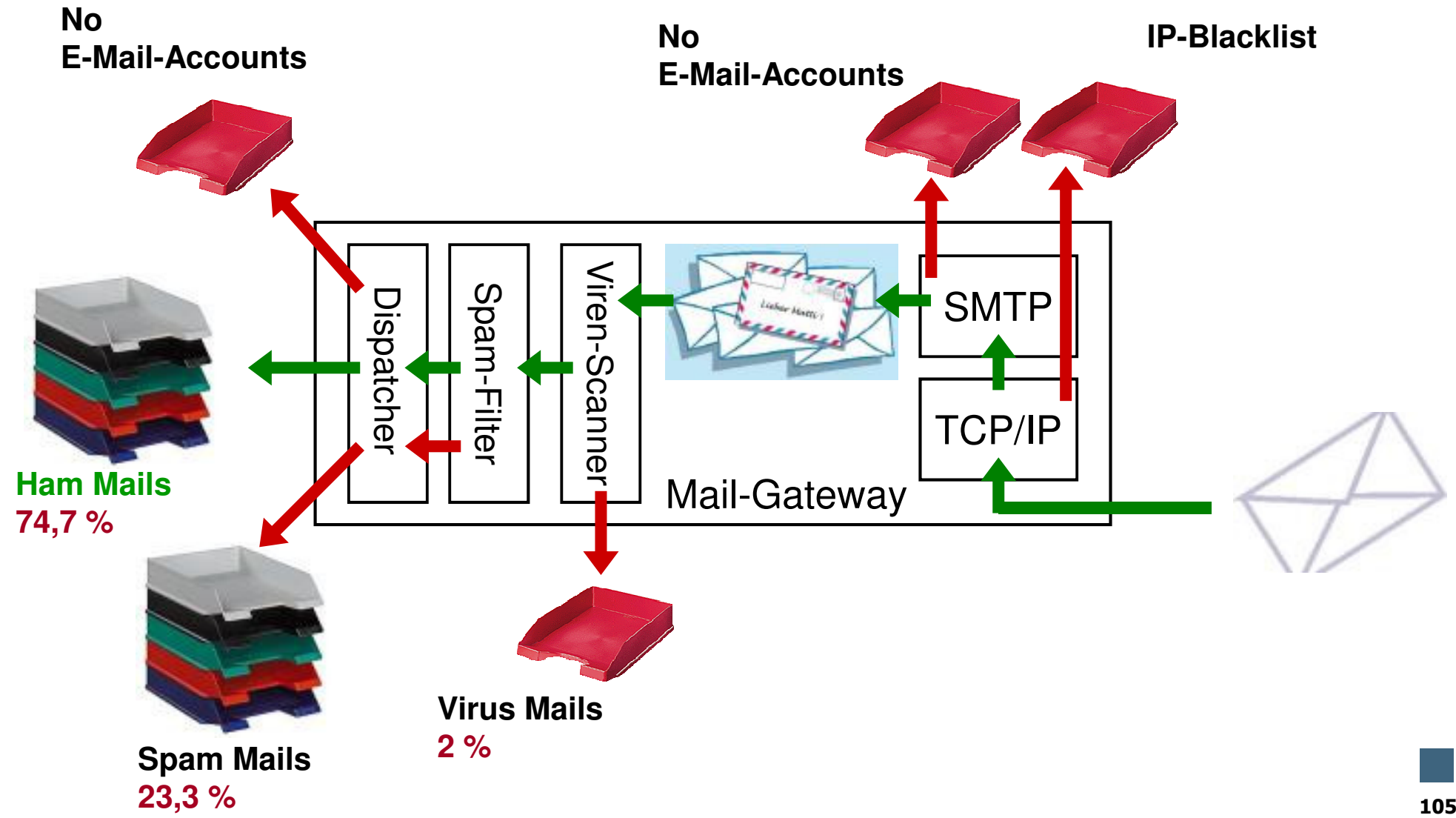




# Generalized View

## → Estimation: User Point of View

© Prof. Dr. Norbert Pohlmann, Institute for Internet Security - if(is), University of Applied Sciences Gelsenkirchen, Germany



# Content

- Aims and outcomes of this lecture
- Terms, Definitions and Damages
- E-Mail Infrastructure
- Sources of Spam
- Anti Spam Techniques
- Generalized View (Survey, Estimation)
- **What can the different Stakeholders do**
- Summary

# Different Stakeholders

## → Idea

- We have different stakeholders in the e-mail field:
  - Enterprises / User
  - E-Mail Service Provider
  - ISPs
  - Government

# Different Stakeholders

## → Enterprises / User

- Careful handling with the e-mail addresses
  - No e-mail addresses in clear text on the website
  - Never send more than 5 e-mail addresses with CC
  - ...
- Never follow links in e-mail of unknown senders!
  - Never follow e-mail advertisement (success for the spammer)
  - ...
- Using of Spam filter at your Mail-Gateway, at your PC
- Build up an appropriated e-mail infrastructure

# Different Stakeholders

## → E-Mail Service Provider

- Check e-mails on IP and SMTP Level
- Offer Spam filter as a service
- Share your Spam knowledge with others to improve the general Anti Spam results
- Build up an appropriated e-mail infrastructure
- Close user groups, like DE-Mail
- Partnerships, like eBay with United Internet (Web.de, GMX, ...)

# Different Stakeholders

## → ISPs

- Check e-mails on IP and SMTP Level
- Work together with other ISPs on an international level
  - e.g. distribution IP reputation system
- Build up an appropriated e-mail infrastructure
- Close port 25 for non e-mail gateways
- ...

# Different Stakeholders

## → Government

- Introduce the appropriate laws
- Enforcement of the laws!
  - Cooperation on an international level
  - ....
- Awareness programs

# Content

- Aims and outcomes of this lecture
- Terms, Definitions and Damages
- E-Mail Infrastructure
- Sources of Spam
- Anti Spam Techniques
- Generalized View (Survey, Estimation)
- What can the different Stakeholders do
- **Summary**



# Anti Spam

## → Summary

- Spam mail is a complex problem of the global internet.
- What can we see on the Spam market:
  - “Professionalization”
  - Easy market entry into spam business
- **The new concept of distributed IP reputation system**
  - makes black and white listing robust, trustworthy and manageable
  - will help to reduce the spam problem and prevent from damage
- **What is needed**
  - International cooperation will be very effective
  - **More analysis of communication behavior** of e-mail senders helps to detect and optimize reputation of IP addresses concerning e-mail

# Anti Spam

→ Introduction, Basis and Research

Thank you for your attention!  
Questions?

Prof. Dr.

**Norbert Pohlmann**

Institute for Internet Security - if(is)  
University of Applied Sciences Gelsenkirchen  
<http://www.internet-sicherheit.de>



if(is)  
internet security.

# Anti Spam

## → Literature

- [1] Christian J Dietrich, Norbert Pohlmann – E-Mail-Verlässlichkeit: Verbreitung und Evaluation, März 2005, Konferenzband DACH Security 2005
- [2] Christian J Dietrich, Norbert Pohlmann – Spam: Situation und Hintergründe, April 2004, Konferenzband BSI Kongress 2005
- [3] Christian J Dietrich, Norbert Pohlmann – IP Blacklisting zur effektiven Spam-Abwehr, September 2005, Datenschutz und Datensicherheit (DuD) 29, Ausgabe 09/2005, S. 548 ff.
- [4] Christian J Dietrich, Norbert Pohlmann – Spam auf dem Rückmarsch?, Oktober 2005, IT-Sicherheit, Ausgabe 04/2005

### Links:

Institute for Internet Security:

<http://www.internet-sicherheit.de/forschung/aktuelle-projekte/e-mail-verlsslichkeit/>

Dynamic blacklisting portal of the Institute for Internet Security

<http://dnsbl.if-is.net/>