Dead Letters to Alice - Reachability of E-Mail Addresses in the PGP Web of Trust

Benjamin Leiding [1] Andreas Dähn [2]

[1]University of Göttingen
Telematics Group
benjamin.leiding@cs.uni-goettingen.de

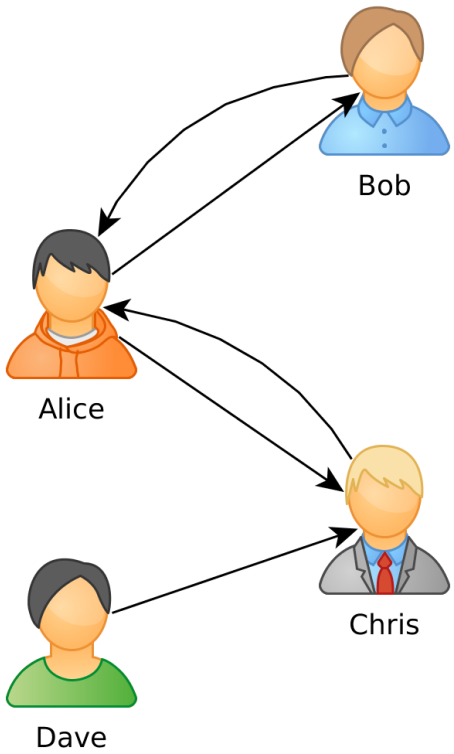
[2]University of Rostock
andreas.daehn2@uni-rostock.de

August 1, 2016

Overview

Introduction

PGP Web of Trust



- Decentralized trust model for public keys
- Mainly used for encrypted email communication
- Users sign other users' public key using their own private key to certify authenticity
- Interpreting each key as node and each signature as directed edge results in a directed graph → The PGP Web of Trust

Motivation

- Many different trust metric calculations applied to underlying graph of the Web of Trust
- Commonly used trust metrics exclude expired/revoked keys and signatures

BUT: Reachability of corresponding email account has not been considered as a criteria so far.

Study Design

Study Setup

1. Preparation
2. Syntax check of e-mail addresses
3. DNS testing
4. Validation
5. E-mail account testing

Preparation and Syntax Check

1. Preparation

- Keyring snapshot retrieved on October 22, 2014
- Extract e-mail addresses
- Remove duplicates

2. Syntax check of e-mail addresses

- Filtering for syntactical validity
- No .onion-addresses (and similar)

DNS Testing, Validation and E-Mail Account Testing

3. DNS testing

- Pull each domain's DNS record and extract mail exchange (MX) server

4. Validation

- Test MX servers' validation policy → 27d89e25a3518f4a7434474c2a7d4f1e43911bc58bec5f1@cia.gov

5. E-Mail account testing

- Actual testing of e-mail addresses

Study Duration

Main-Study

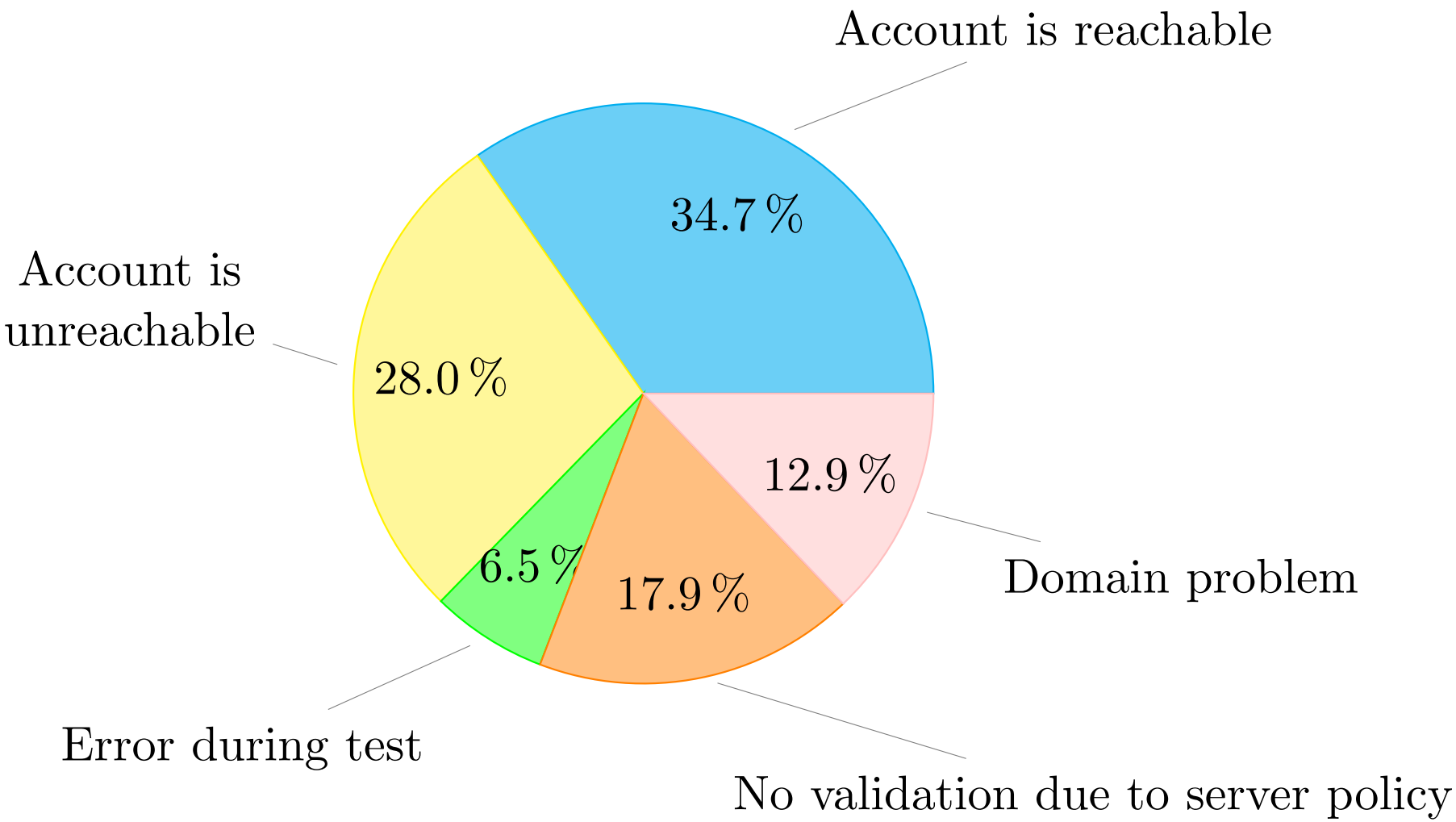
- Between February 12, 2015 and July 24, 2015

Sub-Study

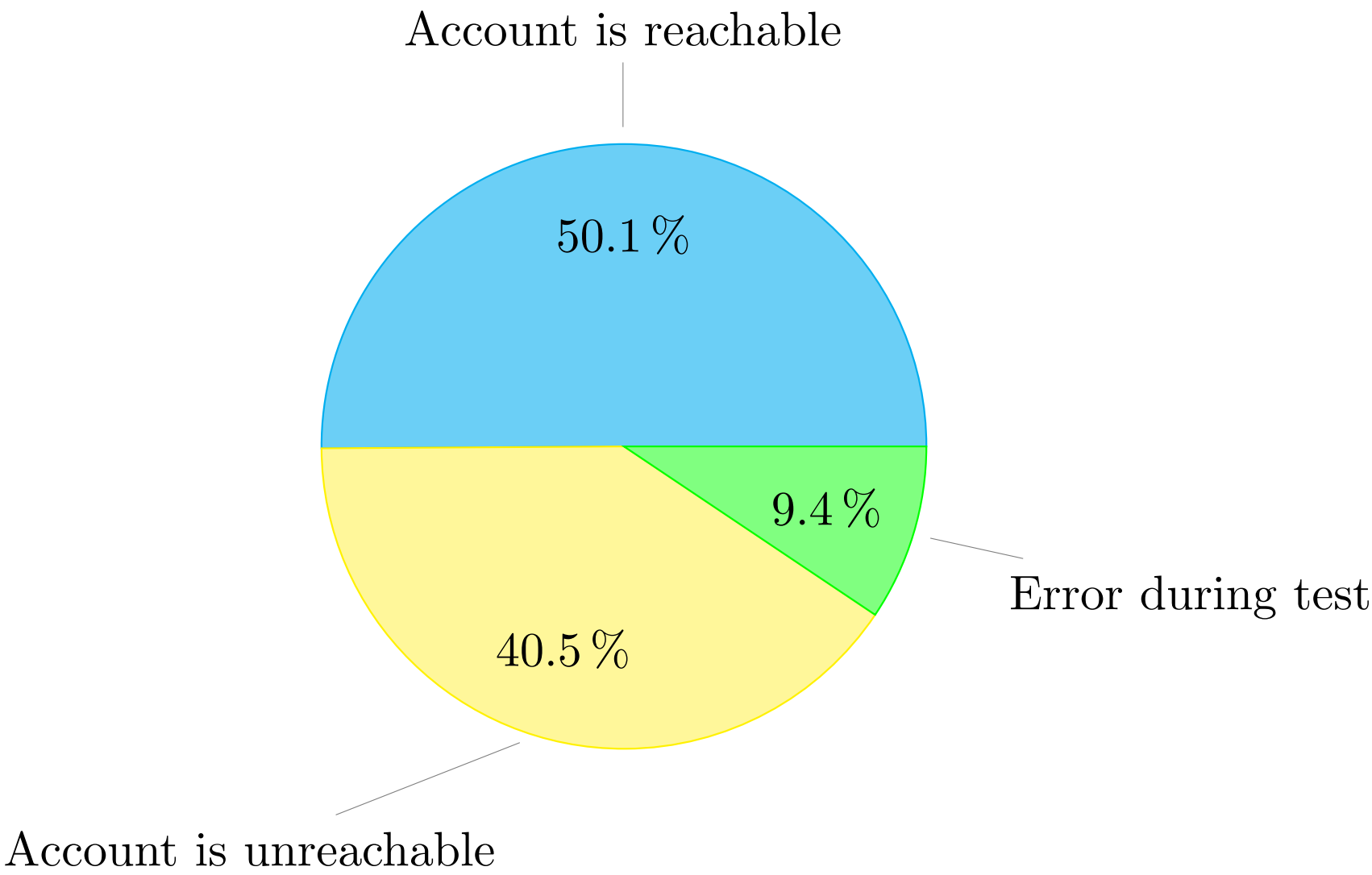
- Between August 11, 2015 and August 30, 2015
- Used a subset (1%) of the syntactical correct and unique addresses
- Almost same results (88%)

Results

Findings - Syntactical Correct Unique E-Mail Addresses



Findings - Server Allowed Validation



Top 5 Mail Exchange Domains

MX domain	Overall		Tested dead	Tested alive
	abs.	rel^{+}	abs.	abs
google.com	472,528	14.84 %	64,268	369,816
googlemail.com	142,350	4.47 %	44,401	64,599
hotmail.com	125,857	3.95 %	50,036	53,128
gmx.net	106,818	3.35 %	18,943	63,796
yahoodns.net	83,747	2.63 %	745	476

$^{+}$ to number of syntactic correct e-mail addresses the WoT

Selected Mail Exchange Operating Companies

Provider	Overall		Tested dead		Tested alive	
	abs.	rel^{+}	abs.	rel^{++}	abs	rel^{+++}
Google	614,878	19.31 %	108,669	12.18 %	343,415	31.11%
United Internet	216,999	6.82 %	49,569	5.56 %	115,479	10.46 %
Microsoft	183,104	5.75 %	54,076	6.06 %	57,360	5.20 %

$^{+}$ to number of syntactic correct e-mail addresses in PGP web of trust

$^{++}$ to number of e-mail addresses found unreachable in the study

$^{+++}$ to number of e-mail addresses found reachable in the study

Conclusion

Conclusion

Conclusion

- Extracted about four million e-mail addresses and tested three million of them
- 40% of the e-mail addresses are unreachable
- 46% of the reachable e-mail addresses are operated by one of three organizations

Questions?