Network architecture and ROA protection of government mail domains: A case study

Alberto Bartoli

Department of Engineering and Architecture, University of Trieste, Italy

ARTICLE INFO

Keywords:
BGP
DNS
ROA
Email
RPKI

ABSTRACT

Email is a crucial technology used in daily interactions of citizens, enterprises and organizations with their respective governments. In this work we are concerned with the *country-wide* network architecture of *mail domains* of *public administrations*. We analyze a dataset of government mail domains in Italy, Germany, the United Kingdom and the United States of America in order to investigate the opportunities for a *network attacker* to violate security properties of email communication, including availability, in large portions of a country. Issues of this kind are particularly relevant in times of high international tension and in which every country should treat its networks as a potential target for other countries.

We define a framework for describing the opportunities for a network attacker in the resolution of mail domain names, resolution of mail server names, access to a mail server. Based on this framework, we investigate in detail a number of issues related to redundancy and distribution of dependencies among networks and autonomous systems. We also analyze the usage in the access to mail domains of *Route Origin Authorization (ROA)*, an important defensive technology for detecting attacks at the IP routing level. Our analysis allows gaining important insights into the actual network architecture of such an important piece of critical infrastructure as government mail domains.

1. Introduction

Email has been an important application of the Internet since its inception and has become a crucial component of many processes in a broad variety of fields, including daily interactions of citizens, enterprises and organizations with their respective government. Email was not designed with security in mind, thus several technologies have been proposed and developed for improving its security and trustworthiness [1–4]. These technologies are usually integrated with the DNS: the sending and receiving mail domains may infer from the DNS whether the partner supports or requires specific email security enhancements, along with the corresponding policies and parameters. Attacks to selected DNS zones, as well as *network attacks* that divert traffic from the legitimate nameservers to malicious nameservers, have thus the potential of neutralizing several email security technologies, thereby leading to violation of the corresponding security guarantees, i.e., secrecy, integrity, authentication. Furthermore and most importantly, email security technologies defend against network attackers acting as man-in-the-middle, whether in communicating with nameservers or with mail servers, by *preventing* the delivery of legitimate email messages. Such network attackers may thus affect *availability* of receiving mail domains, perhaps in a selected or time-varying way, which may have disrupting effects on the involved organizations.

Indeed, many recent incidents have amply demonstrated the feasibility and practical relevance of such network attacks with large scale impact. For example, in June 2022 the US Cybersecurity and Infrastructure Agency has issued an advisory for alerting of long-lasting intrusions in telecommunication and network providers by state-sponsored attackers [5]. Similar long-lasting intrusions in telecommunication providers have been uncovered in the recent years, e.g., [6,7]. An attack that affected connectivity provided by a state-owned telecommunication company occurred at the onset of the Ukrainian war [8]. A significant portion of the Belgium government IT services was shut down by a denial of service attack against a government-funded ISP [9]. Similar denial of service attacks hit two telecom providers serving Israel government services [10] and several small privacy and security-centric email services [11]. The fact that a single denial of service attack may have a disruptive effect on a very large set of organizations was indeed proven years ago, e.g., [12]. Several Internet service providers and telecommunication companies have been affected in their operational capabilities also by financially motivated attacks [13–17]. More broadly, network attacks for acquiring man-in-the-middle capability on a large scale can be executed by manipulating DNS records, e.g., [18–20,20–22] or by propagating malicious routing information at the BGP level, e.g., [23–29]. Interestingly, attacks at the routing level may allow

E-mail address: bartoli.alberto@units.it.

obtaining man-in-the-middle capability not only for a web servers or mail servers, but even for nameservers, e.g., [30].

In this work we are concerned with the *country-wide* structure in terms of *networks* and *autonomous systems* of *mail domains* of *public administrations*. Specifically, we will investigate the following research questions.

RQ1 How many mail domains of the public administration of *a whole country* could be affected by a network attacker capable of controlling a *single* IP address range or autonomous system? How are those attack opportunities distributed along the various steps of mail delivery, i.e., mapping a mail domain name to a mail server name, mapping the latter to an IP address, communicating with that address?

RQ2 What redundancy levels tend to be used for the email infrastructure of a given country? Do such levels tend to be identical at all the steps of mail delivery or different redundancy levels tend to be used at each step?

RQ3 Autonomous systems that support the email infrastructure of a given country tend to be managed by government organizations or by private companies? Are such companies owned by foreign countries?

RQ4 Are there any architectural patterns or design choices that are common across different countries?

RQ1 requires analyzing, in particular, whether mail domains of a given country tend to be highly spread across many different networks and autonomous systems or highly concentrated in a few of them. RQ1 and RQ2 address a fundamental and unavoidable tradeoff: higher redundancy implies higher resiliency country-wide but at the cost of a wider security perimeter.

Answering these research questions is critical to understanding such an important component of critical infrastructure as public administration email services. Issues of this kind have ever-growing importance, especially in times of high international tension and in which every country should treat its networks as a potential target for other countries.

We address our research questions by defining a framework for describing the dependencies of mail domains from DNS zones, IP address ranges and autonomous systems. Based on this framework, we collect the relevant dependencies of email domains of the public administration in Italy, Germany, the United Kingdom and the United States. We identify the entities involved in the mapping from mail domain names to IP addresses of mail servers and then those involved in the access to those IP addresses, in order to list the opportunities for a network attacker. Then we aggregate and analyze the resulting data from several points of view.

An important component of our study consist in analyzing the usage of *Route Origin Authorization (ROA)*, a security mechanism aimed at detecting *route origin hijacks* in which a network attacker attempts to impersonate the autonomous system responsible for a certain IP prefix by sending malicious BGP routing messages [31–33]. A ROA is a binding between an IP address block and the autonomous system (AS) that owns that block, digitally signed by that AS. The structure and validation rules for such signed bindings are standardized within the *Resource Public Key Infrastructure (RPKI)* framework. ROA information is stored in public RPKI repositories distributed across the Internet. ASes are supposed to execute *Route Origin Validation (ROV)*, for detecting and discarding received BGP routing messages that are inconsistent with ROA [34]. While ROA is not a defense against *invalid path announcement* attacks, in which the network attacker advertises routing messages with a path to a valid ROA passing through an attacker-controlled AS, ROA does address a foundational weakness of the Internet routing system and can significantly improve BGP security. Unfortunately ROA usage is still low, although growing rapidly: 31.39% vs 39.42% of the IPv4 address space in August 2021 and August 2022, respectively, according to the NIST RPKI monitor.[1] In this work we analyze in detail the deployment of ROA:

RQ5 What is the actual deployment of ROA in networks and autonomous systems responsible for the email infrastructure of the 4 countries in our dataset?

This analysis allows gaining important insights into the actual deployment of this important defensive technology.

1.1. Related work

We are not aware of any research examining the network architecture of large sets of email domains managed by different and geographically dispersed organizations. Several studies have examined the actual deployment of email security technologies: SMTP over TLS (STARTTLS) [2], Mail Transfer Agent-Strict Transport Security (MTA-STS) [4], email signing and encryption (S/MIME, DKIM, DMARC) [2,3], sender policy framework (SPF) [2,3] and secure DNS (DNSSEC) [35–38]. We do not analyze the deployment of these technologies in our dataset. We note that these technologies do *not* neutralize attacks and leave normal operations proceed undisrupted: they transform attacks aimed at violating secrecy, authentication, integrity to a form of denial of service.

The network architecture of the DNS has been analyzed in a number of studies, e.g., [39–41]. The robustness of *second level* DNS domains was analyzed by [42]. The cited work analyzed, among other things, number and placement of nameserver replicas as well as usage of shared infrastructure for name resolution, i.e., groups of nameservers used by many different domains. The framework and methodology used by [42] form the basis for our study. Another significant work in this area is [43], which showed that many websites of the Alexa Top 1 Million list share the same infrastructure for name resolution, with a significant group of 12.000 different websites that actually share all their nameservers. In our dataset we analyzed the shared infrastructure for name resolution and for actual access to mail servers, similar to [44] which examined the robustness of DNS paths to government web sites in the same countries of the present work.

Dependencies among DNS zones are analyzed in depth in [45], by defining different kinds of dependencies based on the possibility of resolving names in a zone when other zones are not reachable and by analyzing the resulting dependency graph for the Alexa top 1 million sites according to several dependency metrics proposed for this purpose. We propose a dependency graph for mail domains that includes dependencies among heterogeneous entities, i.e., mail domains, mail servers, zones, nameservers, IP address ranges and autonomous systems. We use a single dependency rule for zones and do not attempt to quantify the availability of a zone based on the availability of other zones (or of other entities).

A comprehensive analysis of government DNS domains for over 190 countries and based on historical data spanning 10 years is provided by [46]. The cited work provides a broad and deep analysis of the considered domains, including their network configuration in terms of number of nameserver replica and of dependency from third parties, which is an aspect crucial in our work. Our study of the DNS issues related to mail domains is undoubtedly much less comprehensive than the one in the cited work, though.

A large-scale analysis of the actual deployment of ROA since it was first deployed in 2008 is given in [47]. Tools for the live assessment of ROA deployment are made publicly available by NIST [48] and by APNIC Labs [49]: the former provides global statistics while the latter provides information at the IP address range level. A detailed analysis

[1] https://rpki-monitor.antd.nist.gov/ROV

of the factors that hamper ROA deployment along with proposals for improving the current situation can be found in [50,51].

To our knowledge, studies on ROA deployment do not consider the class of applications hosted in the corresponding IP address ranges. We analyze ROA deployment in networks and autonomous systems involved in mail domains of public administrations of four countries. Such a focus allows gaining useful insights into the actual state of ROA deployment.

Effectiveness of ROA requires that autonomous systems indeed fetch the relevant data (the *Resource Public Key Infrastructure (RPKI))* and execute *Route Origin Validation (ROV)* (also called *RPKI filtering*), to detect and discard received BGP routing messages that are inconsistent with ROA. An analysis of the correctness and completeness of ROA data is provided by [52] while the actual deployment of ROV is estimated in [34]. A tool for live assessment of ROV usage at the granularity of autonomous system is provided by APNIC Labs [53]. Robustness and availability of ROA information in terms of its dependency from the DNS is analyzed in [54]. The cited work emphasizes that a large amount of RPKI deployments could be disabled by several forms of DNS attacks, e.g., by making one single nameserver unavailable or by injecting forged DNSSEC responses in components that are involved in the DNS name resolution chain and do not validate such responses correctly. An attack that prevents the fetching of ROA information, thereby disabling such a defensive mechanism, is presented in [55]. Such an attack can be executed by an offpath attacker and is effective against a large fraction of the existing RPKI repositories. ROA is not able to detect all kinds of routing hijacks attacks, for example, ROA cannot detect *invalid path announcement* attacks, in which the network attacker advertises routing messages with a path to a valid ROA that passes through an autonomous system controlled by the attacker. A categorization of BGP hijacking attacks is given in [56].

Scenarios in which Internet traffic to a large portion of a full country is managed by a small set of autonomous systems, thereby creating a significant risk of selective observation and tampering country-wide, are identified in [57]. We analyze the issue of country-wide dependency from selected autonomous systems for only four countries, only for mail domains of public administrations. A methodology for identifying state-owned autonomous systems is presented in [58]. Such a methodology is not suitable for our study, though, because it excludes autonomous systems that are instead crucial from our point of view, such as those managed by research institutions and those that provide connectivity to government services. We thus assess whether autonomous systems are managed by public or private organizations by simply analyzing their respective description.

2. Analysis framework

Consider an email addressed to a mailbox in mail domain d. The email will pass through several hosts, from the sending user agent to a mail server responsible for receiving emails addressed to d, say $M(d)$. Delivery of the email to $M(d)$ will occur by means of an SMTP interaction between a process acting as SMTP client and $M(d)$. This interaction conceptually consists of the following stages:

(1) *Name resolution (mail domain):* Mapping the name of d to the name of $M(d)$.
(2) *Name resolution (mail server):* Mapping the name of $M(d)$ to its IP address $IP(M(d))$.
(3) *Access:* Communicating with $IP(M(d))$.

We intend to analyze the opportunities for a *network attacker* at each of these steps in order to violate one or more of availability, secrecy, integrity of mail delivery.

We are not concerned with attacks that occur before the last hop of mail delivery, or during the interaction between the destination user agent and $M(d)$, or during any possible forwarding that follows mail delivery.

2.1. Dependency graph

We define a *dependency graph* for describing the dependencies between the entities involved in the execution of the last hop of mail delivery. This definition extends the framework proposed in [44] for websites, which in turn was an extension of the framework in [42] for analyzing the DNS architecture of second-level DNS domains.

Each *node* of the dependency graph corresponds to one of the following entities:

- mail domains,
- mail servers,
- name servers,
- DNS zones,
- alias names (see below),
- IP/24 address ranges (*networks*, for short),
- autonomous systems.

Nodes are connected by oriented *arcs*. An arc from node n_1 to node n_2 means that the entity represented by the former *depends* on the entity represented by the latter. Dependency rules are defined below. According to these rules, nodes with inbound degree zero are mail domains whereas nodes with outbound degree zero are autonomous systems.

Let n_s be the name of either a mail server or of a name server. If a CNAME DNS record exists that maps name n_s to value n_v, then we say that n_v is an *alias name*.[2]

- A mail domain named m depends on: (i) the zone that m belongs to; and (ii) the mail servers specified in the MX DNS record associated with m.
- A mail server, or a name server, or an alias name named n depends on: (i) the zone that n belongs to; and, (ii) either a network (when a DNS record $n\ A\ n_{IP}$ exists; the network will be the one to which n_{IP} belongs to) or on an alias name (when a DNS record $n\ CNAME\ n_A$ exists; the alias name will be n_A).
- A zone depends on: (i) its name servers, (ii) its parent zone in the DNS tree, and (iii) the zones of the names of its name servers (with the exception that a zone does not depend on itself).
- A network depends on the autonomous system responsible for that network.
- An autonomous system does not depend on any other entity.

2.2. Name resolution paths and access paths

Given a mail domain d, consider the set of *paths* in the dependency graph that start at d and end at an autonomous system. It is simple to realize from the dependency rules in the previous section that such a set can be partitioned as: (i) paths that include a name server (that we call *name resolution* paths); and (ii) paths that include a mail server (*access* paths). Each name resolution path describes entities and dependencies that may be involved for obtaining the IP address of a mail server of d. Each access path describes entities and dependencies involved in the routing of packets for communicating with a mail server of d, i.e., when the IP address of that server is known. With respect to the three stages in the last hop of mail delivery (Section 2):

(1) *Name resolution (mail domain):* This step corresponds to a sequence of one or more name resolution paths (one for each nameserver contacted for mapping the mail domain name to mail server names).
(2) *Name resolution (mail server):* This step also corresponds to a sequence of one or more name resolution paths (one for each nameserver contacted for mapping the mail server name to an IP address).

[2] Actually one could state that n_s is an alias for n_v; we preferred to say that n_v is an alias for simplifying the terminology.

(3) *Access:* This step corresponds to an access path.

Name resolution paths and access paths are *not* meant to describe the actions actually executed by an SMTP client for resolving names or accessing a mail server: they are meant to describe opportunities for a *network attacker*. A network attacker in control of an entity may alter the behavior of that entity, thereby provoking a violation of availability, secrecy, integrity for mail domains whose paths pass through that entity (see next section).

To further clarify our framework, note that name resolution paths describe *potential* dependencies. For example, the set of nameservers to be contacted for the mapping operations of a given SMTP interaction will depend on the contents of the DNS caches at the hosts involved in that interaction. Furthermore, for a given mail domain, there may be multiple redundant name resolution paths and mail access paths. In other words, the actual impact of an attack on a given entity will depend on how many SMTP interactions will actually involve that entity. This issue is beyond the scope of this work.

2.3. Threat model

We consider a *network attacker* that may take control of one or more *networks* and/or *autonomous systems* in the dependency graph. Such an attacker may observe, modify, discard, forge any packet in the controlled entity. The method by which the attacker takes control of an entity and modifies the behavior of that entity in order to achieve the described results is irrelevant to our analysis. Network attackers that take control of zones and/or nameservers are also practically relevant but beyond the scope of this work.

Let d be a mail domain and let e be an entity in one of the paths of d. An attacker in control of e may achieve one or more of the following results:

- *Denial of Service*. The entity stops functioning, i.e., all packets are dropped. Such a result may affect *availability* of d.
- *Snooping*. Emails addressed to d will be observed by the attacker, along with all the associated SMTP traffic. Such a result may affect *secrecy* of emails addressed at d.
- *Impersonation*. Emails addressed to d will be deposited at an IP address in control of the attacker; and, they may or may not be deposited also at their legitimate location; if they are, they may or may not be modified by the attacker. Such a result may affect *secrecy* and *integrity* of emails addressed at d, as well as *availability*.

2.4. Dataset

We constructed a list of mail domains of public administrations in Italy (IT), Germany (DE), the United Kingdom (UK) and the United States (US) as follows. For IT we considered all the domains of the Italian certified email system (Posta Elettronica Certificata, PEC[3]). This service allows citizens to send emails with legal value equivalent to a registered letter with return receipt. We downloaded the corresponding list of mail domains from the Opendata catalog of the Italian public administration on January 27-th 2022.[4] The list contained 60098 distinct email addresses corresponding to 5064 distinct email domains. All these email domains indeed belong to the Italian PEC system. For DE, UK, US we could not find any similar list thus we proceeded as follows. We considered the lists of websites of public administrations used in [44], that were downloaded from publicly available lists.[5] For each domain in these lists, we dropped every 'www' prefix and used each resulting domain name as a candidate mail domain. The resulting number of live mail domains is summarized in Table 1.

Table 1
Dataset summary.

Country	Mail domains
IT	4866
DE	150
UK	482
US	411

We then constructed the dependency graph for these mail domains by actively contacting the DNS. We did not analyze AAAA records; we fetched records from a single location and only once (except for masking transient failures). We mapped networks to autonomous systems based on a public database updated hourly.[6] We obtained ROA data for the networks of our interest from the APNIC ROA generation report tool,[7] a publicly available website updated daily and reporting measurements obtained from 600 distinct vantage points.

3. Analysis

3.1. Direct zones

3.1.1. Methodology

Given a mail domain d, let $Z^D(d)$ denote the zone containing the name of d and let $Z^S(d)$ denote the set of the zones containing the names of the mail servers of d. We say that zones in $Z^D(d) \cup Z^S(d)$ are *direct* zones for d. We restricted *all* our analyses on name resolution paths by focussing *only* on networks and autonomous systems in which the nameservers of *direct zones* reside.

To motivate this choice, consider an attacker that controls a certain entity e on a name resolution path for d. The actual impact of this attack on d will depend on: (i) whether there are name resolution paths for d that do *not* pass through e; and, (ii) the number of name resolution procedure executions that actually pass only through those paths. While estimating such impact is beyond the scope of this work, it is clear that attacks on entities of direct zones will be most effective because *all* the name resolution paths certainly pass through entities of direct zones (while only a fraction of those paths passes through entities of zones that are not direct).

For each mail domain d in our dataset we determined the corresponding direct zones; then, we determined the name servers of those zones, the networks where those name servers are placed, the autonomous systems responsible for those networks.

3.1.2. Findings

Table 2 summarizes the number of entities involved in name resolution paths and in access paths, separately for each dataset.

Regarding *name resolution paths*, the following observations can be made:

- The security perimeter of mail domain resolution for the UK dataset is extremely small, as there is only one zone involved. This fact is certainly the result of careful planning and implementation. On the other hand, the number of zones involved in the resolution of mail server names is much higher and comparatively similar to the number of zones in DE and US datasets.
- The previous remark can be made also in terms of name servers, networks and ASes: an extremely small security perimeter for mail domain resolution but an architecture for mail server resolution similar to that of DE and US.

[3] https://en.wikipedia.org/wiki/Certified_email

[4] https://indicepa.gov.it/ipa-dati/dataset/elenco-pec

[5] https://www.gov.uk/government/publications/list-of-gov-uk-domain-names, https://home.dotgov.gov/data/, https://github.com/robbi5/german-gov-domains.

[6] https://iptoasn.com/

[7] https://stats.labs.apnic.net/roas

Table 2
Entities involved in name resolution.

Name resolution paths: Mail domains (direct zones only)				
Country	#Zones	#Name servers	#Networks	#ASes
IT	4550	1522	899	229
DE	127	160	111	52
UK	1	7	8	3
US	411	569	237	92

Name resolution paths: Mail servers (direct zones only)				
Country	#Zones	#Name servers	#Networks	#ASes
IT	39	78	67	32
DE	83	145	112	50
UK	123	318	203	74
US	100	274	181	68

Access paths			
Country	#Mail servers	#Networks	#ASes
IT	64	30	15
DE	117	74	36
UK	438	197	68
US	380	173	69

- The IT dataset exhibits an architecture for name resolution that is somewhat the opposite of the one of UK: a very large number of zones for resolving mail domains (almost one zone for each mail domain) and a much smaller number of zones for resolving the names of mail servers. The latter is the results of the technical and organizational requirements for the Italian PEC system, that allow only 18 *PEC providers* to offer such a service and require each public administration to subscribe to one of those providers. Interestingly, though, there are 39 zones involved in the mapping of mail server names to IP addresses for those 18 PEC providers.

Regarding *access paths*, the salient aspect is the difference between IT and the other datasets: IT exhibits the smallest amount of entities: 64 mail servers and 30 networks manage the entirety of the email of the Italian public administration (4866 mail domains). Such a reduced size is particulary significant having considered the much larger size of this dataset: on the average, IT has 76 and 30 mail domains for each mail server and for each network, respectively; the corresponding figures for UK are 1.1 and 2.4, respectively (the distributions will be analyzed in detail in Sections 3.6 and 3.7). Such a much stronger concentration for IT reflects the usage of PEC providers as discussed above.

3.2. Path redundancy

In the next two sections we analyze path *redundancy* in each dataset at the level of *networks* and of *autonomous systems*. Consider an attacker targeting a mail domain d and capable of attacking a network (the reasoning for autonomous systems is conceptually identical). The attacker may impact d at each of the three stages in the last hop of mail delivery (Section 2). Broadly speaking, the lower the path redundancy, the higher the potential impact of a successful attack on a single network. For example, in the resolution of mail domain, if all the nameservers of $Z^D(d)$ are in the same network, then a denial of service attack on this single network may suffice to render d completely inaccessible (except for clients that can exploit DNS caching); and, taking control of packet routing in that single network may suffice to snoop all the emails addressed to d (by resolving the name of d to the name of an attacker-controlled mail server). On the other hand, if the nameservers of $Z^D(d)$ are distributed over multiple networks, then a successful attack on only one network will impact only the resolutions of the name of d that happen to pass through the affected network. In the next sections we will analyze path redundancy at the level of each full dataset, separately for the three stages: name resolution domain, name resolution server, access.

It is important to remark that while higher redundancy improves resiliency w.r.t. to *successful* attacks, it may also increase the likelihood of an attempted attack to actually succeed: the higher the path redundancy, the larger the security perimeter to defend. For a given defensive budget, enlarging the security perimeter implies diluting defensive resources thereby obtaining a lower barrier to attacker. On the other hand, a large security perimeter does not necessarily correspond to a low defensive barrier, depending on the overall defensive budget available.

It follows that the objective of the next sections is *not* ranking the analyzed countries based on how much redundancy they exhibit at each step. The objective is highlighting the overall architecture of each of the three steps and, if possible, determining to which extent those architectures are the result of explicit and centrally-driven design choices or are just the outcome that happened to emerge from decentralized and uncoordinated actions by the involved organizations.

3.3. Redundancy of name resolution paths

3.3.1. Methodology

We base our analysis on the framework proposed by [42] for categorizing second-level DNS domains based on the redundancy of the respective nameservers. The DNS specification requires that each zone maintains at least two nameservers [59] and that these nameservers should be both geographically and topologically diverse [60]. The framework in [42] approximates the requirement for geographical and topological diversity with membership to different networks and partitions second-level DNS domains in three categories depending on the degree to which they meet those requirements. A conceptually similar categorization in terms of ASes rather than in terms of networks was used in [44].

Following the above ideas we performed four different categorizations of each mail domain, one for each combination of: (i) direct zones of mail domains or of mail servers; and, (ii) number of networks or of ASes in which the corresponding nameservers are distributed. Given a zone z, let $\#n(z)$ and $\#AS(z)$ be the number of different networks and of different ASes in which nameservers of z are distributed. We categorized each mail domain d as follows:

(1) *Name resolution (mail server), networks:* For each zone $z \in Z^S(d)$ we determined $\#n(z)$; we selected the minimum of those values, denoted $\#n^{min}(d)$; then:
   - $\#n^{min}(d) \geq 3 \Rightarrow d$ in the *exceeds* category;
   - $\#n^{min}(d) = 2 \Rightarrow d$ in the *strictly meets* category;
   - $\#n^{min}(d) = 1 \Rightarrow d$ in the *does not meet* category;

(2) *Name resolution (mail domain), networks:* Same procedure as 1, by replacing the set of zones $Z^S(d)$ by the single zone $Z^D(d)$.
(3) *Name resolution (mail server), ASes:* Same procedure as 1, by replacing $\#n(z)$ by $\#AS(z)$.
(4) *Name resolution (mail domain), ASes:* Same procedure as 3, by replacing $\#n(z)$ by $\#AS(z)$.

To clarify, consider a mail domain d in the "does not meet" category for all the four categorizations:

(1) *Name resolution (mail server), networks:* There is at least a mail server of d that can be fully controlled by taking control of one single network: either that mail server can be made inaccessible (i.e., by preventing name resolution for the mail server name from completing) or all the messages addressed to that mail server can be snooped (i.e., by resolving its name to an attacker-controlled mail server); similarly, that mail server can be made inaccessible by a denial of service attack on one single network.

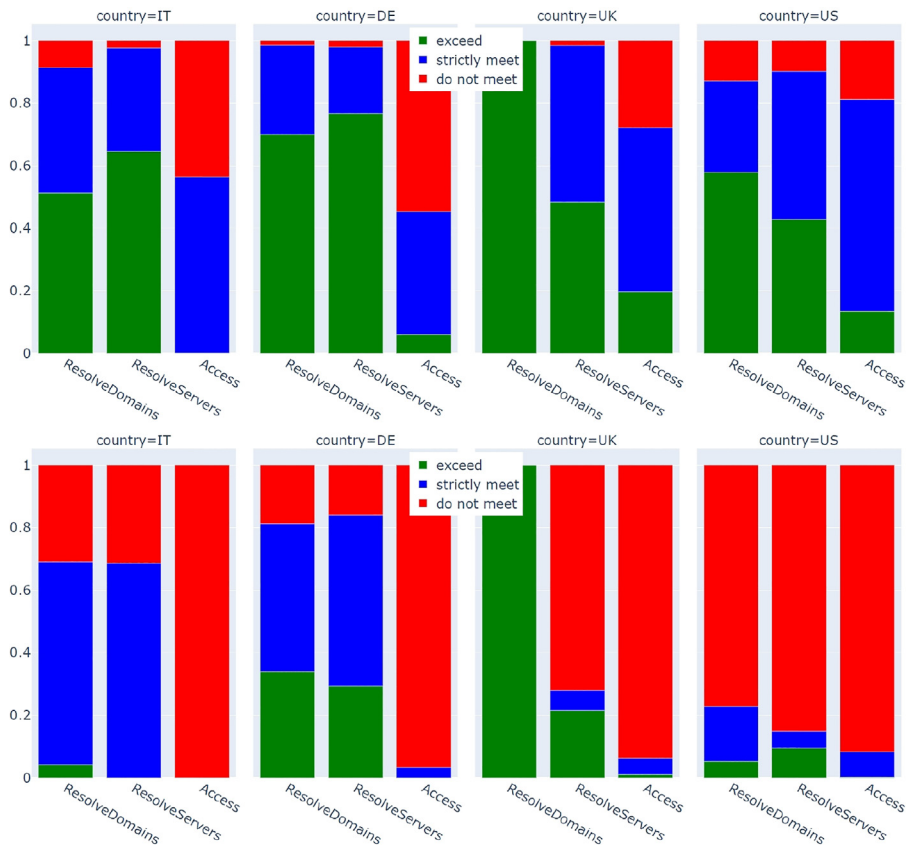

Fig. 1. Categorization of mail domains based on redundancy of name resolution paths and of access paths (network redundancy up, AS redundancy down).

(2) *Name resolution (mail domain), networks:* The entire mail domain can be fully controlled by taking control of one single network.
(3) *Name resolution (mail server), ASes:* There is at least a mail server of d that can be fully controlled by taking control of one single AS.
(4) *Name resolution (mail domain), ASes:* The entire mail domain can be fully controlled by taking control of one single AS.

3.3.2. Findings

The resulting categorizations are summarized in Fig. 1 (redundancy of access paths will be discussed in Section 3.4). We make the following observations:

- Name resolution for mail domains in the UK dataset exhibits the highest redundancy, with all domains that exceed the robustness requirement not only in terms of networks but also in terms of ASes. This is a remarkable result, in particular, because it is coupled with a very small security perimeter in terms of zones as observed in Section 3.1.
- In the IT dataset, there is a slight increase in network redundancy when comparing mail domain resolution to mail server resolution, which may be explained with the usage of PEC providers specialized in offering mail management. On the other hand, such a difference is not observed in terms of AS redundancy, i.e., PEC providers do not tend to enforce AS redundancy more than DNS providers of public administrations.
- US and UK exhibit a remarkably low redundancy in terms of ASes (only for mail server resolution, in the case of UK). We have no elements for telling whether this fact is a result of specific design choices or just properties emerging from such factors as market dynamics, normative requirements and alike. However, by analyzing the most widely used ASes in the US dataset (Section 3.6), it turns out there is a widespread usage of ASes managed by US companies specialized in denial of service protection and content distribution networks, such as Microsoft, Akamai, Cloudflare, Google and Amazon. Usage of a single AS, thus, is likely to be a specific design choice aimed at keeping the security perimeter low while maintaining high resilience to network attacks.
- Some 10% of mail domains for the US dataset have all their name servers concentrated in a single network ("does not meet" category). By looking at those domains, it appears that most of them are in the area of public health services.
- In a recent analysis of 192.6K DNS government domains in 190 countries, the percentage of those domains corresponding to the "does not meet" category was 28.5% and 67.1% in terms of networks and of ASes, respectively [46]. Results in Fig. 1 exhibit much smaller values for such a category, with only two exceptions for name resolution of servers where the values for UK and US are higher.
- The analysis of government websites in IT, DE, UK, US [44] analyzed redundancy in name resolution with the same methodology as the one used here. The results were very similar to those obtained here, with one significant exception for AS redundancy: it was around 50% for UK websites, while it is close to 0 and more than 70% for UK mail domains and UK mail servers, respectively.

A complementary view of the data summarized in Fig. 1 is given by Fig. 2, where name resolution paths are grouped based on the number of ASes and networks in which nameservers are distributed (mail domains up, mail servers down). We observe what follows:

- There is no clear architectural pattern common across all datasets (the four plots in each row are quite different from each other).

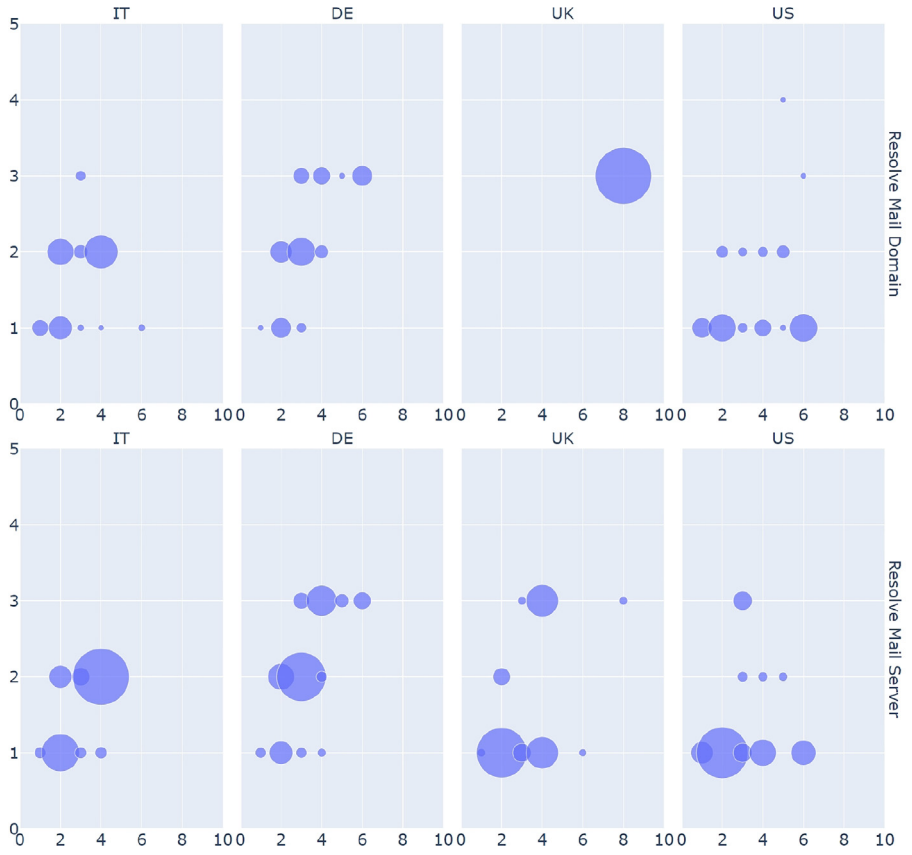

Fig. 2. Redundancy of name resolution paths (mail domains up, mail servers down). The size of each dot is proportional to the percentage of mail domains in the respective dataset associated with that dot. Each mail server is weighed by the percentage of mail domains in the respective dataset that use that server.

- The most significant example of centralized coordination is UK-domains (as observed above), with all the mail domains having the same redundancy but with a significant difference from UK-servers.
- DE is the only dataset where the distributions for mail domains and mail servers are quite similar to each other. This fact might result from specific technical requirements but we have no elements in support of this claim.
- IT-servers (i.e., PEC providers) is the dataset that exhibits the smallest redundancy.
- UK and US exhibit a strong prevalence of name resolution for mail servers in two networks and in a single AS (indeed, US has even a significant amount of domains in a single network). This fact is probably explained by a widespread usage of a cloud provider that has chosen to implement name resolution for its mail servers with small network/AS redundancy (i.e., Microsoft, Sections 3.6.2 and 3.6.3). In IT and DE, instead, the stronger prevalence of name resolution for mail servers corresponds to four networks over two ASes. In this case, higher redundancy could be explained by the fact that the corresponding organizations cannot be as confident as Microsoft to ensure high availability of name resolution with small network/AS redundancy.

3.4. Redundancy of access paths

3.4.1. Methodology

We used the same approach as in the previous section by replacing nameservers by mail servers. Given a mail domain d, let $\#n_A(d)$ and $\#AS_A(d)$ be the number of different networks and of different ASes in which mail servers of d are distributed. We categorized each mail domain d as follows:

(1) *Access, networks:*

- $\#n_A(d) \geq 3 \Rightarrow d$ in the *exceeds* category;
- $\#n_A(d) = 2 \Rightarrow d$ in the *strictly meets* category;
- $\#n_A(d) = 1 \Rightarrow d$ in the *does not meet* category;

(2) *Access, ASes:* Same procedure as 1, by replacing $\#n_A(d)$ by $\#AS_A(d)$.

For ease of presentation, we chose to use for the three categories the same names as for name resolution paths. We emphasize, though, that these categories (as well as their names) are somewhat arbitrary because access to mail servers is not related to DNS requirements, unlike name resolution paths.

3.4.2. Findings

The resulting categorizations are summarized in Fig. 1, along with those for name resolution paths. We make the following observations:

- Not surprisingly, access path redundancy is much smaller than name resolution redundancy in all the datasets.
- At the network level, redundancy is much higher in US/UK than in IT and DE. IT and DE have large amounts of domains in the “does not meet” category (43.5% and 54.6%) while those amounts are much smaller in UK and US (27.8% and 18.7%). Both UK and US have significant amounts of domains in the “exceeds” category (19.7% and 13.3%) while IT has none.
- At the autonomous system level there is almost no redundancy, but DE, US, UK indeed have some of their mail domains with mail servers distributed over 2 different autonomous systems: between 3% and 8%.
- All the datasets exhibit much higher redundancy in access paths to mail servers than in access paths to webservers [44].

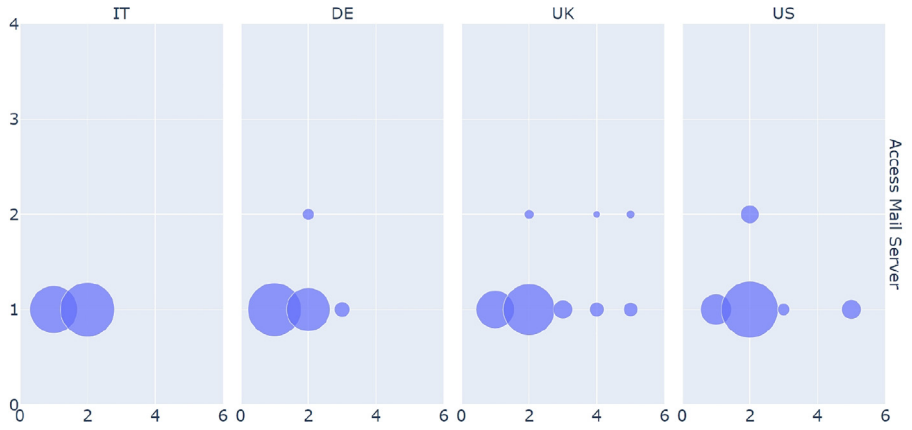

Fig. 3. Redundancy of access paths. The size of each dot is proportional to the percentage of mail domains in the respective dataset associated with that dot. Each mail server is weighed by the percentage of mail domains in the respective dataset that use that server.

We then grouped access paths based on the number of ASes and networks in which mail servers for a given mail domain are distributed (Fig. 3):

- There are more similarities across countries in name access paths than previously found for name resolution paths (Fig. 2): in this case, in all countries, most mail domains are concentrated in the region with one or two networks and only one AS.
- IT is the dataset that exhibits the smallest redundancy, with no access path spread across multiple ASes or beyond more than two networks.
- US and UK exhibit a significant amount of access paths with higher redundancy, in part even across two ASes.
- DE has most access paths with one network only. No access path with more than three network and some access paths with two networks in different ASes.

3.5. ROA usage

3.5.1. Methodology

We say that a network is ROA-protected if there exists a (valid) ROA attestation for that network (we outline how ROA protection works in the introduction). We check whether a given network is ROA-protected by querying the APNIC ROA generation report.[8] To place the results of this section in some perspective, we observe that the percentage of IP addresses associated with a ROA attestation was estimated in June 2022 as:

- APNIC Labs[9]: 29.02%, 36.29%, 37.04% and 17.22% for IT, DE, UK and US, respectively.
- NIST RPKI Monitor[10]: 36.95% globally.

Notice that, according to these estimates, US lags significantly behind the other countries analyzed and behind the global average.

First, we analyzed ROA usage at the level of *mail domains*. For each mail domain d, we determined:

- *Resolution domain:* the percentage of ROA-protected networks in the set of networks where the nameservers in $Z^D(d)$ reside.
- *Resolution server:* the percentage of ROA-protected networks in the set of networks where nameservers in $Z^S(d)$ reside.
- *Access:* the percentage of ROA-protected networks in the set of networks where mail servers in $A(d)$ reside.

We then categorized mail domains in four bins, depending on the percentage value: all, more than half, less than half, none.

Next, we analyzed ROA usage at the level of *networks*, separately for each kind of paths. We denote by C a dataset:

- *Resolution domain:* let N_C^D be the set of networks where nameservers of zones in $\bigcup_{d\in C} Z^D(d)$ reside; we determined the percentage of ROA-protected networks in N_C^D.
- *Resolution server:* let N_C^S be the set of networks where nameserver of zones in $\bigcup_{d\in C} Z^S(d)$ reside; we determined the percentage of ROA-protected networks in N_C^S.
- *Access:* let N_C^A be the set of networks where mail servers of all mail domains in C reside; we determined the percentage of ROA-protected networks in N_C^A.

Finally, we analyzed ROA usage at the level of *ASes*, separately for each kind of paths. We denote by C a dataset:

- *Resolution domain:* let $N_C^D(AS_i)$ be the subset of N_C^D containing networks of the AS_i; we computed the percentage of ROA-protected networks in $N_C^D(AS_i)$.
- *Resolution server:* the same computation w.r.t. N_C^S.
- *Access:* the same computation w.r.t. the set of networks where mail servers of all mail domains in C reside.

We categorized ASes in four bins depending on the percentage of their networks that are ROA-protected. Such an analysis is useful for understanding whether ASes tend to use ROA systematically, i.e., for all of their networks, or not.

3.5.2. Findings

Fig. 4 provides the categorization of *mail domains* based on the amount of networks in the respective path that are ROA-protected, i.e., networks where nameservers for mail domain resolution, nameservers for mail server resolution and mail servers reside. We observe what follows:

- ROA usage for IT, UK servers and, in particular, DE, is well below half of the mail domains. We believe this is an important, and disappointing, result.
- The UK domains dataset once again appears to be tightly coordinated even from the point of view of ROA usage, being the only dataset with all domains having more than half of the networks in their name resolution path ROA-protected.
- US exhibits the largest amount of ROA-protected networks (except for UK domains), despite the fact that ROA usage for generic IP addresses in the US lags significantly behind the other countries analyzed. This fact suggests a form of awareness and coordination

[8] https://stats.labs.apnic.net/roas
[9] https://stats.labs.apnic.net/roas
[10] https://rpki-monitor.antd.nist.gov/ROV

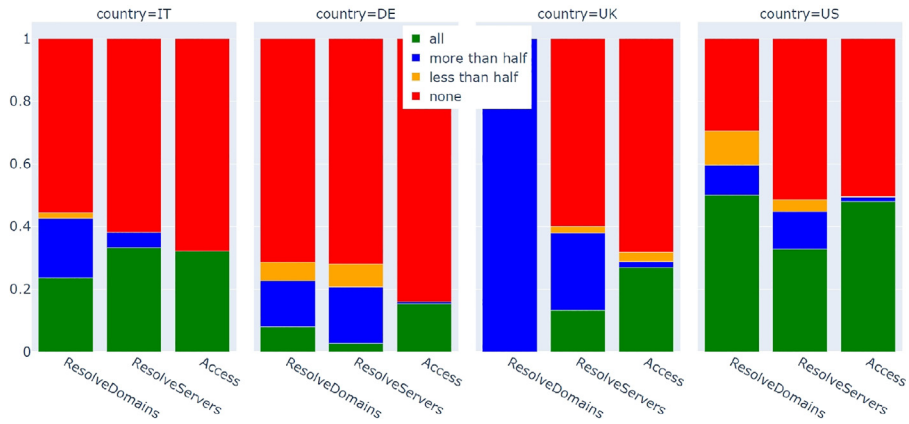

Fig. 4. Categorization of mail domains based on the amount of networks in the respective path that are ROA-protected.

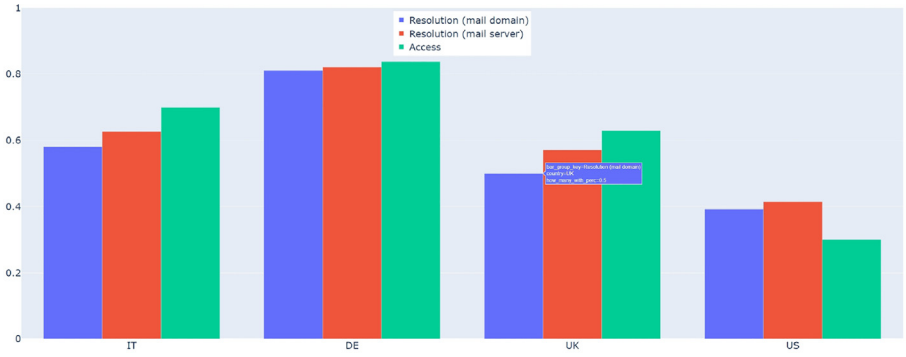

Fig. 5. Percentage of ROA-protected networks.

in the US towards the relevance of ROA for protecting the mail infrastructure.

- ROA usage in access paths is very similar to usage in name resolution paths for mail servers: almost the same for US and slightly worse in the other datasets.

ROA usage at the level of *networks* is in Fig. 5. We observe what follows.

- In each dataset, the percentage of ROA-protected networks is significantly higher than the percentage of ROA-protected IP addresses (see the beginning of this section). This fact suggests that ROA is indeed perceived in every country as a defensive mechanism worth implementing. On the other hand, as observed above, large amounts of mail domains have little or none ROA protection in their name resolution and access paths.
- The ranking of ROA protection among datasets reflects ROA usage in the respective country (see the beginning of this section), except for UK where ROA usage is smaller than in IT/DE while at the level of IP addresses it is the opposite. In particular, it is worth noting the lowest usage of ROA in the US at the network level.
- UK and US have a relative amount of ROA-protected networks significantly smaller than those in IT and DE, yet the overall reach of that protection over mail domains is higher (as just observed in Fig. 4). From a different point of view, one could argue that a smaller global overhead results in a higher global protection, due to more centralization and planning.
- In each country, there is no significant difference in ROA protection between name resolution for mail domains, name resolution for mail servers, access to mail servers, as the difference between the three categories is always smaller than approximately 10%.

Finally, categorization of ASes based on the respective amount of ROA-protected networks is in Fig. 6. We observe what follows.

- Almost all ASes involved in this analysis use ROA for all or none of their respective networks.
- The differences between countries are very similar to those observed at the network level (Fig. 5), in this case with a more pronounced difference between US and the other countries. Indeed, a very large percentage of the ASes involved in the US dataset do not provide any ROA protection in spite of the fact that the overall ROA protection for US is the largest among the considered countries (Fig. 4).
- Interestingly, none of the ASes involved in UK domains use ROA for all of its networks.[11] Moreover, there are more ASes that use ROA for all or for more than half of their networks in UK servers and UK access than in UK domains.

3.6. Distribution of name resolution paths

3.6.1. Methodology

We analyzed the distribution of name resolution paths across networks and autonomous systems, in order to determine the number of mail domains that would be affected by an attack on a single entity (i.e., a network or an autonomous system), as well as to identify those entities that are potentially more critical for a country as a whole.

We say that mail domain d *fully depends* on a network n for domain resolution (server resolution) if n contains *all* the nameservers in $Z^D(d)$

[11] The reason why all domains have a majority of ROA-protected networks (Fig. 4) is because, looking at the raw data, it turns out there are only three ASes involved in UK domains and 2 of them that do not have any ROA-protected network.

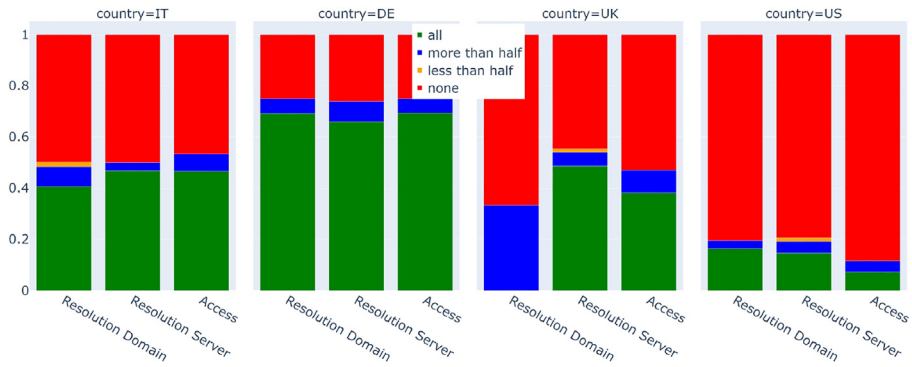

Fig. 6. Categorization of ASes based on the respective amount of ROA-protected networks.

($Z^S(d)$); in this case we say that n *fully controls* d in domain (server) resolution. We say that d *partly depends* on n for domain resolution (server resolution) if n contains at least *one* of the nameservers in $Z^D(d)$ ($Z^S(d)$); in this case we say that n *partly controls* d in domain (server) resolution.

For each network we determined, separately for each dataset, the percentage of mail domains that fully or partly depend by that network in domain resolution or in server resolution. We extended to autonomous systems the above definitions made for networks and computed the analogous distributions for autonomous systems.

We also analyzed the distribution of name resolution paths across *groups* of networks: we determined the groups of 2 or more networks that *fully control* a given mail domain or mail server, i.e., that contain all the name servers for that mail domain or mail server (only for direct zones, as in all our analyses). This analysis allows determining the number of mail domains and mail servers that could be affected by an attack on a group of entities, whether networks or ASes, as well as the groups of entities that are potentially more critical for a country as a whole.

3.6.2. Findings (dependency from networks)

For each dataset, the percentage of mail domains that fully depend on a single network is smaller than 1%, whether in domain resolution or in server resolution: thus, taking control of a single network will allow an attacker to control all name resolution paths of only a tiny fraction of mail domains or of mail servers. By looking at the raw data, the corresponding mail domains do not appear to be particularly critical (with only one exception discussed below), thus the fact that those mail domains full depend on single network does not appear to be a design choice dictated by security-related reasons (e.g., very tight perimeter). The scenario for partial dependency is quite different, as illustrated in Fig. 7 (up) for mail domains. Data for mail domains in UK are particularly interesting:

- All mail domains in UK partly depend on 8 networks; and, all those mail domains partly depend on the same set of 8 networks. This fact confirms the centralized and coordinated approach used for UK domains already observed from several points of view.
- Partial dependency for UK is concentrated in only three ASes and only one of them is managed by a private company (702 UUNET managed by Verizon): JANET and SURFNET correspond to the research and education networks in UK and in the Netherlands, respectively. It is also worthwhile remarking the partial dependency of all mail domains from a public organization of a foreign country.
- Usage of ROA does not appear to follow a systematic pattern: nearly all (yet not all) networks of the JANET AS are ROA-protected; neither the network of a private company nor the one of a foreign public company is ROA-protected.

Regarding the other countries:

- In IT there are 4 networks each of which partly controls 30%–40% of the mail domains. All these networks are ROA-protected and the respective ASes are managed by the private company Aruba. Interestingly, ASN 24806 corresponds to an autonomous system of a company of the Aruba group but located in the Czech republic, which could be an interesting issue from a strategic point of view.
- The 4 networks just discussed constitute the only case of strong concentration in IT, DE, US: all the other networks contain nameservers for a much smaller percentage of mail domains.
- The second and third network in DE, each of which partly controls approximately 20% of mail domains, are in autonomous systems managed directly by the BSI — Bundesamt fuer Sicherheit in der Informationstechnik (Federal Office for Information Security). This fact appears to be the result of an explicit design choice. No similar choice related to an explicit involvement of Information Security agencies appears to emerge from the other datasets, at least not with such a relatively high level of partial dependency.
- Networks ranked five to seven in US (7.5% of mail domains each) are managed by the Dept. of Health and Human Services and by the Dept. of Justice. These are the only examples of ASes directly associated with public administrations, except for BSI in DE, JANET and SURFNET in UK.
- All the networks in the US dataset, except for 5–7, are in ASes managed by private companies specialized in *denial of service (DoS)* protection.
- All the 10 networks in the DE dataset are ROA-protected.
- IT is the only dataset in which no AS of the top 10 networks in Fig. 7 appears to be directly managed with a public institution.

Fig. 7 (down) provides partial dependency data for resolution of mail servers:

- The scenario for UK is quite different from the one for mail domains, both in concentration and in nature of the ASes involved. There are two networks that partly control 34% of mail servers each and several networks that partly control approximately 10% of mail servers each. All the ASes involved are managed by private companies.
- DE exhibits a stronger concentration of partial dependency from the top three networks and these networks are the same as those for resolution of mail domains. This fact appears to be an explicit choice, in particular considering the nature of the corresponding ASes (DFN is the research and education network while BSI is the federal office for information security).
- Both IT and US exhibit a somewhat stronger partial dependency on the respective top networks.

3.6.3. Findings (Dependency from ASes)

After analyzing the distribution of name resolution paths across networks, we analyzed the corresponding distributions across ASes. In

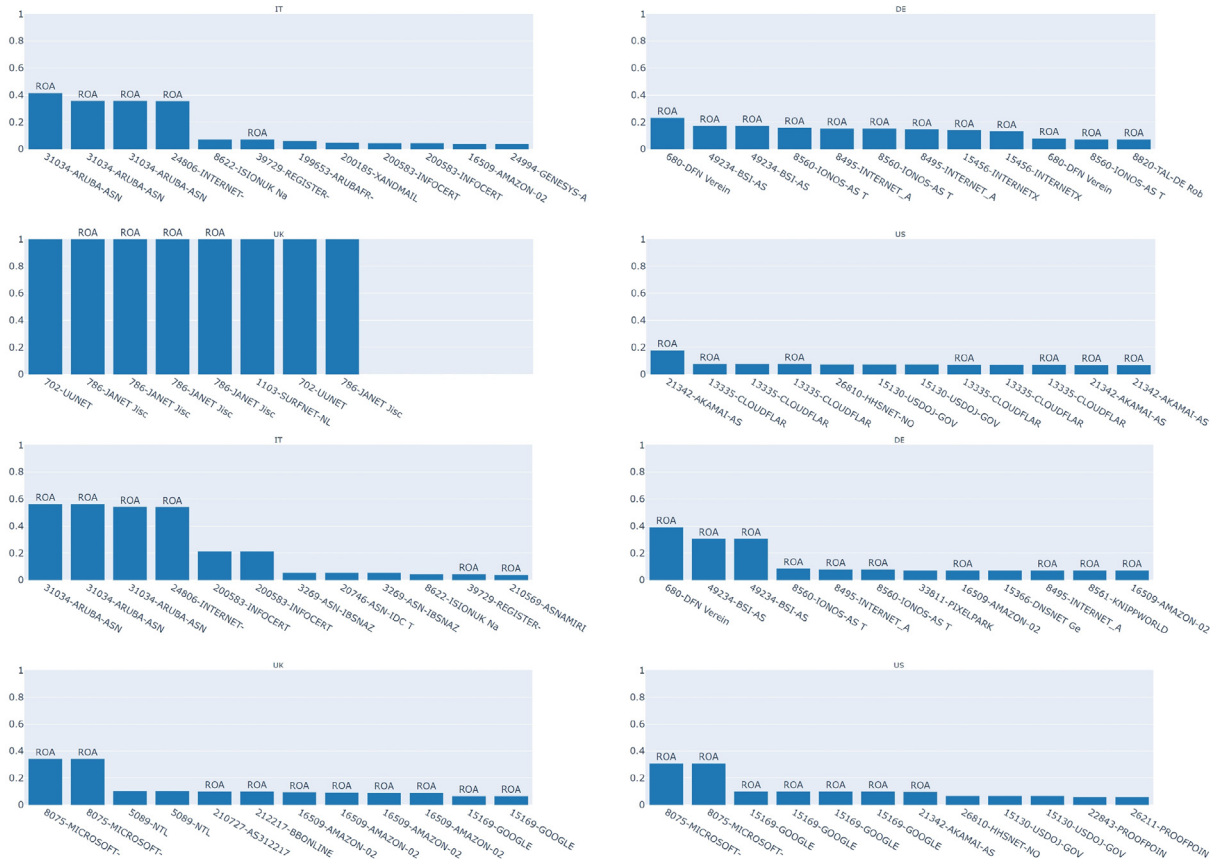

Fig. 7. Partial dependency from networks in mail domain resolution (up) and in mail server resolution (down). Networks are identified by the respective AS and those that are ROA-protected are annotated as such.

this respect, we analyzed ASes that fully or partly control resolution of mail domains or of mail servers. We analyze each of the four resulting combinations below.

Regarding *full control* over resolution of *mail domains*, we remark what follows (Fig. 8, up):

- In UK there is no mail domain that fully depends on a single AS.
- In IT, DE and US, each AS fully controls much less than 10% of mail domains (except for an AS by Akamai that fully controls 17.6% of mail domains in US).
- In US, 5 of the top twelve ASes that fully control some mail domains are managed by public organizations; no network in those AS is ROA-protected.
- In the top twelve AS list for DE, only the top-ranked AS is managed by a public organization (approx. 6.7% of mail domains); all networks of this AS are ROA-protected. Interestingly, the AS managed by the Federal Office for Information Security does not fully control any mail domain.
- In top twelve AS list for IT, only two ASes are managed by public organizations, with less than 1.5% of mail domains each. One of those ASes has no ROA-protected network, the other (managed by the italian academic and research community GARR) has more than half of its networks that are ROA-protected.

The scenario with respect to full control over resolution of *mail servers* is quite different (Fig. 8, down):

- In UK, there is a significant percentage of mail servers that fully depend on a single AS managed by a private company: 32.8%, 8.7%, 6.4% on an AS managed by Microsoft, Amazon and Google, respectively (all ROA-protected networks).
- In US there is a scenario similar to UK: 30.5%, 9.5%, 8.8% on an AS managed by Microsoft, Google and Akamai, again with all ROA-protected networks.
- in IT, 21.3% of mail servers fully depend from a single AS (with no ROA-protected network) and full dependency from ASes managed by public organizations is negligible.
- In DE we observe a much less pronounced concentration of full dependency of mail servers from a single AS: 6.7% from the AS managed by the public research institutions and much smaller amounts for other ASes.
- In all datasets, each AS managed by a public organization fully controls a negligible percentage of mail servers.
- In US and UK we observe a significant percentage of mail servers whose ASes are fully controlled by Microsoft and Google while in IT and DE such phenomenon is negligible.

Regarding *partial control* over resolution of *mail domains*, we remark what follows (Fig. 9, up):

- In UK, all mail domains partly depends on only three ASes. This fact is certainly the result of an explicit choice carefully designed and implemented. Two of those ASes are managed by public organizations but one of those is managed by a foreign country (NL). Two of those ASes have no ROA-protected network while the third one has more than half of its networks ROA-protected.
- In US, there is a much smaller concentration of partial dependency from single ASes: 17.8% of mail domains partly depend on an AS managed by Akamai while each of the other ASes partly controls less than 10% of mail domains.
- IT exhibits a strong concentration of partial dependency from two ASes, that partly control 45.5% and 35.7% of mail domains, respectively (the set of those mail domains could overlap). The remaining ASes partly control less than 7% of mail domains each.

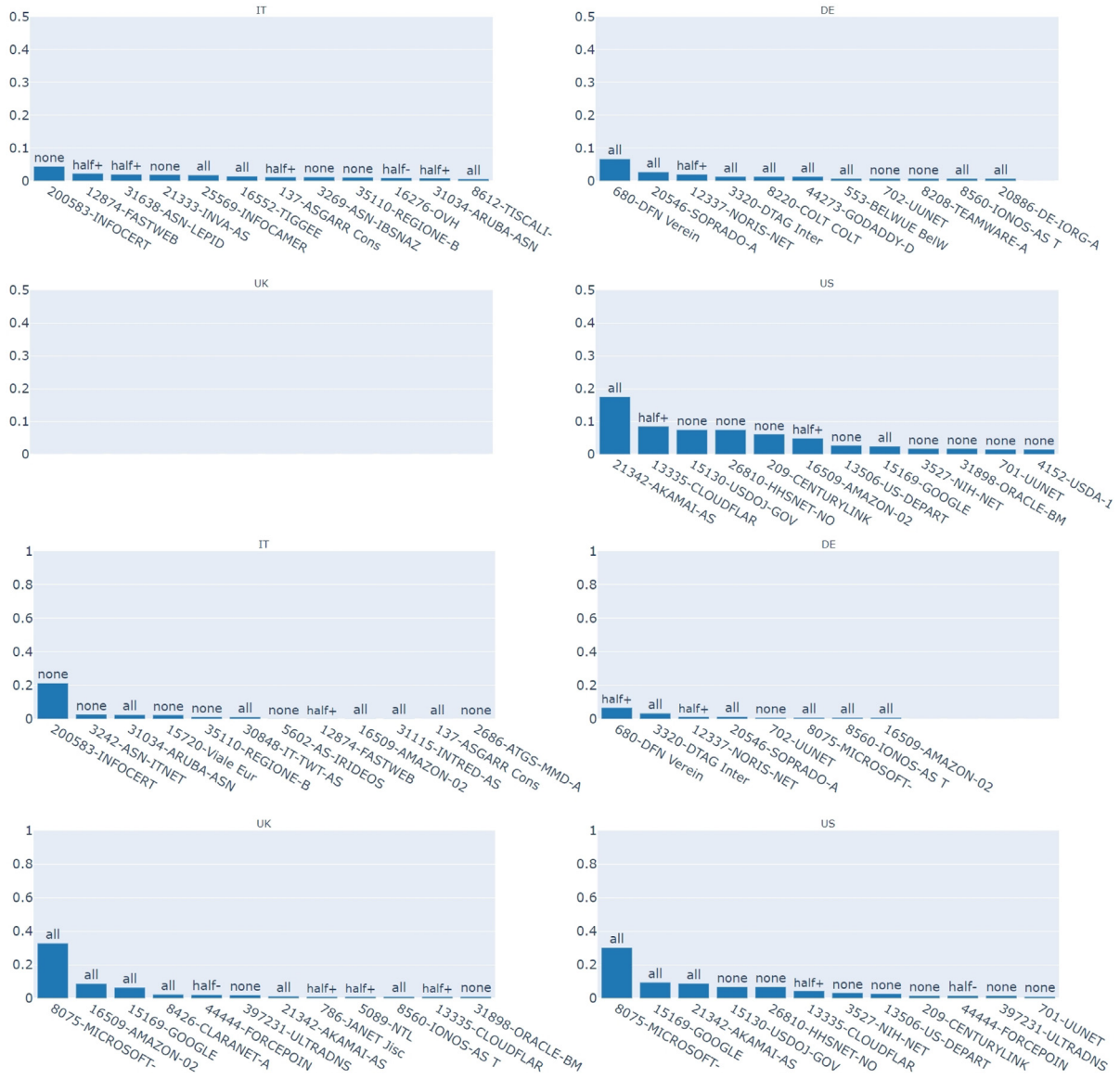

Fig. 8. Full dependency from ASes in mail domain resolution (up) and in mail server resolution (down). ASes are annotated with a label indicating the percentage of their respective networks that are ROA-protected.

- DE exhibits a pattern somewhat in between IT and US. While the AS managed by the Federal Office for Information Security does not fully control any mail domain (as observed above), it partly controls 17.8% of mail domains and is the AS ranked third in this respect.
- DE is the only dataset with a strong ROA usage: the top 6 ASes in this dataset have all networks that are ROA-protected. Usage of ROA in the other datasets is much smaller.
- The presence of ASes managed by public organizations in the analyzed distributions is almost negligible, with the only exception of DE: the top-ranked AS is managed by the public research institution while the third-ranked one by the Federal Office for Information Security, that partly control 32.7% and 17.3% of mail domains respectively. The distribution for US contain three ASes managed by public organizations, that partly control in between 7.5% and 2.7% each. Interestingly, in DE, all the networks of those ASes are ROA-protected, while in US no network of those ASes is ROA-protected.

Finally, regarding partial control over resolution of *mail servers* (Fig. 9, down):

- In UK the distribution of partial dependency is very different from the one for mail domains: 34% of mail servers partly depend on an AS managed by Microsoft while the other ASes partly control less than 10% of mail servers each. None of the ASes in the top 12 list are managed by a public organization.
- In IT, US and DE, the distribution for mail servers is instead quite similar to the one for mail domains, the key difference being a much stronger concentration of partial dependency from the top ranked ASes.
- In all datasets, ROA usage tends to be higher than for mail domains.
- The presence of ASes managed by public organizations is almost negligible, similar to the pattern observed for mail domains. However, in DE, there is a stronger concentration: the AS managed by the public research institution and the one managed by the Federal Office for Information Security partly control 40.3% and 30.7% of mail servers respectively (as opposed to 34% and 10% of mail domains).

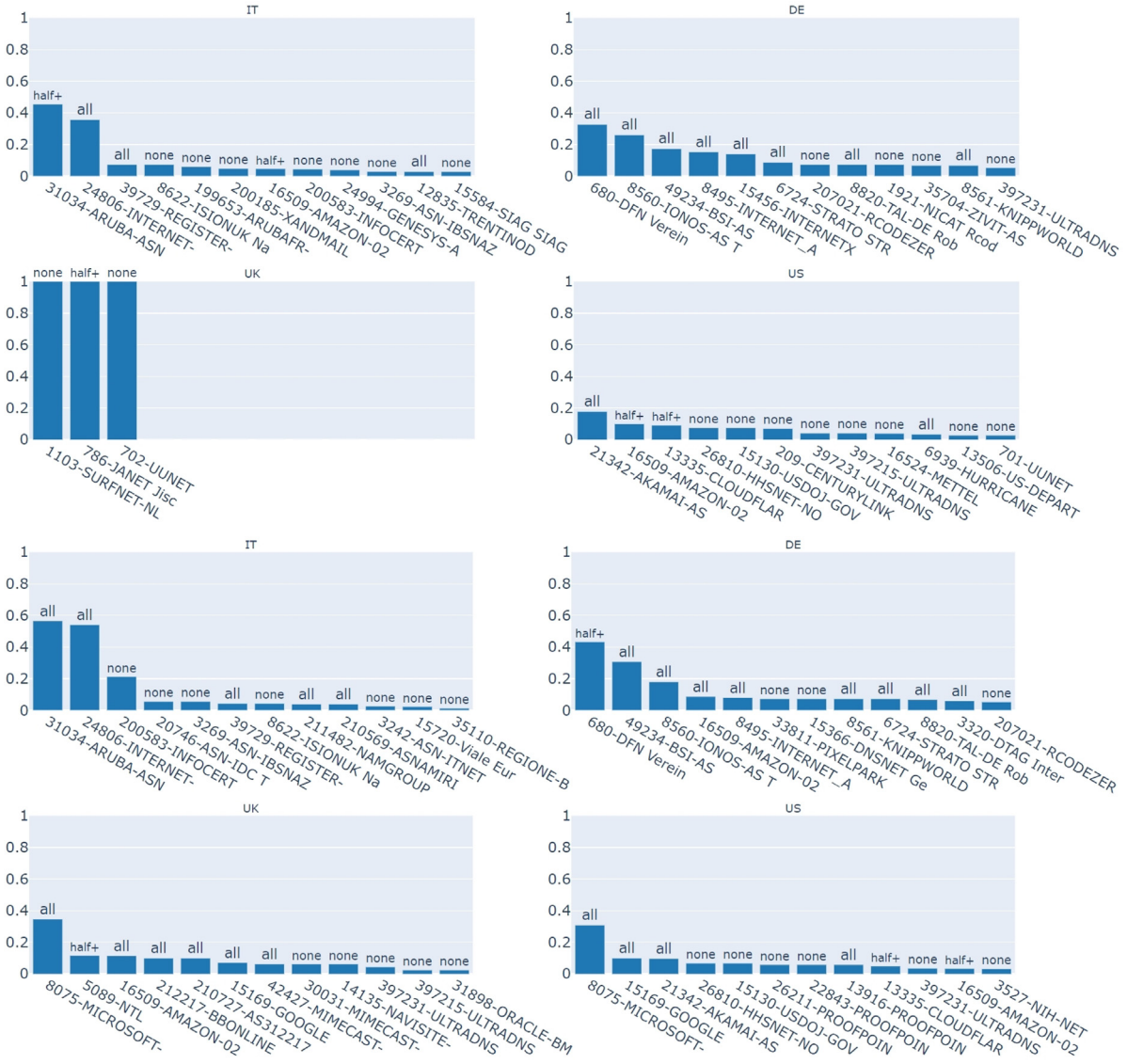

Fig. 9. Partial dependency from ASes in mail domain resolution (up) and in mail server resolution (down). ASes are annotated with a label indicating the percentage of their respective networks that are ROA-protected.

3.6.4. Findings (Dependency from groups of networks)

For brevity, we focus only for groups that fully control more than 9.5% of mail domains or mail servers in the respective dataset: there is no significant dependency from groups of networks beyond those listed in Table 3. For each such group we determined the number of networks, the number of ROA-protected networks and the ASes that manage those networks.

Regarding resolution of *mail domains*, we observe what follows (Table 3, up):

- All the groups are redundant in terms of ASes.
- All the groups have all their networks ROA-protected, with the only interesting exception of UK in which only half of the networks are ROA-protected.
- In UK, as already observed in the previous sections, *all* mail domains depend on a single (highly redundant) group of networks.
- In the other datasets there is a significant concentration only in a group of the IT dataset. There are only 2 further groups that fully control more than 10% of mail servers in the respective dataset (DE).

The scenario for resolution of *mail servers* is somewhat different (Table 3, down):

- Only three of the seven groups are redundant in terms of ASes. The single AS of the UK and US groups is managed by Microsoft or Google, while the one of the IT group is managed by INFOCERT (a company specialized in the trust solutions and digitalization of business processes).
- Not all the groups have all their networks ROA-protected: a UK group has half of its networks ROA-protected while an IT group has none (this group is the one with no AS redundancy where the AS is managed by INFOCERT).
- In UK there is a group that fully controls a significant percentage of mail servers yet it is much less redundant than the group that fully controls mail domains.
- IT, DE, US exhibit higher concentration than for mail domains (there was no group in US that fully control more than 9.5% of mail domains).

It is also important to remark that the ASes in Table 3 play a crucial role in the respective countries well beyond the groups of networks listed here: as previously observed in Fig. 9, those ASes partly control many more mail domains and mail servers and are indeed top ranked in the respective distributions.

Table 3
Groups of networks that fully control name resolution of more than 9.5% of mail domains (up) and of mail servers (down).

Country	Percentage in country	#Networks/ #networks with ROA	AS number and description
IT	0.356	4/4	31034-ARUBA-ASN 24806-INTERNET-CZ
DE	0.173	3/3	49234-BSI-AS 680-DFN-Verein
DE	0.133	6/6	15456-INTERNETX 8560-IONOS-AS-T 8495-INTERNET_A
UK	1	8/4	1103-SURFNET-NL 702-JANET 786-UUNET
IT	0.54	4/4	331034-ARUBA-ASN 24806-INTERNET-CZ
IT	0.213	2/0	200583-INFOCERT
DE	0.307	3/3	49234-BSI-AS 680-DFN-Verein
UK	0.324	2/2	8075-MICROSOFT
UK	0.098	4/2	210727-BUYLIMITED 5089-NTL 212217-BBONLINE
US	0.282	2/2	8075-MICROSOFT
US	0.095	4/4	15169-GOOGLE

3.7. Distribution of access paths

3.7.1. Methodology

We analyzed the distribution of *access paths* for mail domains across networks and autonomous systems. We used the same approach as the one for name resolution paths, modified by replacing nameservers by mail servers. This analysis allows determining the number of mail domains that would be affected by an attack on a single entity (i.e., a network or an autonomous system), as well as to identify those entities that are potentially more critical for a country as a whole.

Given a mail domain d, we say that d *fully depends in access* on a network n if all the mail servers of d are in n; we say that d *partly depends in access* on n if at least a mail server of d is in n. Conversely, we say that n *fully* or *partly controls access to* d when d fully or partly depends in access on n. For each network we determined, separately for each dataset, the percentage of mail domains that fully or partly depend in access by that network. We extend to autonomous systems the above definitions made for networks and computed the analogous distributions for autonomous systems.

3.7.2. Findings (Dependency from networks)

Regarding *full dependency* from *networks* in access paths (Fig. 10, up):

- Full dependency in access from a single network is not very common with the only notable exception of IT, in which a remarkable 21.3% of mail domains have all their mail servers concentrated in the same network. Interestingly, this network is not ROA-protected. There is only one other network that fully controls in access more than 10% of mail domains in the respective dataset (DE); that network is ROA-protected
- Full dependency is particularly infrequent in UK and US, where the top network of the distributions fully controls only 3.1% and 2.4% of mail domains, respectively.

Partial dependency in access follows a pattern similar to the one for full dependency (Fig. 10, down):

- Partial dependency in access from a single network is not very common with the only notable exception of IT and to a lesser extent DE. Each of these datasets has two networks that partly control a significant amount of mail domains: 51.4% and 29.3% respectively, much more than the amount of mail domains fully controlled by the respective top networks (21.3% and 10%)
- Unlike full dependency, there are two networks also in US that partly control in access a significant amount of mail domains (21.4%). As it turns out from the analysis in terms of groups of networks (Table 4 below), these two networks form a group containing all the mail servers of the same mail domains and that thus *fully* controls in access a significant amount of mail domains.

Table 4
Groups of networks that fully control access paths of more than 9.5% of mail servers.

Country	Percentage in country	#Networks/ #networks with ROA	AS number and description
IT	0.54	2/2	31034-ARUBA-ASN
DE	0.293	2/2	49234-BSI-AS
UK	0.098	2/0	5089-NTL
US	0.202	2/2	8075-MICROSOFT

3.7.3. Findings (Dependency from autonomous systems)

Dependency from *autonomous systems* in access paths is summarized in Fig. 11. For simplicity, we provide only the distribution for *full dependency*: most access paths are concentrated in a single autonomous system (Fig. 3), thus the distribution for partial dependency is very similar and does not provide any additional insights.

- The top ASes in IT fully control in access 57.1% and 21.3% of mail domains, respectively. This is by far the strongest concentration in terms of ASes across the datasets.
- In each of the other datasets, the top AS fully controls in access around 30% of mail domains and the distribution decades rapidly, from the second AS.
- Regarding ROA usage, in IT, the top AS has more than half of its networks ROA-protected while the second AS has none. The top AS in DE, UK, US has all of its networks ROA-protected.
- The presence of ASes managed by public organizations is very marginal in all the datasets, with the only notable exception of DE: the top AS is managed by the Federal Office for Information Security (29.3% of mail domains) and the third one by the public research organization (8.7%).
- The top AS in US and UK is the same. This AS is managed by Microsoft and corresponds to a cloud-hosted email service. The AS of the analogous service by Google is ranked at the fourth and at the second place in UK and US, respectively. Both ASes have all of their networks ROA-protected.
- The two previously mentioned ASes are the *only* ASes in the top 12 positions for US that use ROA.

Finally, we examined the distribution of access paths across *groups* of networks: we determined the groups of 2 or more networks that *fully control* a given mail domain, i.e., that contain all the mail servers for a mail domain. Similar to what observed for name resolution paths, there is a significant dependency only from few groups of networks (Table 4). The properties of these groups are very similar to those observed for name resolution paths, except that in this case the groups are all composed of only two networks. Interestingly, none of these groups exhibits redundancy in terms of ASes.

4. Discussion

4.1. Research questions

We summarize below the key findings relevant to each research question listed in the Introduction and duplicated here for ease of reading.

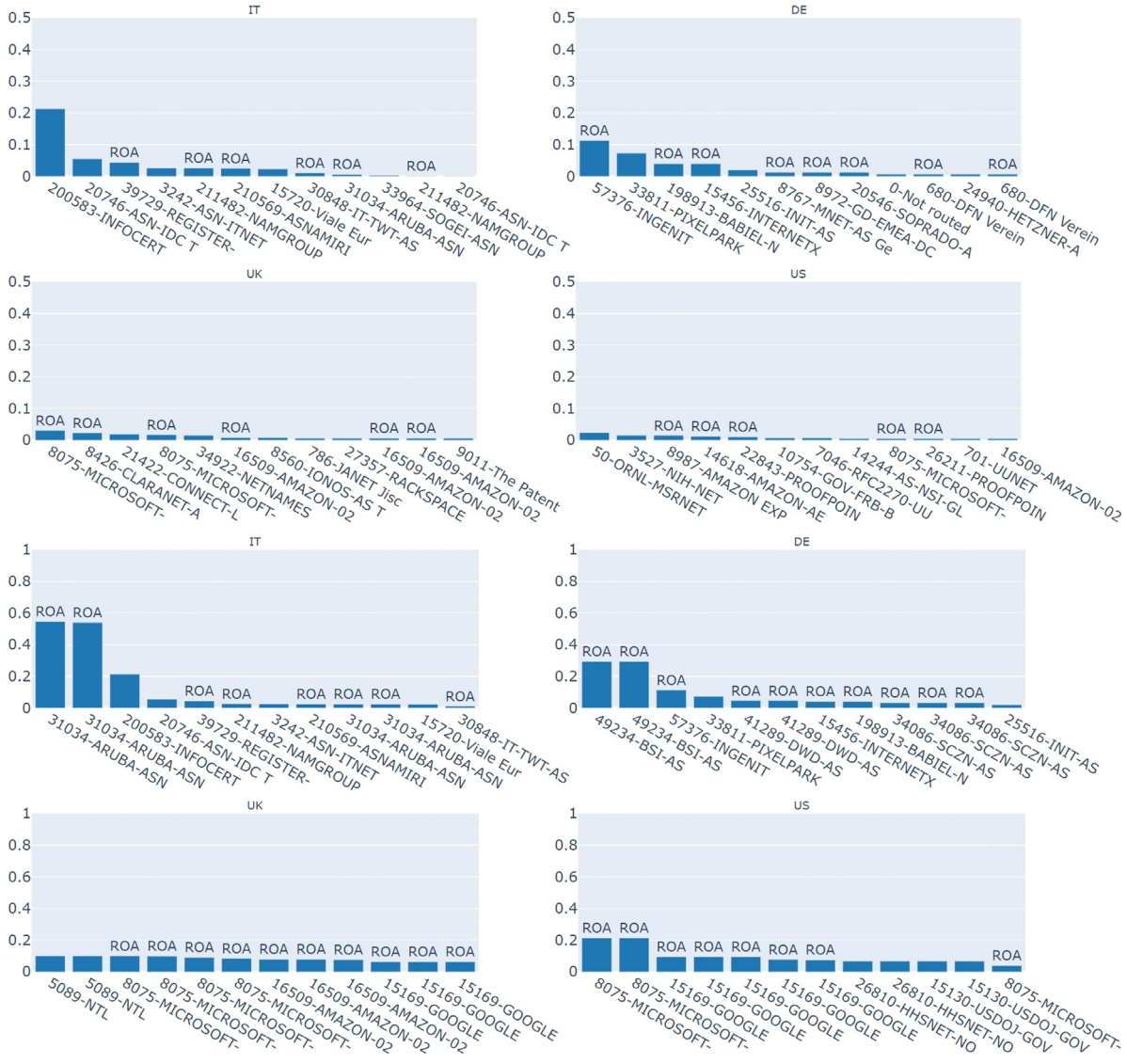

Fig. 10. Dependency from networks in access paths: full dependency up, partial dependency down (note the different scale on the Y axis). Networks are identified by the respective AS and those that are ROA-protected are annotated as such.

4.1.1. RQ1

How many mail domains of the public administration of a whole country could be affected by a network attacker capable of controlling a single IP address range or autonomous system? How are those attack opportunities distributed along the various steps of mail delivery, i.e., mapping a mail domain name to a mail server name, mapping the latter to an IP address, communicating with that address?

- *Full* dependency of *name resolution* from a single network is very rare in all datasets: taking control of a single network, thus, may allow controlling all name resolution paths of just a tiny fraction of mail domains or mail servers.
- *Full* dependency of *mail domain resolution* from a single AS is rare in IT, DE, US (much less than 10% of all mail domains) and absent in UK. Regarding *mail server resolution*, such a full dependency is rare in DE but is significant in UK, US, IT (20%–30% of all mail domains).
- *Partial* dependency of *name resolution* from a single network is somewhat frequent: in each dataset there are a few networks that partly control either mail domain or mail server resolution for a significant percentage of the respective mail domains.
- A single *group of networks* that *fully* control *name resolution* for a large percentage of the respective mail domains is present in all the datasets (with the only exception of mail domain resolution in the US dataset). Regarding mail domain resolution, such groups are highly redundant even in terms of ASes and have all of their networks ROA-protected. Regarding mail server resolution, not all such groups are redundant in terms of ASes and not all of their networks are ROA-protected; some of those groups (in US and UK) correspond to providers of cloud-based email services (Google and Microsoft) and do not have any AS redundancy.
- *Full* dependency of *access path* from a single network is rare: taking control of a single network, thus, may allow controlling all access paths to mail servers of just a tiny fraction of mail domains in the respective dataset (with the exception that the IT and DE datasets each have a network that fully controls in access path more than 10% of the respective mail domains).
- *Partial* dependency of *access path* from a single network is also rare (again, with the exception that the IT and DE datasets each have two networks that partly control in access path more than 30% of the respective mail domains).
- *Full* dependency in *access path* from a single AS for more than 30% of the respective mail domains is present in all the datasets (with

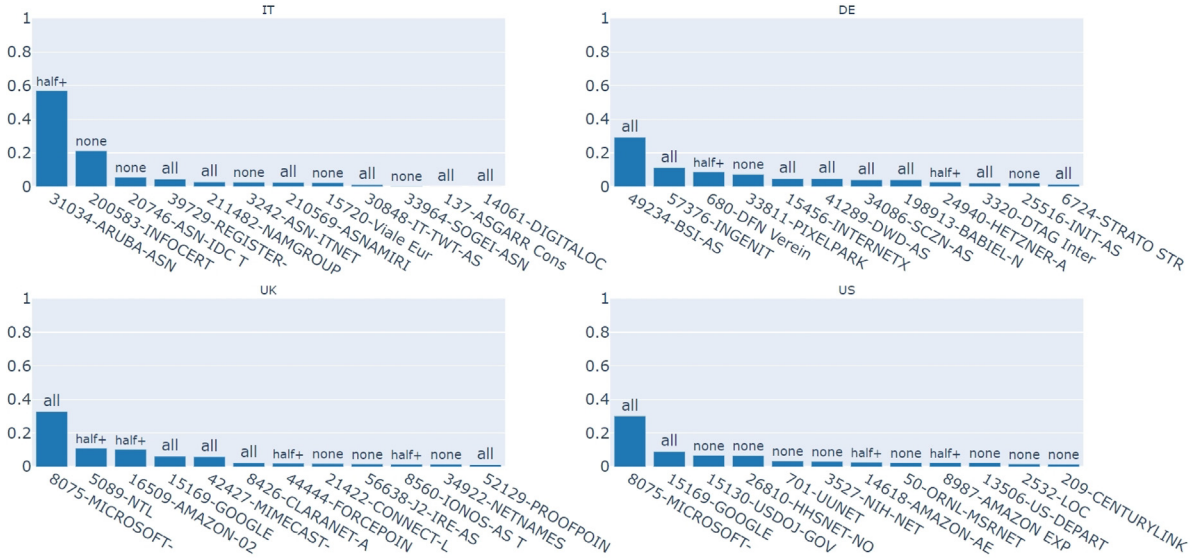

Fig. 11. Full dependency from ASes in access paths. ASes are annotated with a label indicating the percentage of their respective networks that are ROA-protected.

a strong concentration in the IT dataset, where the two top ASes fully control 57.1% and 21.3% of mail domains each).

- Overall, the UK dataset exhibits an extremely small security perimeter of *mail domain resolution*, with *only one zone* involved. This dataset has a much higher number of zones involved in the resolution of mail server names, though. The IT dataset exhibits the opposite approach to that of the UK: almost one zone for each mail domain but a much smaller number of zones for resolving the names of mail servers; and, the smallest amount of entities involved in access paths to mail server across all the datasets.

4.1.2. RQ2

What redundancy levels tend to be used for the email infrastructure of a given country? Do such levels tend to be identical at all the steps of mail delivery or different redundancy levels tend to be used at each step?

- Regarding *name resolution* paths, in each dataset the redundancy patterns for the resolution of mail domains and of mail servers are *quite different from each other* (except for the DE dataset). Furthermore:
  - The UK dataset exhibits the highest redundancy for the resolution of mail domain names: all mail domains exceed the robustness requirement both for ASes and for networks (i.e., the respective nameservers are replicated over three or more distinct ASes or networks).
  - US and UK exhibit a strong prevalence of name resolution paths for mail server names where the respective nameservers are spread over two networks in the same AS.
  - PEC providers in the IT dataset do not tend to enforce AS redundancy for mail domains more than DNS providers of public administrations.
  - Some 10% of mail domains in the US dataset have the respective name servers concentrated in a single network.
- Regarding *access paths*, their redundancy is *much smaller* than redundancy of name resolution paths—a large amount of mail access paths are replicated over only 1 or 2 networks. Furthermore:
  - There is almost no redundancy of access paths at the AS level.
  - Redundancy of access paths at the network level is much higher for US and UK than for IT and DE.
  - Redundancy of access paths is much higher than redundancy of access paths to web servers of the public administrations of the respective country.

4.1.3. RQ3

Autonomous systems that support the email infrastructure of a given country tend to be managed by government organizations or by private companies?

- *Full* dependency from ASes managed by *public* organizations is rare in mail domain resolution and almost negligible in mail server resolution. Furthermore:
  - The distribution of ASes managed by public organizations in the lists of top twelve ASes is remarkably different among the datasets: IT, two/three ASes (less than 1.5% of mail domains or mail servers each); DE, the top-ranked AS (less than 10% of mail domains or mail servers); US, approximately half of the ASes; UK, only one (2% of mail servers; no AS fully controls name resolution of mail domains).
  - ASes managed by private organizations specialized in DoS protection are very frequent in US and in UK (only for resolution of mail servers) and almost absent in DE and IT.
- *Partial* dependency in name resolution is distributed over a mix of ASes managed by public organizations and ASes managed by private organizations (except for the IT dataset in which the presence of public organizations is negligible). Furthermore:
  - An AS managed by the Federal Office for Information Security partly controls name resolution for 17.3% and 30.7% of mail domains and mail servers, respectively. Such AS is not in the top twelve list for full control, though.
  - ASes managed by public research and education organizations play a key role in DE (top ranked AS for both mail domains and mail servers) and UK (all mail domains).
- Regarding *access paths*, full control by ASes managed by public organizations is very marginal. The only notable exception is the DE dataset, where the AS managed by the Federal Office for Information Security is top ranked and fully controls 29.7% of mail domains.

4.1.4. RQ4

Are there any architectural patterns or design choices that are common across different countries?

- There are *far more* architectural *differences* between the analyzed datasets than there are similarities. This result applies at each of the three steps, i.e., resolution of mail domains, resolution of mail servers, access to mail servers. The main key similarities may be summarized as follows:
  - There is *no network* that *fully* controls name resolution of either mail domains or of mail servers for an appreciable amount of mail domains (more than 1%).
  - There are 2–4 ASes that *partly* control name resolution of mail servers for a significant amount of mail domains.
  - The occurrence of nameservers not replicated or concentrated in a single network/AS is much smaller than the average of government domains worldwide and similar to domains of government websites in the respective country.
  - A significant amount of mail domains have *all* their mail servers concentrated in a group of two networks managed by a single AS.
  - Most ASes are managed by *domestic* organizations. ASes of companies specialized in DoS protection (e.g., Akamai, Cloudflare, Google, Microsoft) are mostly used in the US and, to a lesser extent, UK.
  - ROA usage is significantly *higher than the average* in the respective country. The amount of such usage is very similar for name resolution paths and for access paths, both at the level of networks and of ASes.

4.1.5. RQ5

What is the actual deployment of ROA in networks and autonomous systems responsible for the email infrastructure of the 4 countries in our dataset?

- For *more than half* of the mail domains in each dataset, there is *no* ROA usage at all. This result applies at each of the three steps, i.e., resolution of mail domains, resolution of mail servers, access to mail servers. The only key exceptions are:
  - Resolution of mail domains in the UK dataset, where a significant percentage of mail domains has more than half of the respective networks ROA-protected.
  - US dataset, where more than 30% of mail domains have all the respective networks ROA-protected at each of the three steps.
- In each dataset, there is no significant difference in ROA usage across name resolution paths and access paths.
- In each dataset, the percentage of ROA-protected networks is significantly higher than the percentage of ROA-protected IP addresses.
- In terms of protection of mail services, distribution of ROA protection across networks in UK and US is much more efficient than in IT and DE.
- ASes tend to provide ROA either for *all* of their networks or for *none* of them.

4.2. Centralized planning

Broadly speaking, it seems fair to claim that the recommended way for structuring government mail services at the level of a full country should be characterized by: (i) high redundancy, even at the AS level; and, (ii) dependency distributions across networks concentrated in very few organizations. In particular, concentration in a few organizations is important for clearly identifying and defending the security perimeter, as well as for enforcing specific technical requirements, e.g., security technologies, country-wide. Naturally, those organizations should be technically capable of defending their perimeter adequately, of offering significant DoS protection and, most importantly, they should be reliable from a strategic point of view.

The above guidelines should be implemented at each of the three stages involved in network access to a mail domain (resolution of mail domain name, resolution of mail server name, access to mail server), as the opportunities for network attackers are also distributed along these stages. While analysis of attack costs and of attack opportunities beyond those considered in this work (e.g., attacks to nameservers) could justify slightly different architectural choices for the three steps, we believe that the above guidelines should be valid at each of them, with the remark that the redundancy degree in access paths could be smaller than in name resolution paths (also because of the technical difficulties in synchronizing mail server replicas).

As our analysis shows, all the four countries tend to approximate the above guidelines, albeit in very different degrees and generally remaining quite far from them. What we find particularly relevant in this respect is the strong architectural difference between name resolutions of mail domains and of mail servers in UK: while the former is practically coincident with the architecture described above, the second is only a rather distant approximation. In other words, while it has been practically feasible to enforce a certain architecture for the resolution of mail domains country-wide, a similar effort has not been done for the resolution of mail servers, despite the two scenarios being almost equivalent from the point of view of a network attacker.

There are few signals that suggest a centralized planning of the architecture. Furthermore, such signals are quite heterogeneous across the analyzed countries and are important examples of differences among those countries.

- UK name resolution of mail domains: direct dependence of *all* of them from a *single* zone; strong replication of nameservers (across eight distinct networks and three ASes); more than half of those networks ROA-protected (37% of the IP address space in the UK is ROA-protected). The heterogeneous nature of the three ASes is interesting: a public domestic organization, a private domestic organization, a public foreign organization.
- US: in all the three kinds of paths, percentage of mail domains with ROA protection *higher* than in the other datasets (except for name resolution paths of UK mail domains), despite the percentage of ROA-protected networks and ASes being significantly smaller. This fact is quite important, as it suggests a way for obtaining higher global protection even with a smaller global overhead: incentivizing usage of ROA protection in networks that host services widely used or strategically important. This consideration is related, to some extent, to the broader idea of promoting ROA usage in regions of the Internet where activities of users occur (*zone of trust*) [50,51].
- US: strong presence of private organizations specialized in DoS protection and cloud services in top networks and ASes of dependency distributions. The same applies to UK, only for resolution of mail servers and for access paths.
- DE: ROA protection in *nearly all* the networks in the top positions of dependency distributions, both in name resolution paths and in access paths.
- DE: top ASes in the distribution of name resolution paths managed by *public* organizations, including the Federal Office for Information Security.
- DE: *significant concentration* in the top two or three networks in distributions of partial dependency, both in name resolution paths and in access paths; nearly all these networks in ASes managed by *public* organizations; nearly all these networks ROA-protected.
- IT: *strong concentration* in the top two or three networks in distributions of partial dependency and in one or two groups of networks with full dependency, both in name resolution paths and

in access paths; all these networks in ASes managed by *private* companies. All these networks are ROA-protected (with the only exception of the top network in the distribution of access paths, that is fully responsible for almost 20% of mail domains).
- IT: much higher amount of zones for mail domains than for mail servers; ratio #mail domains/#mail servers much higher than in the other datasets. These facts are exactly the opposite of what observed in UK; they do not result from any architectural planning and are instead a side-effect of the technical and legal requirements of the Italian certified email system (PEC), though.
- IT: almost all top networks/ASes in dependency distributions managed by *private* companies. No other dataset exhibits a similar absence of public organizations. We interpret this fact as a consequence of strategic decisions taken centrally.

Strong concentration of dependencies in a network or AS might not result from centralized planning, e.g., it could be a result of market forces. The large amount of such dependencies along with the domestic nature of the respective ASes suggest that centralized coordination, perhaps indirectly through forms of incentivization, has played an important role, though.

5. Concluding remarks

There are many opportunities for localized network attacks that can have global impact on the daily interactions of citizens and businesses with IT government services, thereby affecting security and availability of such important pieces of critical infrastructure. The network architecture of IT government services country-wide is thus an issue that deserves special consideration, as the recent COVID pandemic and the return to times of great tension in relations between states have shown. After many years of decentralized and uncoordinated diffusion of such services, security considerations are increasingly calling for more centralized and coordinated planning and implementation.

Our detailed analysis of the network architecture of mail domains in four countries has shown that all these countries appear to have introduced elements of centralized coordination, with varying degree of depth and pervasiveness. The overall scenario, though, is still characterized more by the architectural differences between countries than by their similarities. Furthermore, each country exhibits important architectural incoherences among the various steps required for communicating with a mail domain. Although the ubiquitous technological trend towards cloud services suggests that such inconsistencies could be mitigated in the future, it seems fair to say that the lack of a common model for the network architecture of government IT services nationwide, coupled with the enormous difficulties to implement concretely any centrally defined project in environments that cannot be rebuilt from scratch and that involve many different actors, will remain a crucial problem for the foreseeable future. We believe that our analysis can contribute to the understanding and awareness of this important issue.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The author is grateful to Leonardo Simonini, Lorenzo Fabbio, Federico Boni for discussions about the dependency graph and for their help in implementing the data collection software. The author is also grateful for Martino Trevisan for his comments on an early draft of this work.

References

[1] S. Rose, J.S. Nightingale, S. Garfinkel, R. Chandramouli, Trustworthy email, Tech. Rep. NIST SP 800-177r1, National Institute of Standards and Technology, Gaithersburg, MD, 2019.
[2] Z. Durumeric, D. Adrian, A. Mirian, J. Kasten, E. Bursztein, N. Lidzborski, K. Thomas, V. Eranti, M. Bailey, J.A. Halderman, Neither snow nor rain nor MITM...: An empirical analysis of email delivery security, in: Proceedings of the 2015 Internet Measurement Conference, 2015.
[3] D. Tatang, F. Zettl, T. Holz, The evolution of DNS-based email authentication: Measuring adoption and finding flaws, in: 24th International Symposium on Research in Attacks, Intrusions and Defenses, 2021.
[4] D. Tatang, R. Flume, T. Holz, Extended abstract: A first large-scale analysis on usage of MTA-STS, in: L. Bilge, L. Cavallaro, G. Pellegrino, N. Neves (Eds.), Detection of Intrusions and Malware, and Vulnerability Assessment, Springer International Publishing, Cham, 2021, pp. 361–370.
[5] People's Republic of China state-sponsored cyber actors exploit network providers and devices, 2022, CISA - Cybersecurity and Infrastructure Security Agency.
[6] D. Palmer, These hackers broke into 10 telecoms companies to steal customers' phone records, 2019, ZDNET.
[7] J. Harries, D. Mayer, A roaming threat to telecommunications companies, 2021, crowdstrike.com.
[8] C. Bing, R. Satter, C. Bing, Ukrainian telecom company's internet service disrupted by 'powerful' cyberattack, 2022, Reuters.
[9] C. Cimpanu, Belgium's government network goes down after massive DDoS attack, 2021, The Record By Recorded Future.
[10] J. Burt, Israeli govt sites hit by huge DDoS attack, 2022, The Register.
[11] C. Cimpanu, DDoS attacks hit multiple email providers, 2021, The Record By Recorded Future.
[12] B. Kreb, Hacked cameras, DVRs powered today's massive internet outage, 2016, Krebs on Security.
[13] D. Goodin, Foul-mouthed worm takes control of wireless ISPs around the globe, 2016, Ars Technica.
[14] C. Cimpanu, Ransomware gang demands $7.5 million from Argentinian ISP, 2020, ZDNET.
[15] C. Cimpanu, Hackers breached A1 telekom, Austria's largest ISP, 2020, ZDNET.
[16] ANSA Agenzia, Tiscali: attacco hacker provoca disservizi web e rete fissa -Sardegna, 2021, Agenzia ANSA, Section: Sardegna.
[17] J. Greig, French telecom company La Poste Mobile struggling to recover from ransomware attack, 2022, The Record By Recorded Future.
[18] C. Cimpanu, DNS hijacks at two cryptocurrency sites point the finger at GoDaddy, again, 2021, The Record By Recorded Future.
[19] W. Mercer, DNS hijacking abuses trust in core internet service, 2019, Talos Intelligence.
[20] C. Cimpanu, Hackers breached Greece's top-level domain registrar, 2019, ZDNET.
[21] M. Hirani, S. Jones, B. Read, Global DNS hijacking campaign: DNS record manipulation at scale, 2019, Mandiant.
[22] P. Rascagneres, DNSpionage campaign targets middle east, 2018, Talos Intelligence.
[23] P. Litke, J. Stewart, BGP hijacking for cryptocurrency profit, 2014, Dell SecureWorks Counter Threat Unit.
[24] M. Apostolaki, A. Zohar, L. Vanbever, Hijacking bitcoin: Routing attacks on cryptocurrencies, in: 2017 IEEE Symposium on Security and Privacy, SP, 2017, pp. 375–392.
[25] H. Birge-Lee, Y. Sun, A. Edmundson, J. Rexford, P. Mittal, Bamboozling certificate authorities with BGP, in: 27th USENIX Security Symposium, USENIX Security 18, 2018, pp. 833–849.
[26] C.C. Demchak, Y. Shavitt, China's maxim – leave no access point unexploited: The hidden story of China telecom's BGP hijacking, Mil. Cyber Aff. 3 (1) (2018) 7.
[27] C. Cimpanu, Russian telco hijacks internet traffic for Google, AWS, Cloudflare, and others, 2020, ZDNET.
[28] D. Goodin, How 3ve's BGP hijackers eluded the Internet—and made $29M, 2018, Ars Technica.
[29] F. Douzet, L. Pétiniaud, L. Salamatian, K. Limonier, K. Salamatian, T. Alchus, Measuring the fragmentation of the Internet: The case of the border gateway protocol (BGP) during the Ukrainian crisis, in: 2020 12th International Conference on Cyber Conflict, Vol. 1300, CyCon, ieeexplore.ieee.org, 2020, pp. 157–182.
[30] D. Madory, BGP / DNS hijacks target payment systems, 2018, Oracle Internet Intelligence.
[31] K. Kirkpatrick, Fixing the internet, Commun. ACM 64 (8) (2021) 16–17.
[32] W. Haag, D. Montgomery, A. Tan, W. Barker, Protecting the Integrity of Internet Routing: Border Gateway Protocol (BGP) Route Origin Validation, Tech. Rep. NIST Special Publication (SP) 1800-14, National Institute of Standards and Technology, 2019.

[33] G. Huston, Measuring ROAs and ROV, 2021, https://labs.apnic.net/?p=1420. (Accessed 26 March 2021).
[34] C. Testart, P. Richter, A. King, A. Dainotti, D. Clark, To filter or not to filter: Measuring the benefits of registering in the RPKI today, in: A. Sperotto, A. Dainotti, B. Stiller (Eds.), Passive and Active Measurement, in: Lecture Notes in Computer Science, Springer International Publishing, Cham, 2020, pp. 71–87.
[35] T. Chung, R. van Rijswijk-Deij, B. Chandrasekaran, D.R. Choffnes, D. Levin, B.M. Maggs, A. Mislove, C. Wilson, A longitudinal, end-to-end view of the DNSSEC ecosystem, in: USENIX Security Symposium, 2017.
[36] M. Wander, Measurement survey of server-side DNSSEC adoption, in: Network Traffic Measurement and Analysis Conference, TMA, 2017, pp. 1–9.
[37] Y.-D. Song, A. Mahanti, S.C. Ravichandran, Understanding evolution and adoption of top level domains and DNSSEC, in: 2019 IEEE International Symposium on Measurements & Networking, M&N, 2019, pp. 1–6.
[38] G. Huston, DNS trends, Internet Protoc. J. 24 (2021) 2–17.
[39] Y. Yu, D. Wessels, M. Larson, L. Zhang, Authority server selection in DNS caching resolvers, Comput. Commun. Rev. 42 (2012) 80–86.
[40] C.A. Shue, A.J. Kalafut, Resolvers revealed: Characterizing DNS resolvers and their clients, ACM Trans. Internet Techn. 12 (2013) 14:1–14:17.
[41] K. Schomp, O. Bhardwaj, E. Kurdoglu, M. Muhaimen, R.K. Sitaraman, Akamai DNS: Providing authoritative answers to the world's queries, in: Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication, 2020.
[42] M. Allman, Comments on DNS robustness, in: Proceedings of the Internet Measurement Conference 2018, IMC '18, ACM, New York, NY, USA, 2018, pp. 84–90.
[43] L. Zembruzki, A.S. Jacobs, G.S. Landtreter, L.Z. Granville, G.C.M. Moura, Measuring centralization of DNS infrastructure in the wild, in: Advanced Information Networking and Applications, Springer International Publishing, 2020, pp. 871–882.
[44] A. Bartoli, Robustness analysis of DNS paths and web access paths in public administration websites, Comput. Commun. 180 (2021) 243–258.
[45] J. Jiang, J. Zhang, H. Duan, K. Li, W. Liu, Analysis and measurement of zone dependency in the domain name system, in: 2018 IEEE International Conference on Communications, ICC, 2018, pp. 1–7.
[46] R. Houser, S. Hao, C. Cotton, H. Wang, A comprehensive, longitudinal study of government DNS deployment at global scale, in: 2022 52nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks, (DSN), 2022, pp. 193–204.
[47] T. Chung, E. Aben, T. Bruijnzeels, B. Chandrasekaran, D. Choffnes, D. Levin, B.M. Maggs, A. Mislove, R. van Rijswijk-Deij, J. Rula, N. Sullivan, RPKI is coming of age: A longitudinal study of RPKI deployment and invalid route origins, in: Proceedings of the Internet Measurement Conference, IMC '19, Association for Computing Machinery, New York, NY, USA, 2019, pp. 406–419.
[48] NIST RPKI Monitor, 2022, URL https://rpki-monitor.antd.nist.gov/ROV.
[49] ROA Use World Map, 2022, URL https://stats.labs.apnic.net/roas.
[50] D. Clark, K. Claffy, C. Testart, Comments before the FCC in the matter of Secure Internet Routing, Tech. Rep., Federal Communications Commission (FCC), 2022.
[51] D. Clark, K. Claffy, Trust zones: A path to a more secure Internet infrastructure, J. Inf. Policy 11 (1) (2021) 26–62.
[52] J. Kristoff, R. Bush, C. Kanich, G.G. Michaelson, A. Phokeer, T.C. Schmidt, M. Wählisch, On measuring RPKI relying parties, in: Proceedings of the ACM Internet Measurement Conference, 2020.
[53] RPKI I-ROV Filtering World Map, 2022, URL https://stats.labs.apnic.net/rpki.
[54] T. Hlavacek, P. Jeitner, D. Mirdita, H. Shulman, M. Waidner, Behind the scenes of RPKI, in: Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS '22, Association for Computing Machinery, New York, NY, USA, 2022, pp. 1413–1426, event-place: Los Angeles, CA, USA.
[55] T. Hlavacek, P. Jeitner, D. Mirdita, H. Shulman, M. Waidner, Stalloris: RPKI downgrade attack, in: 31st USENIX Security Symposium, USENIX Security 22, USENIX Association, Boston, MA, 2022, pp. 4455–4471, URL https://www.usenix.org/conference/usenixsecurity22/presentation/hlavacek.
[56] S. Cho, R. Fontugne, K. Cho, A. Dainotti, P. Gill, BGP hijacking classification, in: Network Traffic Measurement and Analysis Conference, TMA, 2019.
[57] A. Gamero-Garrido, E. Carisimo, S. Hao, B. Huffaker, A.C. Snoeren, A. Dainotti, Quantifying Nations' exposure to traffic observation and selective tampering, in: O. Hohlfeld, G. Moura, C. Pelsser (Eds.), Passive and Active Measurement, in: Lecture Notes in Computer Science, Springer International Publishing, Cham, 2022, pp. 645–674.
[58] E. Carisimo, A. Gamero-Garrido, A.C. Snoeren, A. Dainotti, Identifying ASes of state-owned internet operators, in: Proceedings of the 21st ACM Internet Measurement Conference, IMC '21, Association for Computing Machinery, New York, NY, USA, 2021, pp. 687–702.
[59] P.V. Mockapetris, RFC1034: Domain Names - Concepts and Facilities, RFC Editor, USA, 1987.
[60] R. Elz, R. Bush, S. Bradner, M. Patton, RFC2182: Selection and Operation of Secondary DNS Servers, RFC Editor, USA, 1997.