D. Project Description: CNS Core:Medium: Understanding and Classifying DNS Resolvers

D.1 Introduction and Motivation

The Domain Name System (DNS) [92] is the Internet's de facto distributed name mapping system and it plays a major role in the Internet's critical infrastructure. Internet clients initiate nearly all of their communication requests by first performing DNS look-ups to translate names into IP addresses. When the DNS is unavailable, most reliant services experience a total outage after clients fail to look up connection information.

The DNS protocol is itself quite complex. There are 267 Internet standard documents ("RFCs") that specify, extend, provide operational guidance for or otherwise relate to the DNS. DNS infrastructure includes a vast number of participants, interacting in complex ways. The size of DNS' authoritative namespace is estimated to be 351.8 million zones [137], which are served by millions of *authoritative nameservers* [3]. Each nameserver can be authoritative for one or multiple zones, and can be replicated over many geographic locations to minimize response delay and increase robustness.

DNS clients include both applications used by humans and automated processes, which generate millions of queries per second to the DNS infrastructure. These queries arrive at local *resolvers*, which may reply to the query themselves, forward the query to another resolver or interact with authoritative servers to obtain the appropriate answer to return the client. There are millions of resolvers in the Internet today [76].

Among the many aspects that have enabled the DNS to scale and be successful for the past 30 years in the Internet is the hierarchical delegation system that separates authority and who answers for different branches of the DNS tree (i.e. delegate DNS zones). This is coupled the ability for the resolvers to experiment, implement, deploy, and manage their own infrastructures. Each party is free to make its own decisions, without any formal relationship or coordination with other participating parties aside from the mutually agreed upon DNS standard. While this design's encapsulation of roles has allowed for flexibility, broad deployments, and evolvability over the past 30 years, there is growing evidence that DNS ecosystem has become more of a jungle than a hierarchy, with intertwined dependencies, complex interactions with unforeseen consequences, and exotic, poorly understood behaviors that threaten how the DNS standard is interpreted and implemented.

Recently, DNS infrastructure has both been misused to launch attacks on others, and has itself been frequent attack target [115, 65, 122, 123]. Researchers have observed inconsistent and malformed DNS replies [72, 79, 148] as well as authoritative server misconfigurations [116]. DNS extensions, such as ECS, EDNS0 and DNSSEC, have been observed to lead to response delays [2], incorrect responses [95], and attacks [124]. The DNS community has already signaled its concern about the growing size and unsustainability of the DNS protocol *itself* [58]. Recent initiatives, like the DNS Flag Day [1], illustrate that the operational, standards, and software communities that support the DNS ecosystem have begun taking action to curtail the perennial accretion of corner-cases and bug fixes that have led the DNS protocol to become more cumbersome to maintain, implement, operate, and *understand*.

While prior research has investigated DNS servers [95, 72, 79, 148, 3] and some DNS use patterns [2, 45, 15, 80, 32] in great detail, little is known about the diversity of DNS resolver implementations and behaviors. Yet, resolvers are the linchpin of the DNS infrastructure; their behaviors can affect both DNS clients and authoritative servers. Misconfigured resolvers can be misused in attacks [124], or introduce delays in responses to clients [2, 16]. Resolvers that generate excessive queries to the DNS infrastructure [19, 18, 141] misuse precious resources of authoritative nameservers, possibly for no useful purpose. In part, the lack of insight and understanding into diversity of resolver behaviors have led communities to take leap-of-faith efforts like the DNS Flag Day, which did not go according to plans. The PIs believe that concise understandings of resolver behaviors, concise classification of resolver populations, and codified models of "proper" resolver behaviors are of paramount importance to DNS communities.

Some additional examples of phenomena that illustrate the importance of codifying resolver behaviors (proactively and through measurements) include:

- **Resolver misuse.** DNS resolvers have been heavily misused in recent attacks, ranging from their use to amplify traffic in *reflection attacks* [114, 41], to DNS poisoning attacks for profit [97]. Understanding

all the ways in which a resolver's configuration may lead to its misuse or malfunction is necessary, so that we can design and deploy mitigations to improve the Internet's security and stability.

- **Aggressive resolvers.** The DNS root, the Top Level Domains (TLDs) and other authoritative servers have frequently been targets of large denial-of-service (DDoS) attacks [122, 123]. Because traffic to these servers is sent by resolvers, some of which tend to send large number of queries with seemingly no caching, it is hard during DDoS attacks to differentiate attackers from overly-aggressive resolvers that are failing to properly cache data. This makes it hard to prioritize legitimate queries, forcing the server to either over-provision or drop random portions of queries to preserve bandwidth [138]. If we understood the range of *normal* resolver behavior, it would make it easier to identify attackers as anomalous sources of DNS traffic.

- **Interplay of configurations and use patterns.** The demand on DNS infrastructure is increasing as multitudes of new devices (e.g. IoT devices) are connected to the Internet. The complexity of the DNS protocol and its use is also increasing, thanks to wider implementation of DNSSEC [7, 9, 8, 52], use of DNS clouds (e.g., Route53, DYN), use of DNS to drive traffic to scrubbing centers during DDoS attacks, wider use of load balancing that drives the use of short Time-to-Live (TTLs) values on DNS records, and DNS record types requiring chaining and multiple lookups (e.g. SPF, DKIM, NAPTR, SRV [48, 84, 83, 73, 43]), etc. Configuration and operation of DNS resolvers, and their interplay with other elements of the DNS infrastructure, directly influence the load placed on the DNS root, TLDs and authoritative servers. We need to understand how different use patterns interact together and how they affect the load placed on the DNS infrastructure.

- **Early insight into attacks.** Many attacks on Internet users, such as phishing, malware propagation, and spam generate significant specific DNS query traffic as they unfold. If we could identify these attack query patterns, it would enable us to prevent or limit the harm these attacks inflict on Internet users.

- **DNS abuse.** Some DNS domain *names* are involved in *DNS abuse* behaviors like phishing attacks, counterfeit websites, roBot Network (botnet) Command-and-Control (C2) domains, and other types of malfeasance. A recent policy and legal development in Europe, called the General Data Protection Regulation (GDPR), has affected investigation of cyberattacks by mandating the removal of key information from the global `whois` registration system. This has led to a measurable degradation in cybersecurity abuse protections [118]. The general category of DNS abuse is a critical concern in the cybersecurity threat intelligence communities [23, 61, 117]. If we could link together DNS abuse behaviors with other DNS behaviors by the same or related actors, we could help bridge the gap created by GDPR and we could aid law enforcement officials and network operators in cyberattack prevention and mitigation.

The work of this proposal focuses on addressing the security, stability, and efficiency concerns of the DNS through evaluation-based and model-based understanding and classification of DNS resolver behaviors, and through an informed design of mitigations to alleviate negative effects. In the *investigative* thrust of our research, we will focus on understanding *observable* resolver behaviors. USC/ISI operates the B-root authoritative server (effort lead by PI Hardaker) – one of 13 DNS root servers. Our investigation will begin with analyzing the four-years' worth of queries received by B-root from a variety of sources. While a fraction of these sources may be direct clients themselves, we refer to all as "resolvers" for simplicity. We will first manually identify query patterns that signal anomalous or unexpected resolver behaviors, and then leverage deep learning to identify patterns at scale. We will quantify and prioritize these behaviors by their prevalence and the potential impact on the DNS ecosystem. We will seek to discover the root cause of these behaviors, "labeling" each group tags that identifies both their behavioral groupings and root cause, if found. As our research progresses, we will further refine our labels to include subcategories, and more specific tags.

When we have achieved a stable set of labels, we will extend our research to analyze data from other root servers, collected annually via DITL (Day in the Life of the Internet) effort [36]. This research will produce classification algorithms capable of tagging specific resolvers with one or more labels, depending on the blend of behaviors it has exhibited in recent past. Finally, we will cluster resolvers based on these

tags, resulting in a set of *resolver classes*. We will further quantify prevalence of different resolver classes and study their distribution in the Internet.

We will also link query patterns and actors that relate to abuse with public DNS abuse indicators, and we will seek to identify other, potentially malicious behaviors by the same or related actors. This will help us produce a DNS abuse feed, which we will share publicly.

In the *modeling* thrust of our research, we will seek to understand the range of possible resolver behaviors. We will start from relevant IETF RFC's and develop a detailed model of resolver behavior, taking into account various DNS extensions, resolver configurations and best practice guidelines from various communities in the DNS ecosystem. We will use this model to develop model-guided, abstract scenarios, which target security of DNS participants, or which lead DNS resolvers to behave in suboptimal ways (e.g., generating excessive DNS traffic, or introducing large delays in responses). We will then analyze the potential impact of these scenarios on the DNS infrastructure, and we will develop specific test cases to demonstrate their detrimental impacts. Our model will also help quantify how some observed DNS resolver behaviors agree with the models, thus helping us identifying the root causes of some behaviors seen in the investigative thrust.

In the *mitigation* thrust of our research, we will seek to identify a range of appropriate mitigation actions, which can be taken by stakeholders in the DNS ecosystem: root, TLDs, and other authoritative servers, organizations that host resolvers, clients, DNS software vendors and open source products, popular OS vendors, etc. For each action, we will seek to establish its impact on legitimate and malicious actors in the DNS ecosystem, and its deployment cost. We will then pursue the most promising actions, and will work to transition them to practice through collaboration with DNS infrastructure operators (e.g. authoritative servers and ISPs), software vendors, network operators, etc.

The investigative and modeling thrusts approach the same problem – understanding the range of DNS resolver behaviors and their implications – from two directions. The investigative thrust seeks to understand common and observed behaviors today and their root causes and implications, while the modeling thrust seeks to comprehensively explore the space of DNS resolver behaviors, and potential impact of specific behaviors on the DNS ecosystem. The mitigation thrust will then act on the results of the investigative and modeling thrusts to remedy the identified problems by producing recommended operating guidelines, documents for standardization, or software patches.

D.1.1 Intellectual Merit

Our research will shed light on the vast, complex and mostly unknown space of DNS resolver behaviors. It will lead to a new understanding of how DNS is used today, how it differs from the specifications and what are the root causes that drive unexpected use cases. Our research will improve understanding how configuration and operation of DNS resolvers impacts the security and the efficiency of the DNS infrastructure, and how it interacts with various DNS server configurations. Additionally, our work will produce results and techniques that fundamentally enhance the way that DNS abuse is being discovered and re-mediated. New insights from our work will be useful to harden existing implementations against misuse, to improve DNS efficiency, and to drive future DNS development.

As critical infrastructure for the Internet, understanding the DNS is of paramount concern. Moreover, the DNS is one of the oldest and most evolved protocols that is still running on the Internet today. Our work will provide methodologies that illustrate how to codify, measure, and evaluate protocols that have evolved over time and operate in new, unexpected ways. The PIs believe that the approach of understanding protocols and deployments from their behaviors, and then incorporating models into future designs and enhancements represents the next stage of protocol engineering and design. We intend to use the lessons from this proposal's work to influence and instruct these processes in venues and forums such as the IETF and ICANN.

Other contributions of the proposed work will include:

- quantification of unexpected, malformed or excessive DNS queries
- evaluation-driven models of how misuse by DNS resolvers can be used to impact constituent operators and users,
- root cause analysis that will identify reasons behind unexpected, sub-optimal or aggressive resolver behaviors

- "behavioral labels" for query patterns and resolvers, which will enable DNS server operators and researchers to develop new strategies for handling DNS queries, including query prioritization, selective query filtering, client redirection, reply content tailoring, etc.
- model-guided investigation of DNS resolver behaviors, enabling the discovery of new vulnerabilities in the DNS protocol and implementations, and better understanding how the crucial elements of DNS (e.g., DNSSEC, load-balancing, etc.) work together
- a set of standard DNS use-cases that describe both normal (expected) and abnormal resolver behaviors
- DNS abuse data feed
- our methodologies, which can be reused to study other long-lived Internet protocols

D.1.2 Team Qualifications

Our team brings a long history of working on DNS and on network attacks. PI Mirkovic has 18 years of work in network security, particularly on distributed denial-of-service attacks [91, 87, 86, 88, 89, 99, 85, 133, 50, 68, 33, 34]. PI Hardaker has years of experience in DNS service operation, security, research and standardization [52, 53, 54, 39, 38, 55, 24, 25, 26, 125, 126, 127, 112, 32, 95]. PIs Mirkovic and Hardaker are currently Co-PIs, together with John Heidemann from USC/ISI, on a NSF CICI funded project "DDoS Defense In Depth for DNS". That effort focuses specifically on protecting authoritative servers from direct DDoS attacks, by creating a range of defenses. PI Hardaker is the operational manager for the B-Root DNS server, which brings a core set of data to analyze to this project. PI Osterweil has conducted research on DNS, DNS Security Extensions (DNSSEC), DDoS, inter-domain routing security, and cybersecurity threat intelligence for over 15 years [110, 103, 105, 109, 111, 146, 104, 107, 101, 147, 106, 102, 4, 108, 20, 21]. He has studied and published measurement results on DNSSEC since it was standardized and its global deployment began in 2005. Prior to joining George Mason University, he was a Principal Scientist at VeriSign, Inc., which operates two of the DNS' 13 root server instances, operates `.com` and `.net`, operated a commercial DDoS mitigation service, a commercial-managed DNS product, a cybersecurity threat intelligence product called iDefense, and a public DNS recursive resolver service. While there, he conducted research that involved operators and product owners from all of these services, which provided him critical insights into their respective issues and landscapes. PI Osterweil and PI Hardaker have collaborated together for over a decade in the IETF, NANOG and ICANN to produce DNS and DNSSEC standards, conduct research about DNS deployment and advocate for DNS security technologies like DNSSEC [7, 9, 8, 52] and DANE [47, 39, 38].

D.2 Related Work

There has been a lot of prior work on studying and measuring the DNS, understanding security vulnerabilities and detecting attacks against DNS. We discuss the most related work here.

Measuring resolver behaviors. There have been multiple measurement studies of DNS resolvers, usually analyzing traffic at the DNS root servers. These studies reported on the prevalence of invalid traffic: [19, 18, 14, 141]. Brownlee et al. found 60–85% of bogus queries at F-root in 2001, and 14% of the queries violated the DNS specification. They also identified several interesting resolver misbehavior patterns. Castro et al. [19, 18] analyzed DITL data from 2006–2009 and noticed that only a small fraction of root clients (i.e., resolvers) are responsible for most of the query traffic. They also identified several behaviors of interest such as repeated queries, identical queries, invalid TLD etc. Wessels and Fomenkov [141] also analyzed traffic at the F-root server, with the goal of understanding the behavior of a few selected, very aggressive sources. Additionally, Gao et al. [46] studied DNS resolvers by directly collecting their traffic at the Security Information Exchange (SIE). This let them observe the entire traffic a resolver sent and received, instead of just a fraction of queries that propagated to the DNS root. The Gao et al. study found that 15% of queries to the root remain unanswered. Recently, Foremski et al. [45] investigated both resolver and server traffic at the Security Information Exchange (SIE). They describe traffic composition with regard to query types, response types, popular queries, and server and resolver distribution over networks and geolocations.

These existing studies are valuable foundations on which we will build our proposed work. They suffer from two deficiencies, which we hope to address. First, all but Gao et al. [46] and Foremski [45] were conducted 10–18 years ago, and DNS traffic has evolved significantly since then. Second, these studies are not comprehensive. They identified some interesting, but narrow, resolver behaviors and their root causes opportunistically, but they were not systematic or comprehensive. They also did not comprehensively

evaluate the impact on DNS robustness, security or efficiency from identified misbehaviors. Our proposed research will address these aspects by analyzing recent DNS root traffic, and applying systematic analysis to comprehensively understand resolver behaviors and their impact on DNS ecosystem. Our research will thus collect together and complement the knowledge from past works, extend them and create a corpus of complete resolver behaviors.

Detecting resolver misconfiguration. Schomp et al. [16] investigated DNS resolver behavior using active probing. They discovered many open resolvers, and noted that they are mostly short-lived. They also found that more than 80% of resolvers changed the TTL values in their replies, from the values returned to them by authoritative servers. Schomp et al. [131] investigated how vulnerable the resolvers in the wild were against several kinds of injection attacks and found DNS weaknesses are not rare. Our work will complement the work in [16, 131] by using passive observation of resolver traffic at the root. We will also ask a much broader range of questions about resolver behavior, beside just how they handle TTL values and if they are open resolvers.

Understanding caching. Several DNS studies looked into the effectiveness of caching and investigated caching behavior of various resolvers [71, 142, 11, 129, 130] and authoritative servers' TTL settings [45, 95]. These studies show that changing a caching policy (e.g., changing the TTL of records), or even completely removing shared resolvers, does not greatly impact DNS traffic, because the traffic itself is heavy-tailed, with many requests querying for unique names. Our work will examine the causes of both unique and repeated queries in greater detail, studying when, how and why resolvers generate these queries, and how settings of authoritative server records influence this dynamic (e.g., setting all records with short TTL values).

Detecting malicious domains and activities. Many researchers have investigated how to identify malicious domains by analyzing DNS data [5, 12, 6, 74, 121, 51, 145, 144, 152, 29, 128]. We hope to leverage their work, where possible, to identify malware as a root cause of some behaviors we expect to observe in our investigative tasks. Jones et al. [70] described techniques used to discover unauthorized "root" servers on the Internet. While unauthorized roots also affect DNS security, our research will focus on understanding DNS resolvers.

DNS modeling. Schomp et al. [132] built an empirical model of DNS clients, by analyzing client behaviors. We will develop models of DNS resolvers.

Reducing DNS Resolver Attack Footprints. RFC5358 [31] documents how recursive resolvers are used in reflection attacks, but limits the scope of its recommendations to merely ensuring that resolvers allow access only by their legitimate users. Unfortunately, modern attacks from botnets use legitimate hosts and their local resolvers to launch attacks. Our work will significantly extend recommendations to include safer operational configuration parameters for both recursive resolvers and authoritative servers.

Recent NSF-funded research on DNS. Mark Allman is the PI on recently funded effort by the NSF "NeTS: Small: De-Mystifying and Hardening the Domain Name System". That effort focuses on measurement of DNS client and server activities from myriad of smaller vantage points, and on improving the DNS ecosystem by removing reliance on recursive resolvers. Our work takes a complementary approach by seeking to understand DNS resolvers and how they interact with DNS clients and servers.

D.3 Proposed Work

Our proposed work can be divided into three inter-related evaluation-driven thrusts: investigative, modeling and mitigation. We illustrate their relationship in Figure D.1. The investigative thrust aims to understand and interpret behaviors of DNS participants, which are currently observed in the wild, focusing mostly on DNS resolvers. The modeling thrust aims to understand how DNS resolvers with different configurations and support for DNS extensions may interact with different clients and servers, and when these interactions lead to reduced security, robustness or efficiency of DNS protocol. The mitigation thrust seeks to understand the possible range of actions that can be taken by DNS stakeholders to remedy the problems identified by the investigative and the modeling thrusts. In the rest of this section we provide more details about each thrust.

Our investigative path will start with the analysis of DNS query data, which will yield query patterns matching and identifying DNS resolvers. We will generalize and abstract these patterns into behaviors of

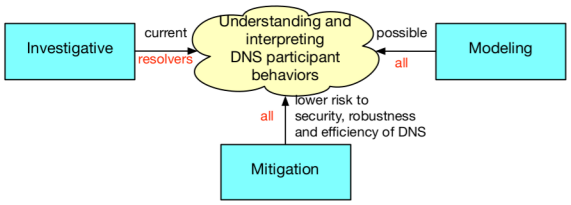

Figure D.1: The three thrusts of our proposed work.

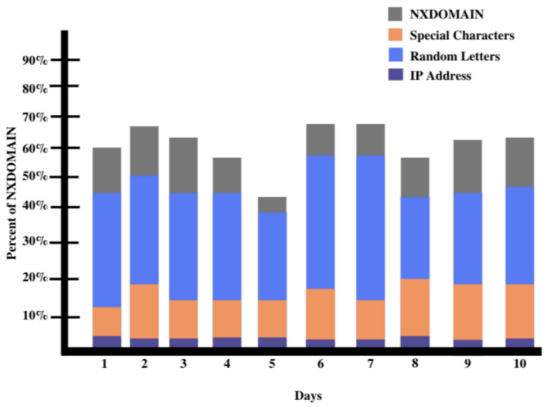

Figure D.2: DITL sample query analysis.

interest. These behaviors will be labeled, and a resolver will tagged with labels representing its exhibited blend of behaviors. We will further cluster resolvers according to their behaviors and investigate clusters that are particularly large or display behaviors we wish to address. We will then narrow down the set of interesting behaviors to only those that impact security, robustness or efficiency of DNS infrastructure or the Internet. For these behaviors, our attribution task will attempt to establish root causes. In parallel, we will seek to link results from our investigation of DNS queries with public DNS abuse feeds. This will both help us improve our classifiers and it will result in new DNS abuse indicators. Our modeling path will start will DNS RFCs, aided by operational guidelines. We will use these sources to produce a DNS resolver model. This model will yield abstract attack or problematic scenarios (akin to behaviors on the investigative path). We will use these abstract scenarios with a protocol fuzzer to generate specific test cases. Both our investigative and modeling paths will create inputs for the mitigation design activity. Mitigations will then be evaluated for effectiveness, and the promising ones will be used in our technology transition efforts.

D.3.1 Investigative Thrust

Task I1: Manual investigation of DNS traffic. In this thrust we will start by manually investigating resolver behaviors observable in four-years of DNS queries collected at B-root. Manual investigation is a necessary step, after which we will develop a range of automated tools to perform behavior classification and interpretation, and establish underlying root causes (*Task I2*). Every query received at B-root is recorded, enabling us to study the DNS packet contents as well as the IP, transport (UDP and TCP) and DNS headers too. Because traffic quantities received at B-root is large ($\tilde{4}$ B requests, $\tilde{3}$00 GB of xz-compressed data per day), we will carefully sub-select traffic to study to achieve a good diversity of patterns. We plan to initially select resolvers from three different categories, based on their query dynamics:

1. *Familiar resolvers (baseline).* A number of resolvers whose identity and purpose are known to us are in the organizations that we work with (USC, ISI, Colorado State University, UCLA, UCSD, GMU, UCDavis, etc.) We can isolate and analyze the traffic from these mostly legitimate, known clients. We will study their behaviors to identify cases where these resolvers act in unexpected ways, forming a baseline for root cause research (*Task I2*).
2. *Aggressive resolvers.* The DNS hierarchy minimizes query traffic in the upper layers via caching. TLD records, served by the root, are long-lived, with their time-to-live (TTL) set to 48 hours. While the TTL specifies the maximum length a resolver may cache a record, we expect resolvers to cache records for at least a few hours. A small number of very aggressive resolvers, though, generate thousands of queries per second for names that should have been cached. These aggressive resolvers are always present in B-root's traffic. In any one minute, 80% of queries received are sent by mere 1% of resolvers. In any one day, 80% of queries are sent by fewer than 0.1% of resolvers. These resolvers will be tagged with the "*aggressive*" label and be further tagged with new labels based on any discovered root cause in Task I2.
3. *Sporadic resolvers.* On another end of spectrum are sporadic resolvers, sending only a few, occasional queries to B-root. In one week of B-root's traffic, half of the sources ask only one query in seven

days. This suggests that these resolvers are not serving large organizations, but may be end hosts performing their own DNS resolution, or distant resolvers not selecting B-Root due to latency issues. We will seek to identify the root cause for these low-volume behaviors based on analysis of their traffic (§ D.3.6 describes successful results studying a similar low-volume use case). New labels will be added to these resolvers based on the investigation's results.

We will reduce our initial search space by sub-selecting queries based on two types of time intervals: (1) randomly selected hours of a random day – we will select multiple instances to equally cover all days in a week and all hours in a day, (2) specific hours and days when the B-root server was under heavy load, or when known malicious campaigns were taking place in the Internet. The random days and times will be used to study resolver behaviors that occur regularly, irrespective of large-scale attacks. We will also perform a longitudinal study of these usually-seen behaviors, to understand how they change over time.

Across the combination of initial resolver classifications and time instances, this task will seek to identify patterns of queries which are abnormal, such as aggressive and repetitive queries, malformed queries, queries that seem to follow a specific pattern of names, etc. We will then abstract and generalize from these patterns to arrive at *behaviors*, which describe each cluster of similar patterns. Each behavior will be labeled. We expect to both enlarge and sub-classify these labels as we progress with our research.

Task I2: Manual root cause attribution and analysis. For each identified behavior of interest, we will seek to attribute it to a root cause, and to analyze when and why such behavior occurs. This attribution is difficult, as we only have the data about what was asked (the DNS query) and who posed the question (a potential resolver at a source IP address) but not why the question was being asked. Thus we will have to hypothesize and develop approaches to confirm or refute our hypotheses. We now outline how we may be able to infer the purpose behind a set of DNS queries, applying three forensic approaches. We hope to identify more approaches during the course of our proposed research.

A vanilla-software analysis approach will serve as a validation stage, whereby the PIs will use evaluation-driven analyses of known software to create a baseline for the project's measurement phases. In this approach, the PIs will pose queries to vanilla installations of popular resolver software, following the same order and timing as the recorded pattern from the trace. We will then observe if we can reproduce the resolver behavior. For example, if the resolver under study posed two related queries to the root (e.g., an address query and a DNSSEC related query), we will replay these queries to our test resolvers, and analyze the results. We can then study if the queries trigger the same behavior seen at B-Root as each resolver version works toward answering the question posed by our client. If we can recreate a pattern of interest, we can then label it accordingly and seek its underlying causes in the resolver software and its default configuration.

In the vanilla-software analysis we may be able to utilize the Resolver testbed [57], an open-source, virtual testbed sponsored by ICANN, containing vanilla versions of popular resolver software. Both PI Hardaker and PI Mirkovic are members of the Root Server System Advisory Committee's (RSSAC's) Resolver Work Study Group. They will use this opportunity to identify synergistic efforts between the proposed work and the group's charter, such as the use of the Resolver testbed.

A pattern-based approach will seek to identify if queries received at B-Root have a specific syntactic pattern, e.g., they all ask about the same or related names. If this matching is successful, we may label the behavior as "*non-caching*", and then sub-divide this label into lower-level labels depending on the observed pattern. For example, we may add the label "*enumeration*", if queries seem to try to enumerate the same domain, or "*malicious*" if names appear related to a known attack instance, or "*random*" if queries seem to exhibit random pattern.

An originator approach will investigate the resolver or resolvers that exhibit a behavior of interest by looking at the organization owning its IP address network block, performing reverse DNS lookups, and blacklist searches to understand the underlying purpose of the machine. If it is known to have participated in prior attacks, we may infer and label the observed behavior of interest as "*malicious*". If it is an open resolver, We may also send it queries to understand its purpose. These probes will be designed carefully, following the established guidelines for ethical active measurement approaches [113, 40].

Task I3: Machine learning and automated attribution. While we will identify some patterns of interest in *Task I1*, the sheer volume of data from resolvers prevents manual approaches from scaling in the long term. In this task we will employ machine learning techniques to identify other patterns from the entire B-Root dataset, via clustering and pattern detection. We will initially apply Deep Belief Networks [77], which

perform unsupervised learning, and have performed well for pattern detection in other domains [49, 10]. Once deep learning techniques identify patterns of interest, we will generalize and abstract the results into behaviors, applying labels based on the discovered root causes. A significant goal of this task will be automation of behavior attribution (root cause identification), through automated testing of vanilla versions of popular software packages, and by automating originator checks for IPs of interest.

Task I4: Applying pattern identification and root cause analysis to DITL datasets. DNS Operation Analysis and Research Center (OARC) collects DNS traces from busy and interesting DNS nameservers through various means, such as the annual Day In The Life of the Internet (DITL) collection effort. Members of OARC can analyze this data on OARC machines (USC/ISI is a member). Once we have pattern signatures, we will apply them to the available DITL datasets at OARC to seek patterns of interest in data contributed by other root nameservers. This will enable us to understand how prevalent specific resolver behaviors are in the Internet, and how well they can be observed from different nameservers. We will further seek to apply our deep-learning pattern identification to the DITL dataset, to identify new patterns. This may require modification of our neural network from step I3, to make it suitable for smaller amount of computational resources at OARC machines.

Task I5: Identifying interesting scenarios. Only some of our resolver patterns of interest and their root causes will have the potential to be misused for attacks or potential to create inefficient use of the DNS infrastructure. In this task we will look to identify these patterns and construct misuse scenarios. We will analyze these scenarios for feasibility and ease of use, as well as for their impact on the DNS participants. We will later test each scenario in a controlled, isolated environment, during our evaluation tasks. These analyses will also be correlated with other threat-intelligence data sources to either corroborate findings and/or increase confidence in detected misuse.

Task I6: Mining information from observed DNS abuse. The PIs will acquire open datasets and feeds of DNS abuse, such as SORBS and SURBL, containing DNS domain names that are indicated to be associated with various forms of abuse: phishing sites, spam domains, roBot Network (botnet) Command-and-Control (C2) domains, and others. Starting from this known abuse we will develop specific classifiers to empirically classify and evaluate query behaviors for specific abuse types. For example, we will develop classifiers that separate spam-specific query patterns from those involved in phishing. This task will also quantify the behavioral differences between cases when queries are issued directly form clients, and cases when they are proxied through upstream resolvers. We will further use DNS abuse data to identify resolvers that frequently engage in queries containing DNS abuse indicators.

Task I7: Detecting DNS abuse from classifications. In this task we will apply the classifiers developed in task I6 on our datasets to identify other queries and resolvers that engage in abuse-specific behaviors. We will also seek to measure the accuracy of our classifiers by trying to predict future abuse indicators. These measurements will collect information about DNS domains that do not yet appear in DNS abuse feeds we monitor, but that are part of DNS abuse as labeled by our classifiers. We will then measure if such domains appear later in the DNS abuse feeds.

Task I8: Pivot from detected DNS abuse domains to discover other related abuse domain names. Detecting otherwise unreported DNS abuse domains is a pressing need in the cybersecurity field. Recent developments from the General Data Privacy Regulation (GDPR) have caused DNS registry service providers to begin redacting critically important information that has long existed in WHOIS service, making detection and attribution more difficult. The work of this task will be to develop and evaluate the ability of using the resolver classifications developed in this proposal, and their linkage with reports of DNS abuse, and external data sources (like Regional Internet Registries' IP allocation information) to detect additional DNS domains that may be involved in DNS abuse. For example, we may start with identified misbehaving resolvers and pivot to examine DNS queries sent by other IP addresses in the same behavioral classification (from Tasks I1-I7), or allocated by the RIR to the same organization. This may lead to abuse domains operated by the same actor or in the same campaign. This effort will also produce new intelligence about domains and resolvers involved in DNS abuse, which will be shared in task MI3.

D.3.2 Modeling Thrust

Task MO1: Building the basic DNS resolver model. In this task, we will build the basic DNS resolver model, starting from those RFCs that are standard and that pertain to DNS resolvers (RFC 1034, 1035, 1123, 3596, 4034, 4035, 4509, 5011 and 6891 [93, 94, 44, 135, 9, 8, 134, 52, 30]). Our resulting model will be a

finite state machine, with configuration parameters that drive states and transitions, as well as transition triggers. The model will further accept client queries and server responses, processing them to adjust the model's state and emit certain events (e.g., sending of a query or a response). We will parameterize this model with various configuration parameters, listed in the RFCs. We will then apply model-based testing [136], to derive functional, abstract scenarios (akin to behaviors in the investigative thrust), which can then be distilled into practical, concrete test cases. These test cases would consist of sequences of DNS queries and/or responses that produce unwanted behaviors in DNS resolvers. This approach has been successfully applied by Jero et al. to identify new attacks on TCP congestion control [69]. The outcome of this task will be new attack scenarios that misuse the DNS to either attack DNS participants, or other targets on the Internet (e.g. using amplification), or to lower the robustness of the resolver to attacks, or to lower the efficiency of the DNS protocol (e.g., by creating unnecessary queries)

Task MO2: Extending the DNS resolver model. In this task, we will survey other RFCs, including the informational and Best Current Practice (BCP) RFCs relating to DNS participants. We will use information gleaned from these surveys to extend our resolver model with additional parameters, and to investigate how different observed practices in the DNS operational world affect correctness and robustness of DNS infrastructure. We will then repeat our model-based testing with the extended DNS resolver model. Since there will likely be many parameters for the resolver model, we will seek to group them by functionality and to understand their dependencies, so that we can reduce the number of tests that we need to evaluate.

Task MO3: Enriching DNS resolver model. In this task, we will survey best practice guidelines for operators of DNS nameservers and DNS resolvers. These guidelines are offered by various DNS software vendors, DNS hosting networks, and operator communities. These practices may impact how we parameterize our resolver model, and in some cases may add new parameters to the model. We will also examine popular DNS resolver and server software to extract additional parameters for use in our model.

Task MO4: Cross-pollination between modeling and investigative thrusts. It is very likely that in the investigative thrusts there will exist many resolver behaviors for which we could not establish a root cause, or where we hypothesized about several causes but could not confirm any with high certainty. We may be able to resolve these undecided behaviors with help of our models. For example, if we observe that a resolver asks the root questions too frequently based on what should have been cached, we may learn from our model that this can happen when the EDNS UDP maximum size parameter is smaller than the nameserver's response. We can then infer the value of the EDNS UDP maximum size parameter for this particular resolver based on its observed behavior and we can correctly tag this behavior with the label "edns-udp-max-size too small".

Task MO5: Develop resolver test suite. Our modeling thrust will produce test cases that exercise a certain vulnerability or produce a certain unwanted behavior. We will run these test cases against popular resolver software, potentially leveraging the Resolver testbed [57]. We will evaluate the behaviors of existing resolver software suites to understand how they handle our test cases, quantify these effects, and then design solutions for mitigation.

D.3.3 Mitigation Thrust

Task MI1: Identify and characterize different mitigation approaches. In this task we will create mitigations to counter the effects of problematic behaviors identified in the investigative (*I1-8*) and modeling (*MO1-4*) thrusts. Each of the results from the investigative and modeling thrusts will be systematically analyzed for potential solutions. We currently foresee the following categories of mitigation approaches:

1. Develop and publish defensive filtering and other networking mitigations to protect authoritative servers against misbehaving or maliciously tagged resolvers [138, 42].
2. Develop and publish software patches to address vulnerabilities we discover and to optimize DNS protocol implementations in popular DNS resolver software
3. Develop and publicize guidelines for default resolver configuration changes
4. Develop new BCP RFCs documenting resolver configurations more robust against attack and misuse
5. Develop new BCP RFCs suggesting authoritative nameserver configurations that discourage or minimize the impacts of attacks and misuse. These could be: (1) selectively refusing to answer queries matching patterns, (2) rate-limiting aggressive resolvers [138], (3) refusing to answer queries by "*sporadic*" or "*aggressive*" resolvers when overloaded, (4) load-balancing to dedicate resources to classes of resolvers, isolating effects of their behaviors within the group.

We will investigate each mitigation approach to evaluate its cost, ease of deployment and effectiveness. We will do this both theoretically, by analyzing expected gains and drawbacks, and practically, by designing evaluation scenarios and performing evaluation in a controlled, isolated environment, such as in DeterLab [120] testbed or in Resolver testbed [57].

Task MI2: Pursue deployment of promising mitigation approaches. We will work on putting promising mitigation approaches into practice. In this process we will leverage our proven approaches for transition of new technology to B-Root. USC/ISI's B-Root service strives to be a leading edge, but operationally robust, DNS root server and research platform. To that end we have developed an extensive testing platform designed to specifically test solutions like the mitigations to be developed under this effort. Under upcoming NSF CNS-1925737 effort, PI Hardaker and John Heidemann are developing a B-Root test bed, which allows easy evaluation and testing of technologies like our mitigations before being deployed to a live environment. B-Root's approach to deploying new technologies involves testing the prospective code under both simulated loads and real-world traffic using replay technologies [151]. Once a new technology has proven both successful and safe in testbeds, it is deployed to a fraction of production B-Root service nodes for live in-line testing. Assuming successful results, it is eventually deployed to the rest of the production environment. After full deployment within B-Root, the success is reported to the rest of the Root Server Operators for potential deployment at the other root servers. PI Hardaker and his B-root operational staff routinely communicate with many of the root and TLD operators and can advertise our successful results to them as well at ICANN and operational meetings. Finally, PI Hardaker also regularly speaks to engineers from popular recursive resolver software packages (e.g. ISC's Bind and NLNet Labs' Unbound) and can communicate suggested software changes to them along with our evidence for change derived from our experiments. All of these existing communication channels serve as the basis for our expected successful technology transfer into operational practice.

Task MI3: Produce DNS abuse indicators. The work of this proposal includes several approaches to detect domain names that are involved in forms of DNS abuse and also to use those detected domains to discover other DNS domains that may be involved in abuse. In this task, the PIs will produce actionable telemetry about these domains to disseminate to cybersecurity researchers, operators, and more broadly to communities. This data will be distributed via threat intelligence feeds, Response Policy Zones (RPZ) [28], and the project's Web portal.

D.3.4 Evaluation Plans

During the evaluation of this work, the PIs will measure the effectiveness of each mitigation approach, comparing the correctness and traffic cost measurements with and without mitigations in place. We will contrast this effectiveness measurement with the memory and processing cost of deploying the mitigation.

Task E1: Correctness of Model. The initial phase of evaluation will be to verify the correctness of the models derived. Common and expected resolver use-cases will be developed and used to verify that the models created by this work properly codify resolver behaviors under specific (testable) conditions and scenarios. We will run such cases against our model, and also against popular resolver software, leveraging the Resolver testbed [57]. We will then investigate any discrepancies and, if needed, update our model.

Task E2: Performance and Impact Evaluation. Our work requires reproducing certain DNS query request and response patterns observed in the investigative thrust, or in the scenarios designed by us in the modeling thrust. When we detect interesting behaviors in the investigative thrust, we will seek to evaluate how they can be misused for attacks on the DNS hierarchy or on other Internet hosts, or to create unnecessary load on the DNS infrastructure. Similarly, when we develop new test cases under the modeling thrust, we will seek to evaluate their impact. Finally, after devising promising mitigation approaches, we will evaluate how well they address a given problem we have identified.

We must reproduce undesirable resolver behaviors (attacks, excessive queries, etc) in a controlled, isolated environment, to ensure predictability and to protect others from unwanted consequences of our experimentation. We will perform the majority of our evaluation on the DeterLab testbed [90, 120]. DeterLab is DHS- and NSF-funded, public testbed for cybersecurity experimentation, hosted at USC/ISI and UC Berkeley. It consists of more than 700 PCs, which can be remotely organized into custom topologies, and allocated by researchers for exclusive use. All experimental traffic is contained within the testbed. We plan to reproduce some of the DNS hierarchy within DeterLab testbed. We do not have to reproduce DNS clients, but will instead programmatically launch queries with the specific timing and content required

by each of our tests, to the resolver being tested. We plan to use the `dns-replay-client` [149] and `dns-replay-controller` [150] tools, publicly released by the ISI's ANT lab [56]. We will deploy several resolvers – each running a specific version of a popular resolver software, with specific configuration. We will further reproduce root and TLD servers, as well as authoritative servers of interest. To achieve scalable testing we may need to multiplex multiple servers on the single hardware node.

In our evaluation, we will gather several metrics to measure the impact of attacks and effectiveness of the developed mitigations. First, we will measure *correctness* – whether or not correct information propagated to the resolver and was returned to the client. Second, we will measure the *traffic cost* – the amount of traffic received and sent by each participant in the DNS hierarchy. Third, we will measure total *latency* experienced by the client between when they sent their query and when the response was received. The traffic and client latency will be measured with desired (expected) resolver behavior and with the specific, unwanted resolver behavior. In cases when a mitigation approach is under test, we will measure the case where both the unwanted behavior and the mitigation are active and will measure the *memory and processing costs*. We will evaluate the impact of unwanted behaviors by evaluating if the correctness metric has been reduced, and if either traffic or latency cost (or both) has increased under the behavior compared to the desired-behavior case. We will evaluate effectiveness of each mitigation approach, comparing the correctness and traffic cost measurements with and without mitigations in place. We will contrast this effectiveness measurement with the memory and processing cost of deploying the mitigation.

Task E3: Scalability Evaluation. When we select approaches for technology transition, we will need to reproduce high loads to evaluate how these mitigation approaches scale when many queries must be processed per second, or when many sources are simultaneously active. We plan to perform this evaluation on the DeterLab testbed and the B-Root testbed. USC/ISI uses a selection of open source tools for stress and regular testing of B-root's design and configuration [149, 150, 35, 37, 22]. These are the same tools that the B-Root operational staff uses to test new hardware and software before deployment to their production environments. Our tests will measure how well resolvers and authoritative servers handle query traffic with and without mitigation approaches, both under normal and attack conditions.

Task E4: Corroborate DNS abuse with external telemetry. Using both the resolver classifications that we derive from the investigative thrust and open data sets with DNS abuse reports (such as SORBS, SURBL, and others), this task will evaluate the accuracy of our investigations to proactively detect and indicate DNS abuse domains. In this task, careful evaluation will be made to measure how well we can predict future abuse reports, and if we have bias toward detecting certain types of DNS abuse better than others.

D.3.5 Technology Transition

In this Section we detail our technology transition efforts.

Task T1: Transition to Software Vendors. When mitigation approaches result in patches to DNS software, or changes to default configurations, we will work with software vendors to incorporate these patches into future releases. PI Hardaker works closely with popular DNS resolver and authoritative server software vendors at IETF, ICANN and other conferences. He will engage these vendors in bi-directional communication, both to receive their feedback on our work and to transition our solutions to practice.

Task T2: Transition to Authoritative Servers. When mitigation approaches result in changes to authoritative server software, we will evaluate them for possible deployment at B-root as described in Task MI2 and will share these results with the operator communities. PI Hardaker has operated B-root server since 2014 and is a respected member of the tight-knit Root Server Operators team. He will perform outreach to root and authoritative server operators. We describe these activities in more detail in Section D.4.2.

Task T3: Transition via academic and network operations community. All team members will participate in outreach to academic and network operations community (Section D.4.2).

D.3.6 Preliminary Results

In this Section we describe some preliminary results relating to various tasks.

D.3.6.1 Machine learning for pattern identification

PI Mirkovic worked with an REU student (through the ISI's REU site) Alexandra Fernandez to identify interesting and prevalent patterns in several samples of DITL data. They counted character frequency over each resolver's query sets and applied a threshold approach to transform these frequency vectors into binary vectors. They then applied density-based clustering over binary vectors to identify resolvers with

similar patterns. Largest clusters were investigated manually to derive a succinct understanding of the observed pattern. Their findings are illustrated in the Figure D.2. The majority of queries for non-existing TLDs fall into the following three patterns: random letters, IP address instead of domain name, and queries with special characters (like commas, spaces, ampersand, etc.)

D.3.6.2 Root cause analysis

Chrome NXDOMAIN queries. The root servers receive a large quantity of requests for TLDs that do not exist. There is a wide range of these that need investigating. By using a *pattern-based approach* we identified one pattern stemming from the Chrome web browser's deliberate use of random DNS strings. In an effort for the application to test if the local ISP is performing Internet redirection, common in hotels and other captive portal deployments, Google's Chrome (and Chromium) browser sends three address queries for randomly generated names with names of length 7 to 15 characters (e.g. `qvshurwrvg`). If the answers returned by the resolver all include real addresses in the response, Chrome knows that it is operating within a captive portal. Because these random strings do not exist within a non-captive resolver's cache, all of these queries leak up to the root servers resulting in a significant portion of root's incoming traffic.

CloudFlare + bind = cropped responses. We observed that a familiar resolver at USC repeatedly asks B-root for an address for `www.cloudflare.com`, when it should be remembering the answer for `.com`'s nameservers within its cache. We applied the *vanilla approach*, installing bind software on a test machine operating as a caching resolver. We then repeatedly asked this test resolver for an answer for `www.cloudflare.com`. We observed that the resolver returns a sufficient but incomplete response to the client. Specifically, in the process of resolving the query, the .com TLD server returns a 577-byte response that includes the names of five authoritative name servers for CloudFlare, and an IPv4 and IPv6 address for each of them. However, vanilla bind's EDNS UDP maximum size parameter is set to 512 bytes, resulting in some of the TLD's response getting cropped from the "Additional" section. As a result, the bind's cache lacks address information for two of CloudFlare's name servers. When one of CloudFlare's NS records in the bind cache expires, any future request to resolve a name within `cloudflare.com` stimulates the resolver to also ask for the missing address records for the two other name servers. This situation can be better optimized by changing either CloudFlare's zone records or by changing the default bind configuration to use a larger default EDNS UDP maximum size. Resolver issues like these need to be understood, documented and communicated to software vendors, OS vendors and network operators of both authoritative servers and resolvers. While we identified this issue through manual investigation, we also expect that our modeling approach should be able to uncover it.

Repeated DNSKEY queries. In late 2019 and early 2019 the DNSKEY that signs the root zone was swapped out in an operational rotation called a "key roll". During this event, a large number of odd resolver behaviors were seen in root server traffic. Extensive analysis was performed to figure out some of the root causes of traffic increases, but many cause are left unanswered. By using our *A vanilla-software analysis approach*, PI Hardaker discovered one root cause of repeated DNSKEY queries in ISC's bind recursive resolver. When configured with certain options and the old DNSKEY as its trust anchor, bind would occasionally trigger significant numbers of DNSKEY queries for the root zone's DNKSEY set. Discovering the issue took repeated experiments using bind version *9.11.5-P4*, as after 20 experiments the query storm only showed up 3 times (15% of the time). Analysis of this key rollover, including a few of the root cause behaviors found (and others that remain unfound), is the subject a paper accepted for publication in IMC-2019.

Datashack misuse. We observed that several IPs sent tens of thousands of queries per minute to B-root in May 2018. The queries exhibited multiple patterns. By employing the *originaltor approach* and examining the ownership of the IP address block, we established that these aggressive resolvers were hosted by DataShack, LLC – a hosting company in Kansas City, MO. A reverse DNS lookup told us they were set up as resolvers for `as32097.net`, which seems to be the domain owned by WholeSale Internet, Inc. Both DataShack and WholeSale Internet have in the past been known to host botnets and malware [75]. Currently, 455 IPs from their address ranges appear on public blacklists.

D.3.6.3 Root cause analysis and mitigation success

In this Section we describe a recent success in analyzing resolver traffic, identifying the root cause of unexpected resolver behavior, and working with the DNS community to mitigate this cause. This showcases that: (1) our manual approaches are effective at identifying interesting behaviors and their causes – in the proposed effort we will greatly enlarge and automate these processes, (2) our current connections with the

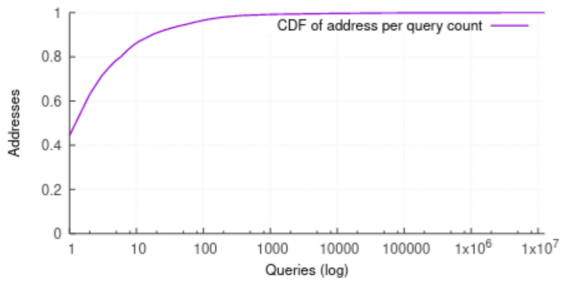

Figure D.3: *Addresses vs queries CDF*

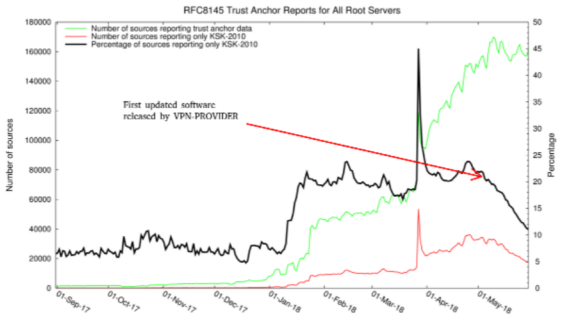

Figure D.4: *Graph showing mitigation results*

DNS community enable us to quickly transition to practice promising mitigation approaches.

In 2016, ICANN began planning [60, 62] the first key change of the DNS' Root trust anchor (TA). The adopted plan [59] established a time-line wherein the new public key would be published on *2017-07-11* and put into production use on *2017-10-11*, allowing roughly 3 months for DNSSEC validating resolvers to obtain the new public key and add it to their TA sets. However, on *2017-09-27* ICANN stopped the rollover process [63] due to adoption rate measurements indicating that approximately 5% of incoming RFC8145 [139] signals sent by validating DNS resolvers for only the older *KSK-2010* DNSSEC root key [140]. The rollover was subsequently delayed until *2018-10-11* [66] while the technical community continued measurement and analysis of the situation [64]. At the 61st ICANN meeting (2018 March), new analysis was released [67] showing an increase, rather than a decrease, of RFC8145 signals received for just *KSK-2010*, coupled with a steady increase in the number of unique sources sending RFC8145 signals.

PI Hardaker hypothesized that we may find the root cause of this behavior by examining all of the traffic from the resolvers that exhibited strange behavior, to the root servers. To reduce the time required for searching through B-Root's large dataset, we narrowed our investigation to *just those sources that sent a single RFC8145 query to B-Root during the March 2018*, reducing the queries to analyze from 1.2 M to 16 K.

With this select set of queries from resolvers, we examined the other queries they sent to B-Root. Many of these addresses sent very few queries, as seen in Figure D.3. This graph (with a logarithmic X axis), shows that most resolvers in this set are sending very few queries during the four-week period (63.06% of the sources sent two or fewer queries to B-Root, and one was the RFC8145 signal), In our proposed research such behavior would be labeled "*sporadic*").

We next focused our analysis only on those resolvers that had exhibited sporadic behavior.We extracted and correlated the top few query names, beyond the RFC8145 queries ("*_ta-4a5c*") and the queries for root-zone data itself ("*. (period)*). The most frequently requested names were a Virtual Private Netowrk (VPN) provider's primary domain (anonymized here as *VPN-PROVIDER*.com) and its alternate domain (anonymized here as *VPN-PROVIDER-ALTERNATE*.com). This commonality in top query names strongly indicated that a bug in the specific VPN software may be the root cause of *KSK-2010*-only signaling.

We next initiated a mitigation activity and disclosed this discovery to the VPN provider's technical staff. They responded rapidly and confirmed that all of their software products were affected by this issue, through hard-coding of root TA. They quickly produced and released a software update.

Our methodology of identifying and classifying resolver behaviors has already proven successful in solving real-world issues. Figure D.4 depicts the downward trend in *KSK-2010*-only signaling at the rightmost part of the black line. This is the direct result of our preliminary analysis and outreach.

D.4 Broader Impacts of the Proposed Work and Collaborations

This section provides details of expected broader impact of our work and intended collaborations.

D.4.1 Broader Impact

The broadest impact of this work will be to improve efficiency, security and robustness of the Internet's Domain Naming System by directly identifying and addressing vulnerabilities and problems in the way DNS infrastructure is used today. Since nearly all Internet activity involves resolving DNS names, the impact of this contribution is significant—spanning commercial, government, scientific, and personal uses

of the Internet. Our labels will help authoritative servers understand the root causes of the traffic they receive, and our mitigation thrust will help administrators understand a range of possible actions available, and how they impact users. Further, our root cause analysis will uncover misuse and vulnerabilities in resolver software and/or configurations. We will work with DNS community to patch discovered issues. Our research will lead to a deeper understanding of the load placed on DNS servers at different levels of hierarchy, and may lead to significant load reduction as DNS traffic gets better distributed. Our DNS abuse indicators will help operators and law enforcement fight cybercrime.

While we focus on how DNS is currently used and specified, our approaches to analyzing DNS traffic and using DNS specification for model-based attack discovery will be relevant to future variants of DNS.

D.4.2 Outreach Plans

Once we identify promising mitigation approaches we will perform outreach to various communities to transition them to practice. PI Hardaker has been integrally involved with the DNS standardization (IETF) and operations communities (e.g. NANOG, LACNOG, DNS-OARC, etc) for 15 years. He has strong relationships with other Root operators, from large commercial entities (Verisign, Cogent), to small businesses (Netnod), and from academic institutions (University of Maryland and the WIDE Project in Japan), to non-profit organizations (ISC, RIPE), and government bodies (DoD, NASA). Hardaker is also a respected member of Internet Governance and policy bodies associated with Internet Corporation for Assigned Names and Numbers (ICANN). These regularly bring together technical experts, TLD and root-server operators, software vendors, registries and registrars, and governments to discuss deployment and operational policies of the DNS at all levels. He is regularly invited to security specific meetings within these contexts (ccTLD security meetings, SSAC, RSSAC, root server operator meetings, etc) where he can freely present our research findings. Hardaker also serves as a member of the Internet Architecture Board (IAB), which provides architectural oversight of IETF activities. We will leverage Hardaker's network of colleagues and these meetings to advertise our work to this community.

We will present our work at academic research conferences such as ACM SIGCOMM, Internet Measurements Conference, IEEE S&P, Usenix Security, etc. We are further closely involved in efforts to secure the DNS by the wider research community. ISI hosted the DNS and Internet Naming Research Directions Workshop in November 2016, PI Hardaker held three DNS "Birds of a Feather" (BoF) brain storming sessions from 2016-2017, and our team has also attended the DNS privacy workshop at NDSS in February 2017 and 2018. These events have helped us establish collaborations with other researchers that work in this field. We will leverage these connections to solicit feedback from the academic community and continue looking for opportunities for further technology transition.

PI Mirkovic is on two technical advisory committees for CENIC. We technologies will present our work at these community meetings to solicit interest for further technology transition.

Finally, we will leverage the PI Osterweil's active participation in ICANN, as a Vice-Chair of the second Security, Stability and Resiliency (SSR2) Review Team, to share our work with ICANN community. We will also give presentations about our work at Network Operational Groups (*NOGs) meetings, which are held in five global regions (Latin America, Europe, Asia/Pacific, North America, and Africa). PIs Hardaker and Osterweil regularly attend these meetings and present about their DNS research.

D.4.3 Educational Plans

We will work to transfer results of our research into education PI Mirkovic teaches and advises PhD students, and PI Hardaker has been very active in advising MS students in research and leading industry tutorials and lectures. They expect to integrate what they learned, as well as any available public datasets, into their curriculum at USC and their research work with students. We will also develop a homework assignment on DNS security, and share it publicly via the DeterLab education portal [119]. PI Mirkovic has developed this portal, and she has led two NSF-funded efforts to develop student exercises, which are shared via the portal. She is dedicated to improving security education with the help of testbeds, to facilitate active learning.

PI Osterweil teaches classes relevant to Internet networking and related security topics. In addition to emphasizing networking fundamentals, these classes also include current Internet evolutions and security work. This effort will have strong relevance to these topics, and he will incorporate its results into lectures and include students from his classes in relevant portions of this work. In particular topics such as DNS abuse and cybersecurity will be enriched with concepts and results from this work.

We will involve undergraduate students in this research through the ISI's REU Site (led by PI Mirkovic), and several other undergraduate research programs at USC, such as Viterbi-India, Viterbi-China and SURE.

D.5 Results from Prior NSF Support

PIs Mirkovic and Hardaker have had numerous projects funded by NSF; we report on recent ones here.

Collaborative research: Hands-on exercises on DETER testbed for security education, (DUE-0920719, $ 112,842, 9/2009-8/2012), PI Mirkovic's collaborative project with Peter Reiher (UCLA), Mooi Choo-Chang (Lehigh University), Brent Huyn Kang (GMU, now at KAIST) and Daniel Massey (Colorado State University, now at UC Boulder). *Intellectual Merit:* This award resulted in 15 hands-on exercises for cybersecurity. *Broader Impact:* To date, these exercises were used in at least 200 courses, by 14 K students.

Revitalizing Cyber Security Education and Research through Competitions, (CNS-1319197, $ 300,000, 9/2014-12/2016) was project by PI Mirkovic and Ron Pike (Cal Poly Pomona). *Intellectual Merit:* This award resulted in six competition exercises for cybersecurity classes. *Broader Impact:* These exercises were used in at least 11 courses so far, and in numerous events to recruit students into cybersecurity.

CICI: RSARC: DDoS Defense In Depth for DNS. (OAC-1739034, $ 997,226, 07/2017 - 07/2020), is collaborative with John Heidemann (USC/ISI) and PI Hardaker (USC/ISI). *Intellectual Merit:* This project developed new techniques to protect authoritative DNS servers against DDoS attacks. *Broader Impact:* Results of this work are currently being tested and implemented at B-root hosted at USC/ISI.

Collaborative Research: Modeling Student Activity and Learning on Cybersecurity Testbeds. (DGE-1723717, $ 250,000, 08/2017 - 08/2020) is PI Mirkovic's project, collaborative with Richard Weiss (Evergreen State College) and Jens Mache (Lewis and Clarke College). *Intellectual Merit:* This project has developed measurement and analysis approaches for how students learn with testbeds, and is now developing intervention techniques to help those students that show slow progress. *Broader Impact:* Learning measurement approaches are currently implemented on the DeterLab testbed and early results are reported in [78].

REU Site: Human Communication in a Connected World. (DGE-1723717, $ 360,000, 05/2017 - 05/2020), is PI Mirkovic's project collaborative with Christophe Hauser (USC/ISI). *Intellectual Merit:* To date, 17 students at this site were involved in research on communication, cybersecurity, information flow and retrieval. *Broader Impact:* Students learned how to perform research and got prepared for graduate school.

CCRI: DNS, Identity, and Internet Naming for Experimentation and Research. (CNS-1925737, $ 1,458,440, 10/2019 - 09/2022), is a DNS research testbed being developed by John Heidemann (USC/ISI) and PI Hardaker. *Intellectual merit:* This project will catalyze a broad community of researchers in Internet naming and identification. *Broader Impact:* This project will result in improvements to Internet naming, trust, and transitional infrastructure that come from the research community.

DETER Research Education and Operations Mission Sustainment (OAC-1842703, $2,000,000, 08/2018 - 08/2020), is PI Mirkovic's project, collaborative with Terry Benzel (USC/ISI). *Intellectual Merit:* This project develops new research, education and operational features for the DeterLab testbed to meet new user needs. To date it resulted in novel ways to include human factor in cyberexperiments, new models of complex networks for cyberexperiments and new binary analysis capabilities. *Broader Impact:* This work created new features and new support for DeterLab's research and educational users, and also improved robustness and security of the DeterLab testbed.

Elements: Software: Distributed Workflows for Cyberexperimentation (OAC-1835608, $ 598,798, 08/2018 - 08/2021), is PI Mirkovic's project, collaborative with Genevieve Bartlett (USC/ISI) and James Blythe (USC/ISI). *Intellectual Merit:* This project develops distributed experimentation workflows (DEWs) for cyber experiments. To date it resulted in translators between bash and MAGI scripts and the DEW format, as well as the design and a prototype of the UI for experiment creation using DEW. *Broader Impact:* This work will improve repeatability, reproducibility and reuse of cyber experiments.

SaTC-CCRI: Sharing Expertise and Artifacts for Reuse through Cybersecurity Community Hub (SaTC-1925773, $ 425,000, 08/2019 – 08/2022), is PI Mirkovic's project, collaborative with Terry Benzel (USC/ISI), Laura Tinnel (SRI), David Balenson (SRI), Tim Yardley (UIUC) and Eric Eide (University of Utah). *Intellectual Merit:* This effort builds a SEARCCH Hub–new community infrastructure that helps with the transfer and reuse of cybersecurity experiment knowledge through sharing of experiment artifacts, tools, datasets and through connecting experimenters into domain-specific communities. *Broader Impact:* This work will enable vertical development of cyber experiments, which will improve quality, maximize efficiency, and reduce the researchers' time and effort in cyber experimentation.

E. References

[1] DNS Flag Day. https://dnsflagday.net/.

[2] Rami Al-Dalky, Michael Rabinovich, and Kyle Schomp. A Look at the ECS Behavior of DNS Resolvers. In *Internet Measurement Conference*, 2019.

[3] Mark Allman. Comments on dns robustness. In *Proceedings of the Internet Measurement Conference 2018*, IMC '18, pages 84–90, New York, NY, USA, 2018. ACM.

[4] Sumayah A Alrwais, Alexandre Gerber, Christopher W Dunn, Oliver Spatscheck, Minaxi Gupta, and Eric Osterweil. Dissecting ghost clicks: Ad fraud via misdirected human clicks. In *Proceedings of the 28th Annual Computer Security Applications Conference*, pages 21–30. ACM, 2012.

[5] Manos Antonakakis, Roberto Perdisci, David Dagon, Wenke Lee, and Nick Feamster. Building a dynamic reputation system for dns. In *Proceedings of the 19th USENIX Conference on Security*, USENIX Security'10, pages 18–18, Berkeley, CA, USA, 2010. USENIX Association.

[6] Manos Antonakakis, Roberto Perdisci, Wenke Lee, Nikolaos Vasiloglou, II, and David Dagon. Detecting malware domains at the upper dns hierarchy. In *Proceedings of the 20th USENIX Conference on Security*, SEC'11, pages 27–27, Berkeley, CA, USA, 2011. USENIX Association.

[7] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. DNS Security Introduction and Requirements. RFC 4033 (Proposed Standard), March 2005. Updated by RFCs 6014, 6840.

[8] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. Protocol Modifications for the DNS Security Extensions. RFC 4035 (Proposed Standard), March 2005. Updated by RFCs 4470, 6014, 6840, 8198.

[9] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. Resource Records for the DNS Security Extensions. RFC 4034 (Proposed Standard), March 2005. Updated by RFCs 4470, 6014, 6840, 6944.

[10] Eric Battenberg and David Wessel. Analyzing Drum Patterns Using Conditional Deep Belief Networks. In *Proc. of the International Society for Music Information Retrieval Conference*, 2012.

[11] Saleem N Bhatti and Randall Atkinson. Reducing dns caching. In *Proceedings of the Global Internet Workshop*, 2011.

[12] Leyla Bilge, Engin Kirda, Christopher Kruegel, Marco Balduzzi, and Sophia Antipolis. Exposure: Finding malicious domains using passive dns analysis. In *In Annual Network and Distributed System Security Symposium (NDSS*, 2011.

[13] National Science Board. Science and Engineering Indicators 2018. https://nsf.gov/statistics/2018/nsb20181/report/sections/higher-education-in-science-and-engineering/undergraduate-education-enrollment-and-degrees-in-the-united-states, 2018.

[14] N. Brownlee, K. C. Claffy, and E. Nemeth. Dns measurements at a root server. In *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, volume 3, 2001.

[15] Timm Bttger, Felix Cuadrado, Gianni Antichi, Eder Leo Fernandes, Gareth Tyson, Ignacio Castro, and Steve Uhlig. An Empirical Study of the Cost of DNS-over-HTTPS. In *Internet Measurement Conference*, 2019.

[16] Tom Callahan, Mark Allman, and Michael Rabinovich. On Modern DNS Behavior and Properties. *ACM Computer Communication Review*, 43(3), July 2013.

[17] Alecia Carter, Alyssa Croft, Dieter Lukas, and Gillian Sandstrom. Womens visibility in academic seminars: Women ask fewer questions than men. *PLOS ONE*, 13, 11 2017.

[18] Sebastian Castro, Duane Wessels, Marina Fomenkov, and Kimberly C. Claffy. A day at the root of the internet. *Computer Communication Review*, 38(5):41–46, 2008.

[19] Sebastian Castro, Min Zhang, Wolfgang John, Duane Wessels, and Kimberly C. Claffy. Understanding and preparing for DNS evolution. In *Traffic Monitoring and Analysis, Second International Workshop, TMA 2010, Zurich, Switzerland, April 7, 2010, Proceedings*, pages 1–16, 2010.

[20] Qi Alfred Chen, Eric Osterweil, Matthew Thomas, and Z Morley Mao. Mitm attack by name collision: Cause analysis and vulnerability assessment in the new gtld era. In *2016 IEEE Symposium on Security and Privacy (SP)*, pages 675–690. IEEE, 2016.

[21] Qi Alfred Chen, Matthew Thomas, Eric Osterweil, Yulong Cao, Jie You, and Z Morley Mao. Client-side name collision vulnerability in the new gtld era: A systematic study. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 941–956. ACM, 2017.

[22] CobbLiu. Dnsperf. https://github.com/cobblau/dnsperf/.

[23] ICANN Government Advisory Committee. Dns abuse mitigation. https://gac.icann.org/activity/dns-abuse-mitigation.

[24] Root Server System Advisory Committee. Report from the 1^{st} RSSAC workshop, January 2016.

[25] Root Server System Advisory Committee. Report from the 2^{nd} RSSAC workshop, June 2016.

[26] Root Server System Advisory Committee. RSSAC statement concerning the impact of the unavailability of a single root server, September 2016.

[27] computerscience.org. Women in computer science: Getting involved in stem. https://www.computerscience.org/resources/women-in-computer-science/, 2019.

[28] Hugo M Connery. DNS Response Policy Zones history, overview, usage and research, 2013.

[29] I. Corona, R. Perdisci, W. Lee, and D. Dagon. Detecting malicious flux service networks through passive analysis of recursive dns traces. In *Computer Security Applications Conference, Annual(ACSAC)*, volume 00, pages 311–320, 12 2009.

[30] J. Damas, M. Graff, and P. Vixie. Extension Mechanisms for DNS (EDNS(0)). RFC 6891 (Internet Standard), April 2013.

[31] J. Damas and F. Neves. Preventing Use of Recursive Nameservers in Reflector Attacks. RFC 5358 (Best Current Practice), October 2008.

[32] Wouter B. de Vries, Ricardo de O. Schmidt, Wes Hardaker, John Heidemann, Pieter-Tjerk de Boer, and Aiko Pras. Verfploeter: Broad and load-aware anycast mapping. In *Proceedings of the ACM Internet Measurement Conference*, London, UK, 2017.

[33] Xiyue Deng and Jelena Mirkovic. Commoner privacy and a study on network traces. In *Proceedings of the 33rd Annual Computer Security Applications Conference*, pages 566–576. ACM, 2017.

[34] Xiyue Deng, Hao Shi, and Jelena Mirkovic. Understanding malware network behaviors using fantasm. In *Proceedings of LASER 2017 Learning from Authoritative Security Experiment Results*, 2017.

[35] Frank Denis. Dnsblast. https://github.com/jedisct1/dnsblast.

[36] DNS-OARC. Day in the life of the internet. https://www.dns-oarc.net/oarc/data/ditl.

[37] DNS-OARC. Dns replay tool (drool). https://github.com/DNS-OARC/drool.

[38] V. Dukhovni and W. Hardaker. SMTP Security via Opportunistic DNS-Based Authentication of Named Entities (DANE) Transport Layer Security (TLS). RFC 7672 (Proposed Standard), October 2015.

[39] V. Dukhovni and W. Hardaker. The DNS-Based Authentication of Named Entities (DANE) Protocol: Updates and Operational Guidance. RFC 7671 (Proposed Standard), October 2015.

[40] Zakir Durumeric, Eric Wustrow, and J. Alex Halderman. Zmap: Fast internet-wide scanning and its security applications. In *Presented as part of the 22nd USENIX Security Symposium (USENIX Security 13)*, pages 605–620, Washington, D.C., 2013. USENIX.

[41] Dyn. Dyn's statement on the 10/21/2016 dns ddos attack. https://dyn.com/blog/dyn-statement-on-10212016-ddos-attack/.

[42] D. Eastlake 3rd and M. Andrews. Domain Name System (DNS) Cookies. RFC 7873 (Proposed Standard), May 2016.

[43] D. Crocker (Ed.), T. Hansen (Ed.), and M. Kucherawy (Ed.). DomainKeys Identified Mail (DKIM) Signatures. RFC 6376 (Internet Standard), September 2011.

[44] R. Braden (Ed.). Requirements for Internet Hosts - Application and Support. RFC 1123 (Internet Standard), October 1989. Updated by RFCs 1349, 2181, 5321, 5966, 7766.

[45] Pawel Foremski, Oliver Gasser, and Giovane Moura. DNS Observatory: The Big Picture of the DNS. In *Internet Measurement Conference*, 2019.

[46] Hongyu Gao, Vinod Yegneswaran, Yan Chen, Phillip Porras, Shalini Ghosh, Jian Jiang, and Haixin Duan. An empirical reexamination of global dns behavior. In *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM*, SIGCOMM '13, pages 267–278, New York, NY, USA, 2013. ACM.

[47] O. Gudmundsson. Adding Acronyms to Simplify Conversations about DNS-Based Authentication of Named Entities (DANE). RFC 7218 (Proposed Standard), April 2014.

[48] A. Gulbrandsen, P. Vixie, and L. Esibov. A DNS RR for specifying the location of services (DNS SRV). RFC 2782 (Proposed Standard), February 2000. Updated by RFC 6335.

[49] Gain Han and Keemin Sohn. Clustering the Seoul metropolitan area by travel patterns based on a deep belief network. In *Proc. of the International Conference on Big Data and Smart City (ICBDSC)*, pages 1–6, 2016.

[50] Ameya Hanamsagar, Simon S Woo, Chris Kanich, and Jelena Mirkovic. Leveraging Semantic Transformation to Investigate Password Habits and Their Causes. In *CHI*, 2018.

[51] Shuang Hao, Nick Feamster, and Ramakant Pandrangi. Monitoring the initial dns behavior of malicious domains. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, IMC '11, pages 269–278, New York, NY, USA, 2011. ACM.

[52] W. Hardaker. Use of SHA-256 in DNSSEC Delegation Signer (DS) Resource Records (RRs). RFC 4509 (Proposed Standard), May 2006.

[53] W. Hardaker. Requirements for Management of Name Servers for the DNS. RFC 6168 (Informational), May 2011.

[54] W. Hardaker. Child-to-Parent Synchronization in DNS. RFC 7477 (Proposed Standard), March 2015.

[55] W. Hardaker, O. Gudmundsson, and S. Krishnaswamy. DNSSEC Roadblock Avoidance. RFC 8027 (Best Current Practice), November 2016.

[56] John Heidemann. The ant lab: Analysis of network traffic. https://ant.isi.edu/.

[57] Paul Hoffman. Resolver testbed. https://github.com/icann/resolver-testbed.

[58] Bert Hubert. The dns camel. https://blog.apnic.net/2018/03/29/the-dns-camel/.

[59] ICANN. 2017 ksk rollover operational implementation plan. https://www.icann.org/en/system/files/files/ksk-rollover-operational-implementation-plan-22jul16-en.pdf.

[60] ICANN. Dnssec: Rolling the root zone key signing key. https://www.icann.org/news/blog/dnssec-rolling-the-root-zone-key-signing-key.

[61] ICANN. Domain abuse activity reporting. https://www.icann.org/octo-ssr/daar.

[62] ICANN. Ensuring the integrity of the top level of the dns through security best practices. https://www.icann.org/news/blog/changing-the-keys-to-the-domain-name-system-dns-root-zone.

[63] ICANN. Ksk rollover postponed. https://www.icann.org/news/announcement-2017-09-27-en.

[64] ICANN. Rfc8145 root trust anchor reports. http://root-trust-anchor-reports.research.icann.org/.

[65] ICANN. Root server attack on 6 February 2007. Factsheet, Internet Corporation for Assigned Names and Numbers, March 2007.

[66] ICANN. Plan for continuing the root ksk rollover. https://www.icann.org/en/system/files/files/plan-continuing-root-ksk-rollover-01feb18-en.pdf, 02 2018.

[67] ICANN. Root ksk rollover update. https://static.ptbl.co/static/attachments/169320/1520904771.pdf?1520904771, 03 2018.

[68] Jelena Mirkovic and Eric Kline and Peter Reiher. RESECT: Self-Learning Traffic Filters for IP Spoofing Defense. In *Proceedings of the 33rd Annual Conference on Computer Security Applications*, 2017.

[69] Samuel Jero, Endadul Hoque, David Choffnes, Alan Mislove, and Cristina Nita-Rotaru. Automated Attack Discovery in TCP Congestion Control Using a Model-guided Approach. In *Proc. of the Network and Distributed Systems Security Symposium*, 2018.

[70] Ben Jones, Nick Feamster, Vern Paxson, Nicholas Weaver, and Mark Allman. Detecting DNS Root Manipulation. In *Passive and Active Measurement Conference*, March 2016.

[71] Jaeyeon Jung, Emil Sit, Hari Balakrishnan, and Robert Morris. Dns performance and the effectiveness of caching. *IEEE/ACM Trans. Netw.*, 10(5):589–603, October 2002.

[72] Andrew Kalafut, Minaxi Gupta, Pairoj Rattadilok, and Pragneshkumar Patel. Surveying dns wildcard usage among the good, the bad, and the ugly. In Sushil Jajodia and Jianying Zhou, editors, *Security and Privacy in Communication Networks*, pages 448–465, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[73] S. Kitterman. Sender Policy Framework (SPF) for Authorizing Use of Domains in Email, Version 1. RFC 7208 (Proposed Standard), April 2014. Updated by RFC 7372.

[74] Maria Konte, Nick Feamster, and Jaeyeon Jung. Dynamics of online scam hosting infrastructure. In Sue B. Moon, Renata Teixeira, and Steve Uhlig, editors, *Passive and Active Network Measurement*, pages 219–228, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[75] Brian Krebs. Homegrown: Rustock botnet fed by u.s. firms. https://krebsonsecurity.com/tag/wholesale-internet-inc/, 2010.

[76] Marc Kührer, Thomas Hupperich, Jonas Bushart, Christian Rossow, and Thorsten Holz. Going wild: Large-scale classification of open dns resolvers. In *Proceedings of the 2015 Internet Measurement Conference*, IMC '15, pages 355–368, New York, NY, USA, 2015. ACM.

[77] Honglak Lee, Roger Grosse, Rajesh Ranganath, and Andrew Y. Ng. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pages 609–616, New York, NY, USA, 2009. ACM.

[78] Paul Lepe, Aashray Aggarwal, Jelena Mirkovic, Jens Mache, Richard Weiss, and David Weinmann. Measuring Student Learning On Network Testbeds. In *Midscale Education and Research Infrastructure and Tools (MERIT) Workshop*, 2019.

[79] Daiping Liu, Shuai Hao, and Haining Wang. All your dns records point to us: Understanding the security threats of dangling dns records. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, CCS '16, pages 1414–1425, New York, NY, USA, 2016. ACM.

[80] Chaoyi Lu, Baojun Liu, Zhou Li, Shuang Hao, Haixin Duan, Mingming Zhang, Chunying Leng, Ying Liu, Zaifeng Zhang, and Jianping Wu. An End-to-End, Large-Scale Measurement of DNS-over-Encryption: How Far Have We Come? In *Internet Measurement Conference*, 2019.

[81] Wired Magazine. Women and minorities in tech, by the numbers. https://www.wired.com/story/computer-science-graduates-diversity/, 2018.

[82] Allison Master, Sapna Cheryan, and Andrew N. Meltzoff. Computing whether she belongs: Stereotypes undermine girls' interest and sense of belonging in computer science. *Journal of Educational Psychology*, 108:424–437, 2015.

[83] M. Mealling. Dynamic Delegation Discovery System (DDDS) Part Three: The Domain Name System (DNS) Database. RFC 3403 (Proposed Standard), October 2002.

[84] M. Mealling and R. Daniel. The Naming Authority Pointer (NAPTR) DNS Resource Record. RFC 2915 (Proposed Standard), September 2000. Obsoleted by RFCs 3401, 3402, 3403, 3404.

[85] M. Mehta, K. Thapar, G. Oikonomou, and J. Mirkovic. Combining Speak-up with DefCOM for Improved DDoS Defense. In *Proceedings of the IEEE International Conference on Communications (ICC)*, 2008.

[86] J. Mirkovic, A. Hussain, S. Fahmy, P. Reiher, and R. Thomas. Accurately Measuring Denial of Service in Simulation and Testbed Experiments. *IEEE Transactions on Dependable and Secure Computing*, 6(2), 2009.

[87] J. Mirkovic and P. Reiher. A Taxonomy of DDoS Attacks and Defense Mechanisms. *ACM SIGCOMM Computer Communications Review*, 34(2), 2004.

[88] J. Mirkovic and P. Reiher. D-WARD: A Source-End Defense Against Flooding Denial-of-Service Attacks. *IEEE Transactions on Dependable and Secure Computing*, 2(3), 2005.

[89] J. Mirkovic, M. Robinson, P. Reiher, and G. Kuenning. Alliance Formation for DDoS Defense. In *Proceedings of the ACM SIGSAC New Security Paradigms Workshop (NSPW)*, 2003.

[90] Jelena Mirkovic and Terry Benzel. Deterlab testbed for cybersecurity research and education. *J. Comput. Sci. Coll.*, 28(4):163–163, April 2013.

[91] Jelena Mirkovic, Sven Dietrich, David Dittrich, and Peter Reiher. *Internet Denial of Service: Attack and Defense Mechanisms*. Prentice Hall PTR, ISBN 0-13-147573-8, 2005.

[92] P. Mockapetris and K. J. Dunlap. Development of the domain name system. In *SIGCOMM '88*, 1988.

[93] P.V. Mockapetris. Domain names - concepts and facilities. RFC 1034 (Internet Standard), November 1987. Updated by RFCs 1101, 1183, 1348, 1876, 1982, 2065, 2181, 2308, 2535, 4033, 4034, 4035, 4343, 4035, 4592, 5936, 8020.

[94] P.V. Mockapetris. Domain names - implementation and specification. RFC 1035 (Internet Standard), November 1987. Updated by RFCs 1101, 1183, 1348, 1876, 1982, 1995, 1996, 2065, 2136, 2181, 2137, 2308, 2535, 2673, 2845, 3425, 3658, 4033, 4034, 4035, 4343, 5936, 5966, 6604, 7766.

[95] Giovane C. M. Moura, John Heidemann, Ricardo de O. Schmidt, and Wes Hardaker. Cache me if you can: Effects of DNS Time-to-Live (extended). Technical Report ISI-TR-734b, USC/Information Sciences Institute, July 2019. Released May 2019, updated Sept. 2019.

[96] Engineering National Academies of Sciences and Medicine (U.S.). Committee on the Growth of Computer Science Undergraduate Enrollments. *Assessing and responding to the growth of computer science undergraduate enrollments*. The National Academies Press, 2018.

[97] Shaun Nichols. AWS DNS network hijack turns MyEtherWallet into ThievesEtherWallet. The Register, April 2018.

[98] Bureau of Labor Statistics. Computer and Information Technology Occupations. https://www.bls.gov/ooh/computer-and-information-technology/home.htm, 2019.

[99] G. Oikonomou and J. Mirkovic. Modeling Human Behavior for Defense against Flash Crowd Attacks. In *Proceedings of the IEEE International Conference on Communications (ICC)*, 2009.

[100] B-Root Operations. Usc statement of intent on dns data sharing. https://b.root-servers.org/statements/data.html.

[101] Eric Osterweil, Shane Amante, Dan Massey, and Danny McPherson. The great ipv4 land grab: resource certification for the ipv4 grey market. In *Proceedings of the 10th ACM Workshop on Hot Topics in Networks*, page 12. ACM, 2011.

[102] Eric Osterweil, Burt Kaliski, Matt Larson, and Danny McPherson. Reducing the x. 509 attack surface with dnssecs dane. *SATIN: Securing and Trusting Internet Names (March 2012)*, 2012.

[103] Eric Osterweil, Dan Massey, and Lixia Zhang. Observations from the dnssec deployment. In *2007 3rd IEEE Workshop on Secure Network Protocols*, pages 1–6. IEEE, 2007.

[104] Eric Osterweil, Dan Massey, and Lixia Zhang. Availability problems in the dnssec deployment, 2009.

[105] Eric Osterweil, Dan Massey, and Lixia Zhang. Deploying and monitoring dns security (dnssec). In *2009 Annual Computer Security Applications Conference*, pages 429–438. IEEE, 2009.

[106] Eric Osterweil, Danny McPherson, Steve DiBenedetto, Christos Papadopoulos, and Dan Massey. Behavior of dnstop talkers, a. com/. net view. In *International Conference on Passive and Active Network Measurement*, pages 211–220. Springer, Berlin, Heidelberg, 2012.

[107] Eric Osterweil, Danny McPherson, and Lixia Zhang. Operational implications of the dns control plane. *IEEE Reliability Society Newsletter*, 2011.

[108] Eric Osterweil, Danny McPherson, and Lixia Zhang. The shape and size of threats: Defining a networked system's attack surface. In *2014 IEEE 22nd International Conference on Network Protocols*, pages 636–641. IEEE, 2014.

[109] Eric Osterweil, Vasileios Pappas, Dan Massey, and Lixia Zhang. Zone state revocation for dnssec. In *Proceedings of the 2007 workshop on Large scale attack defense*, pages 153–160. ACM, 2007.

[110] Eric Osterweil, Michael Ryan, Dan Massey, and Lixia Zhang. Quantifying the operational status of the dnssec deployment. In *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*, pages 231–242. ACM, 2008.

[111] Eric Osterweil and Lixia Zhang. Interadministrative challenges in managing dnskeys. *IEEE Security & Privacy*, 7(5):44–51, 2009.

[112] PARSONS. Dnssec-tools. https://www.dnssec-tools.org/.

[113] Craig Partridge and Mark Allman. Ethical considerations in network measurement papers. *Commun. ACM*, 59(10):58–64, September 2016.

[114] Vern Paxson. An analysis of using reflectors for distributed denial-of-service attacks. *SIGCOMM Comput. Commun. Rev.*, 31(3):38–47, July 2001.

[115] Nicole Perlroth. Hackers used new weapons to disrupt major websites across U.S. *New York Times*, page A1, Oct. 22 2016.

[116] Amreesh Phokeer, Alain Aina, and David L. Johnson. DNS Lame Delegations: A Case-Study of Public Reverse DNS Records in the African Region. In *AFRICOMM*, 2016.

[117] Dave Piscitello. Apwg and m3aawg survey finds icann whois changes impede cyber investigations. https://www.securityskeptic.com/2018/10/apwg-and-m3aawg-survey-finds-icann-whois-changes-impede-cyber-investigations.html.

[118] Dave Piscitello. Facts & figures: Whois policy changes impair blocklisting defenses. https://www.securityskeptic.com/2019/03/index.html.

[119] Deter Project. Deterlab education portal. https://isi.deterlab.net/education.php.

[120] Deter Project. Deterlab testbed. https://isi.deterlab.net/.

[121] Moheeb Abu Rajab, Fabian Monrose, Andreas Terzis, and Niels Provos. Peeking through the cloud: Dns-based estimation and its applications. In Steven M. Bellovin, Rosario Gennaro, Angelos Keromytis, and Moti Yung, editors, *Applied Cryptography and Network Security*, pages 21–38, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[122] Root Server Operators. Events of 2015-11-30. Technical report, Root Server Operators, Dec. 4 2015.

[123] Root Server Operators. Events of 2016-06-25. Technical report, Root Server Operators, June 29 2016.

[124] Christian Rossow. Amplification Hell: Revisiting Network Protocols for DDoS Abuse. In *In Proceedings of the 2014 Network and Distributed System Security Symposium, NDSS*, 2014.

[125] RSSAC. RSSAC037: A proposed governance model for the dns root server system, June 2018.

[126] RSSAC. RSSAC038: Rssac advisory on a proposed governance model for the dns root server system, June 2018.

[127] RSSAC. RSSAC025: RSSAC statement on root server operator independence, May 2019.

[128] Kazumichi Sato, Keisuke Ishibashi, Tsuyoshi Toyono, and Nobuhisa Miyake. Extending black domain name list by using co-occurrence relation between dns queries. In *Proceedings of the 3rd USENIX Conference on Large-scale Exploits and Emergent Threats: Botnets, Spyware, Worms, and More*, LEET'10, pages 8–8, Berkeley, CA, USA, 2010. USENIX Association.

[129] Kyle Schomp, Mark Allman, and Michael Rabinovich. DNS Resolvers Considered Harmful. In *ACM SIGCOMM HotNets*, October 2014.

[130] Kyle Schomp, Tom Callahan, Michael Rabinovich, and Mark Allman. On Measuring the Client-Side DNS Infrastructure. In *ACM SIGCOMM/USENIX Internet Measurement Conference*, October 2013.

[131] Kyle Schomp, Tom Callahan, Michael Rabinovich, and Mark Allman. Assessing DNS Vulnerability to Record Injection. In *Passive and Active Measurement Conference*, March 2014.

[132] Kyle Schomp, Michael Rabinovich, and Mark Allman. Towards a Model of DNS Client Behavior. In *Passive and Active Measurement Conference*, March 2016.

[133] Hao Shi and Jelena Mirkovic. Handling anti-virtual machine techniques in malicious software. *IEEE/ACM Trans. on Privacy and Security*, 2018.

[134] M. StJohns. Automated Updates of DNS Security (DNSSEC) Trust Anchors. RFC 5011 (Internet Standard), September 2007.

[135] S. Thomson, C. Huitema, V. Ksinant, and M. Souissi. DNS Extensions to Support IP Version 6. RFC 3596 (Internet Standard), October 2003.

[136] Mark Utting and Bruno Legeard. *Practical Model-Based Testing: A Tools Approach*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2007.

[137] Verisign. The Domain Name Industry Brief (DNIB). *The Verisign Domain Report*, Volume 16(Issue 2), May 2019.

[138] Paul Vixie. Response rate limiting in the domain name system (dns rrl), June 2012.

[139] D. Wessels, W. Kumari, and P. Hoffman. Signaling Trust Anchor Knowledge in DNS Security Extensions (DNSSEC). RFC 8145 (Proposed Standard), April 2017.

[140] Duane Wessels. A look at rfc 8145 trust anchor signalling for the 2017 ksk rollover. https://indico.dns-oarc.net/event/27/session/1/contribution/11, 09 2017.

[141] Duane Wessels and Marina Fomenkov. Wow, that's a lot of packets. In *Proc. of Passive and Active Measurement Workshop*, 2003.

[142] Duane Wessels, Marina Fomenkov, Nevil Brownlee, and kc claffy. Measurements and laboratory simulations of the upper dns hierarchy. In Chadi Barakat and Ian Pratt, editors, *Passive and Active Network Measurement*, pages 147–157, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

[143] Jun Xu, Jinliang Fan, Mostafa H. Ammar, and Sue B. Moon. Prefix-preserving IP address anonymization: Measurement-based security evaluation and a new cryptography-based scheme. In *Proceedings of the 10th IEEE International Conference on Network Protocols*, pages 280–289, Washington, DC, USA, November 2002. IEEE.

[144] Sandeep Yadav and A. L. Narasimha Reddy. Winning with dns failures: Strategies for faster botnet detection. In Muttukrishnan Rajarajan, Fred Piper, Haining Wang, and George Kesidis, editors, *Security and Privacy in Communication Networks*, pages 446–459, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[145] Sandeep Yadav, Ashwath Kumar Krishna Reddy, A.L. Narasimha Reddy, and Supranamaya Ranjan. Detecting algorithmically generated malicious domain names. In *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*, IMC '10, pages 48–61, New York, NY, USA, 2010. ACM.

[146] He Yan, Eric Osterweil, Jon Hajdu, Jonas Acres, and Dan Massey. Limiting replay vulnerabilities in dnssec. In *2008 4th Workshop on Secure Network Protocols*, pages 3–8. IEEE, 2008.

[147] Hao Yang, Eric Osterweil, Dan Massey, Songwu Lu, and Lixia Zhang. Deploying cryptography in internet-scale systems: A case study on dnssec. *IEEE Transactions on Dependable and Secure Computing*, 8(5):656–669, 2010.

[148] Ming Zhang, Yaoping Ruan, Vivek Pai, and Jennifer Rexford. How dns misnaming distorts internet topology mapping. In *Proceedings of the Annual Conference on USENIX '06 Annual Technical Conference*, ATEC '06, pages 34–34, Berkeley, CA, USA, 2006. USENIX Association.

[149] Liang Zhu. Dns trace replay client. https://ant.isi.edu/software/ldplayer/dns-replay-client/index.html.

[150] Liang Zhu. Dns trace replay client. https://ant.isi.edu/software/ldplayer/dns-replay-controller/index.html.

[151] Liang Zhu and John Heidemann. Ldplayer: Dns experimentation at scale. In *Proceedings of the SIGCOMM Posters and Demos*, SIGCOMM Posters and Demos '17, pages 60–62, New York, NY, USA, 2017. ACM.

[152] Zhaosheng Zhu, Vinod Yegneswaran, and Yan Chen. Using failure information analysis to detect enterprise zombies. In Yan Chen, Tassos D. Dimitriou, and Jianying Zhou, editors, *Security and Privacy in Communication Networks*, pages 185–206, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.