Prediction of the Domain Name System (DNS) Quality Attributes

Marwan Radwan
University of Leicester
University Road
Leicester, LE1 7RH
+447833913122
mmmr1@le.ac.uk

Reiko Heckel
University of Leicester
University Road
Leicester, LE1 7RH
+44(0)1162523406
reiko@mcs.le.ac.uk

ABSTRACT
The Domain Name System (DNS) has a direct impact on the performance and dependability of nearly all aspects of interactions on the Internet. DNS relies on a delegation-based architecture, where resolution of a name to its IP address requires resolving the names of the servers responsible for that name. The graphs of the inter-dependencies that exist between name servers associated with each zone are called *Dependency Graphs*. We constructed a *DNS Dependency Model* as a unified representation of these Dependency Graphs. We utilize a set of *Structural Metrics* defined over this model as indicators of external quality attributes of the domain name system. We explore the inter-metric and inter-quality relations further in order to quantify the indicative power of each metric. We apply some machine learning algorithms in order to construct *Prediction Models* of the perceived quality attributes of the operational system out of the structural metrics of the model. Assessing these quality attributes at an early stage of the design/deployment enables us to avoid the implications of defective and low-quality designs and deployment choices and identify configuration changes that might improve the availability, security, stability and resiliency postures of the DNS.

CCS Concepts
• **Networks→ Network Services → Network Management**

Keywords
Domain Name System, Dependency Graphs, DNS Qualities, Predictive Models.

1. INTRODUCTION

The Domain Name System (DNS) is one of the most fundamental infrastructures of today's Internet. The critical importance of the DNS makes it demanding for its availability, stability, security and resilience. The DNS is a distributed database for storing information on domain names, the primary namespace for hosts on the Internet. The name space is organized in a hierarchical structure to ensure domain name uniqueness. Each node in the DNS tree corresponds to a zone. Each zone belonging to a single administrative authority is served by multiple authoritative name servers.

DNS relies on a delegation-based architecture, where resolution of a name to its IP address requires resolving the names of the servers responsible for that name. Delegation is crucial in achieving DNS name space's scalability. DNS delegation dependencies were introduced in [4], which observed that resolving a single domain name often requires traversing multiple other domains. Failure to resolve the domains in the dependency chain, e.g., due to misconfiguration or attack, or resolution to an incorrect address, may impact all the dependent domains. Delegation dependencies also nullify effectiveness of DNSSEC, [2], and hurdle its adoption [13]. If name servers or other resources of a signed zone are placed under unsigned domains, the DNS resolver will not be able to establish the security of the signed records, and the security will depend on the security of the weakest link in the dependency graph. Another notable side effect of a large dependency graph is the fact that it introduces more latency to resolution of DNS records, and increase the network traffic due to sending queries to many multiple name servers.

While DNS plays a critical role for the operation of the Internet, DNS zone administration relies heavily on error-prone manual configurations. A mistake in configuring a specific DNS zone may potentially have adverse impacts on the global Internet [3,8,11,17]. Operational guidelines [11, 20, 22] require that a zone have multiple authoritative name servers, and that they be distributed through diverse network topological and geographical locations to increase the reliability of that zone as well as improve overall network performance and access. It also makes DNS services robust against unexpected failures. Recent work [2, 10, 13] outlines the need for zone operators to understand how many delegation dependencies they may inadvertently be incurring through the deployment and sharing of DNS secondary servers. Choosing servers with names under other zones provides zone redundancy but may incur security and resiliency threats to the zone. Deciding on where to physically locate the servers should ensure a certain degree of resistance against different types of failures. Peering with external organizations for secondary server hosting should take into consideration the impact of transitional trust and administrative complexity [2, 4, 7].

This research is motivated by the need to avoid the implications of misconfigurations and bad deployment choices made by system administrators that may lead to data inconsistencies, vulnerable

configurations or even failure of resolution at an early stage of the design/deployment of the DNS. Efforts to improve risk management related to DNS security, stability and resiliency must be guided by an ability to predict these characteristics.

In this paper, a three-step process is conducted:

1. Investigating whether DNS model structural metrics are correlated with perceived DNS quality attributes through the employment of classical statistical correlation techniques.
2. Building prediction models for the various quality attributes out of the DNS model structured metrics.
3. Evaluating the developed predictive models in terms of their accuracy, sensitivity and other performance indicators.

Results obtained from the study support the idea that significant correlation exists between a set of structural metrics of the DNS model and the subjective perception of the participants about the quality attributes of the DNS. It also confirms that prediction models can be effectively built for the purpose of predicting DNS quality attributes using this set of structural metrics at early stages of the system design and deployment.

The rest of the paper is structured as follows: Section 2 discusses relevant background about the structure and operation of the DNS system. Section 3 presents the DNS dependency model with its main components, features and relationships. Section 4 introduces the DNS structural metrics suite and the metrics interpretation model. Section 5 details the definition of the concerned four quality attributes. Section 6 details our experiment and the developed predictive models for the various DNS quality attributes. Section 7 concludes the paper with directions of future work.

2. BACKGROUND

DNS is responsible for the mapping of human-friendly domain names to the corresponding machine-oriented IP addresses. Operators of each zone determine the number of authoritative name servers and their placement and manage all changes to the zone's data content.

2.1 General Operation of the DNS

Figure 1 shows the process by which an application looks up the domain name www.le.ac.uk and how it is mapped to the DNS data, control and management operational planes. To find the IP address of www.le.ac.uk, the client (e.g. a web browser) submits a DNS query to a DNS resolver (step 1). Assuming that the corresponding IP is not in the resolver cache, it will ask one of the root name servers for the translation (step 2). The names and IP addresses of root name servers are locally stored within each server. The root name servers will respond with a "referral", telling the resolver to query the DNS servers of the .uk zone for an answer (step 3). The resolver then repeats this process for the .uk name servers and get a referral to ask the .ac.uk name servers which in turn answers with a referral to as the le.ac.uk name servers (step 4 -7). The resolver next asks one of the le.ac.uk name servers for the translation (step 8), and gets the answer in step (9), and finally forwards the answer to the requesting client (step 10) and gets the answer in step (9), and

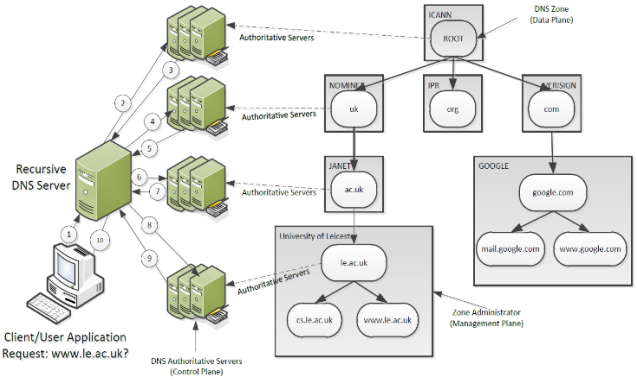

Figure 1 An Illustration of the DNS Resolution Process.

finally forwards the answer to the requesting client (step 10) who will use this information to connect to the web server hosting the web site www.le.ac.uk. Throughout the process, resolvers may encounter name servers hosted under other zones whose names need to be resolved before contacting them about the original request.

2.2 Dependency Graphs

Inter-dependencies are common in the DNS and stem from the hierarchal structure of the DNS, the DNS protocol as well as from different motivations and goals [1, 2 , 6 ,12]. A zone is said to depend on a name server if the name server could be involved in the resolution of names in that zone. The dependencies among name servers that directly or indirectly affect a zone are represented as a dependency graph. A dependency graph [16] is a directed connected graph with a distinguished node (r) which is the root zone. Each node in the graph represents a zone name, and each edge signifies that its source is directly dependent on its target for proper resolution of itself and any descendant domain names.

Since many of the misconfigurations can't be detected from the zone file directly, there is a need for a model that encompasses all information related to the zone file and the server deployments in one conceptual graph. The instance of the model (the dependency graph) will enable us to detect zone integrity violations as well as violations in the deployment of name servers and the choice of peering organizations and management structures. The conceptual graph representation facilitates modelling at multiple levels of details simultaneously.

3. DNS DEPENDENCY MODEL

The DNS Dependency Model is an attempt to describe the Domain Name System for the purpose of a particular goal of detecting violations of the design and deployment principles at the authoritative level. The model is composed of the following elements:

- Operational Entities (e.g. resource records, zones, servers and organizations)
- Properties of entities such as (in-bailiwick and out-of-bailiwick name servers)
- Relations between the entities (e.g. access attributes such as dependability, containment, delegation and management)

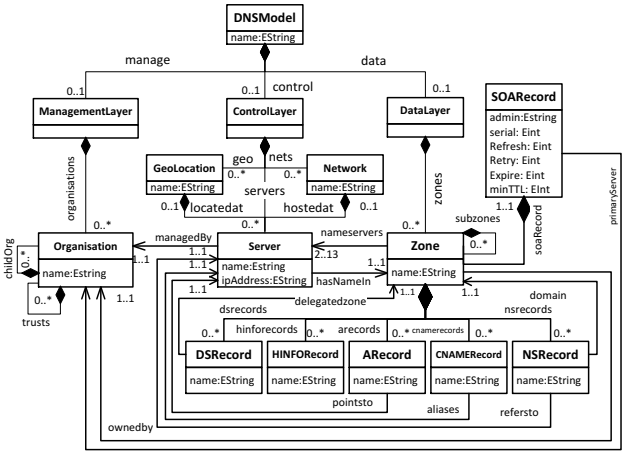

Figure 2: DNS Dependency Model Using Ecore Meta-Model.

The DNS entities that appear in our model as shown in Figure 2 fall into two categories: primitive and composed entities. Composed entities have an identity and a set of properties. In addition to these, composed entities have a list of contained entities, which are primitive or composed entities. A composed entity type is one that contains other entities. The model supports the following composed entities: Organization, Server, Zone and Resource Record.

In order to describe a composed entity we have to specify its properties, containment structure (i.e. the entities that it contains), relations and container entity. As an example, we can look at the server component where it can be managed (contained) by organizations. Multiple servers can be managed by one organization. The server can host many zone files and it has the name and IP address as attributes. There are many types of servers and in this context we are concerned with in-bailiwick server whose name is within the zone file hosted at that particular server and out-bailiwick server who has a name from a zone hosted in another name server.

Three specific dependencies are present within the DNS operational planes and they are the following:

- Parent Dependency: resolving the name of a domain name is always dependent on resolving its parent name since the resolver must learn the authoritative servers for a zone from referrals from the zone's hierarchical parent.
- Authoritative Name Server (NS) Dependency: A zone is said to depend on a name server if the name server could be involved in the resolution of names in that zone.
- CNAME Aliasing Dependency (Name pointing to another Name): the resolution of an alias is always dependent on the resolution of its target CNAME. If a resolver receives a response indicating that the name in question is an alias to another name, it must subsequently resolve the target of the alias, and so on until an address is returned.

4. DNS STRUCTURAL METRICS

Given the fact that that a single metric cannot capture all of the various aspects of a certain DNS quality attributes, in this study we propose a set of simple, yet intuitive, structural metrics defined over the DNS model as measurement references for various DNS quality attributes. Interesting metrics are per-server and per-zone distributions such as the number of zones that are served from multiple name servers in different network autonomous systems or diverse geographical locations (Server redundancy). The number of zones (direct and third party) that influence the resolving of domain names within a particular zone (*zone influence*).

Table 2 shows a subset of the *DNS structural Metrics Suite* covering the five main categories of size, data coherence, structural complexity, dependency and delegation/inheritance metrics. The complete suite is not shown here for space limitations but can be accessed through www.dnsmodel.ps/dnsmetrics/dnsmetrics.pdf. The main idea behind the design of these metrics has been comprehensiveness and simplicity. We have tried to cover as many structural characteristics of the DNS operational model as possible. To achieve this, structural metrics proposed in the areas of DNS management, object-oriented software design, software model design, and even business process models have been considered. For each structural metric defined over the DNS model, we give the metric definition, implementation and usability, how to measure, formula for computing as well as giving an example of such a metric.

Let's consider the *Administrative Complexity* [5] as an example of such metric. One important necessity for DNS proper operation is careful coordination between zone administrators and system managers hosting the authoritative name servers of the zone. Lack of such coordination can result in increased risk of failure. The coordination spans both hierarchically (i.e., between parent and child zones) and laterally, between organizations hosting each other's zone data (i.e, between name servers operators). There are two metrics used to quantify the complexity of a DNS zone. The first metric which measures the lateral complexity of a zone is *Administrative Complexity (AC)* which describes the diversity of a zone, with respect to organizations administering its authoritative servers. The second metric that measures the hierarchical complexity of the zone is the H*ierarchical Reduction Potential* (HRP) [16], which quantifies how much the ancestry of a zone might be reasonably consolidated to reduce hierarchical complexity. The interpretation model of the *Administrative Complexity* metric is shown in Table 1.

Table 1: Metric Administrative Complexity Interpretation Model.

Metric	Administrative Complexity
Definition	Describes the diversity of a zone with respect to the organisations administering its authoritative name servers.
Usability	The advantage mutual hosting of zones between organizations is an increased availability but at the same time increased potential of failure and instability of the zone resolution process.
How to Measure	Count the number of authoritative name servers managed by each organization involved in the dependency graph of zone (z).
Metric Notations	O_z: set of organizations administering authoritative name servers hosting zone (z); n: total number of authoritative name servers of zone (z); NS_z^o : the

	subset of name servers administered by organization o in O_z.
Formula	$Ac(z) = 1 - \sum_{o}^{n} \left(\frac{NS_z^o}{NS_z}\right)^n$

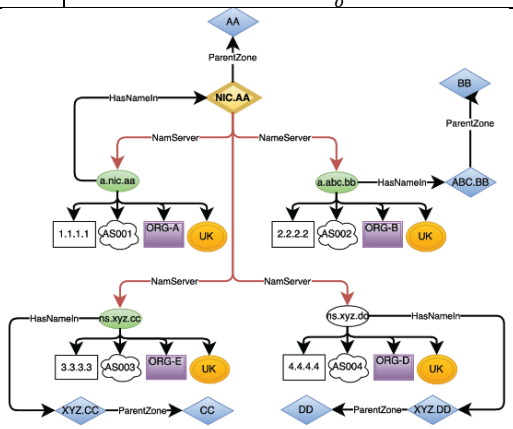

Figure 3: Measuring the Administrative Complexity (AC) Metric

Figure 3 shows an example of a simple *Dependency Graph* model for the zone *NIC.AA*. The Administrative Complexity of this zone is calculated by counting the number of directly configured authoritative name servers of the zone (z) that are managed by the same organisation and apply the formula in Table 1.

$$AC(\text{“}NIC.AA\text{”}) = 1 - \sum_{o=1}^{4} \left(\frac{1}{4}\right)^4 = 0.984375$$

The value of this metric is high since each authoritative name server of the zone is managed by a totally different organisation so the amount of coordination (or lateral complexity) needed for such a configuration is expected to be very high.

5. DNS QUALITY ATTRIBUTES

Measuring the quality attributes of the DNS and investigating its potential weaknesses is a question of crucial interest [19]. Best practices [11, 22] for ensuring availability and security of the DNS infrastructure recommend: (1) defining a number of name servers for each domain, (2) configuring these name servers under at least two different parent domains and (3) placing the physical name servers, hosting the zone files for the domain, in separate networks.

The redundancy provides for stability of the domain and prevents single point of failure. In particular, if one of the parent domains is not accessible, the domain will remain functional via the other parent domain; in case one of the networks, hosting the name servers, is under attack, the other name server, located in available networks, can be reached. On the flip side, while ensuring availability, this redundancy introduces new dependencies which can be utilized to attack the domain. Specifically, if a vulnerability exists in a network or a name server hosting the domain, it can be exploited to attack the domain, e.g., inject spoofed DNS record for domain hijacking.

Currently, there is little consensus on the right measures and acceptable performance levels for the DNS as a whole related to availability, security, stability and resiliency. Due to the fact that we are modelling the DNS system from the prospective of authoritative system administrators and zone managers, the *perceived quality* is defined as the quality of the DNS system as anticipated by the system administrator during the process of designing, configuring and deployment of the DNS system. Quality attributes have different definitions based on the point of view of the DNS user. For example, *resiliency* is viewed by users as availability and viewed by providers as a combination of detection, response, resistance and recovery processes that increase overall confidence in relying on and investing in the Internet over the long-term [19].

In this research, we focus on four quality attributes of the DNS as perceived by authoritative zone managers and system administrators and they are:

- **DNS Availability** is defined as the ability of the group of authoritative name servers of a particular zone (e.g., a TLD), to answer DNS queries. For the service to be considered available at a particular moment, at least two of the delegated name servers registered in the DNS must have successful results to each of their public-DNS registered “IP addresses” to which the name server resolves.
- **Security** is the ability of the components of the system to protect the integrity of DNS information and critical system resources.
- **Stability** is the consistency of the names of the authoritative servers names within the system and the consistency of the performance of the system components over time. That is, if the names of the authoritative servers within a system change with high frequency, the system is unstable and if a query takes 10 milliseconds to respond in one instance and 1000 milliseconds to respond in a second instance, resolution time is unstable which means the system is also unstable.
- **Resiliency** is the ability of the system to provide and maintain an acceptable level of name resolution service in the face of faults and changes in normal operating conditions.

Given the fact that the DNS protocol enables administrators and zone operators a high level of flexibility in configuring their zone and the deployment structure for their systems, it can be anticipated that low-quality configurations and deployment choices can ripple through to many operational domain name systems. Therefore, the need for early indicators of external quality attributes is recognized in order to avoid the implications of defective and low-quality design and deployment during the late stages of system operation. In the remainder of this paper, we will use an empirical experiment with advanced statistical analysis techniques to evaluate the efficiency of these metrics for external quality attribute prediction of the DNS.

6. EMPIRICAL EXPERIMENT

The purpose of this study is to identify any significant relationship between a set of structural metrics defined over a DNS model and the subjective perception of domain experts of the DNS quality attributes. Another purpose of the study is to investigate any inter-relationship between the various DNS quality attributes under study and to evaluate how well different prediction models based on the

proposed structural metrics can perform in indicating the perceived quality attributes of the system. These objectives will be achieved by conducting a controlled experimentation and employing a set of statistical analysis techniques.

6.1 Hypotheses

H1: A correlation exists between a set of DNS dependency model structural measures and a set of perceived quality attributes of the DNS.

H2: Prediction models based on the proposed structural metrics can be built to predict the quality attributes of the system.

6.2 Variables

In order to proceed with the experiment, the defined hypotheses need to be mapped onto a set of measurable independent and dependent variables. An independent variable is the variable that is changed or controlled in a scientific experiment to test the effects on the dependent variable. These variables are evaluated in the experiment and will be used in the analysis phase.

Independent Variables: Representative set of eleven structural metrics defined over the DNS model as shown in Table 2.

Table 2: List of structural metrics defined over the DNS Model.

Measure	Symbol	Description
Attack Surface	AS	Total number of unique zones, name servers and organizations.
Number of Name servers	ANS	Total number of authoritative name serves of a zone.
Network Diversity	NETD	Distinguished networks AS numbers which host the name servers of the zone.
Geographical Diversity	GEOD	Distinguished number of servers' geographical locations.
Redundancy	RED	The size of the smallest set of servers that if failed will render the zone unresolvable.
Administrative Complexity	AC	Administrative complexity related to the number of directly configured organizations.
Average Query Path	AQP	Average number of requests needed to query the name under the zone.
Direct Zones	DCZ	Number of directly configured zones influencing the name resolution of the zone.
Third Party Zones	TPZ	Number of zones influencing the zone resolution as a result of hosting the zone in servers with names under other zones.
Directly Configured Organizations	DCO	Number of organization directly configured by the zone administrator.
Third Party Organizations	TPO	Number of third party organizations involved in the query process of the zone.

To get metrics measurements, we used 10 different model instance of the DNS model and measured those metrics on each of them. We don't have a pre-defined store of such models and have to build them using our *DG-Builder* tool. We also built the dependency graphs for 15 Top-Level-Domains (TLDs) that are managed by the participants of our experiment. This group of TLDs has a diverse range of dependency graphs from small and compacted ones to large and widely spread ones.

Dependent Variables: Four external quality attributes of the DNS system (i.e. availability, security, stability and resiliency as defined in Section **Error! Reference source not found.**) are considered to be the dependent variables.

6.3 Collection of Data

The subjective opinions of the participants about the quality attributes of the DNS system were collected using an online questionnaire. During the period of the survey, the participants had the opportunity to ask questions to the experimenter. The questionnaire consisted of 45 questions divided in 3 sections. (1) Each participant was asked to answer about 10 general questions related to their experience with the DNS system as well as the TLD they are responsible for. (2) Then, the participant was asked to evaluate the perceived quality attributes of a set of 9 dependency graphs presented as instances of the DNS model. (3) Finally, the participants were asked to assess the quality attributes of the TLDs under their own management.

6.4 Participants

The participants were all TLD administrators responsible for managing one or more top-level domains. They were from different geographical locations with 5 from the Middle East, 5 from Europe, 1 from the Americas, 4 from the Asia Pacific region and 2 from Africa. Those administrators have a good range of DNS experience ranging from 2 to 10 years of experience. The TLDs managed by those administrators have various number of registered domain names ranges from a couple of thousand up to millions of domain names. It is clear that the set of participants are representative of a good spectrum of DNS operators around the world and their views can be effectively used in our experiment.

In order to establish the extent of consensus among the subjective opinions provided by the participants, we perform an inter-rater reliability analysis. We employ an intra-class correlation (ICC) which is used to assess the consistency, or conformity, of measurements made by multiple observers measuring the same quantity [23]. Table 3 reports the results of this statistical test based on a two way random effects model with a confidence interval of 95%.

Table 3: Intra-Class Correlation (ICC).

Quality Attribute	ICC Single Measure
Availability	0.705
Security	0.712
Stability	0.709
Resiliency	0.68

As seen in this table, the single measure reliability of the four quality attributes is higher than 0.67, which shows that a reasonable agreement between the participants exists in terms of the perceived values for these attributes for each of the objects of the study.

6.5 Metric-Quality Correlation Analysis

In this section, we will evaluate the first hypothesis which states that a meaningful correlation can be found between a set of DNS model structural measures and a set of quality attributes of the DNS system (H1). In order to test this hypothesis, we asked the participants to key in their views regarding the perceived quality attributes of a set of 9 DNS Dependency Model instances (i.e. Dependency Graphs). The models varied in terms of their metric values as shown in Table 4. The empirical data that were collected

are also quantitatively reasonable from the perspective of the amount of data. We obtained 540 data points from the subjective opinions of the participants regarding the models (9 dependency models, 15 participants, 4 quality attributes).

Table 4: Measurements of Metrics on the 9 DNS Model Instances

M#	AS	ANS	NETD	GEOD	Red	AC	AQP	DCZ	TPZ	DCO	TPO
M-1	34	3	3	1	3	0.89	4	3	5	3	6
M-2	21	4	4	1	4	0.98	3	4	4	4	3
M-3	13	4	4	1	4	0.84	2	2	0	4	1
M-4	10	4	1	4	4	0.43	2	2	0	1	1
M-5	19	3	2	1	3	0.44	4	4	1	2	3
M-6	10	4	4	4	4	0.5	2	1	1	4	0
M-7	15	6	2	2	2	0.89	2	2	0	2	1
M-8	16	2	1	1	2	0.5	2	4	1	1	2
M-9	21	8	8	8	8	0.84	2	2	0	8	1

The metric-quality correlation analysis shows that some of the metrics are in fact correlated to certain quality attributes with various coefficients. The technique that we explore is the use of Spearman's Rho correlation, namely to identify relationship between the measured metrics of the models and the four quality attributes. Spearman's Rho correlation coefficient is a statistical measure of the strength of a monotonic relationship between paired data and its value ranges from -1 to 1.

Table 5: Metric-Quality correlations (Spearman's Rho).

Metrics	Availability	Security	Stability	Resiliency
AS	-.819*	-.685*	-.757*	0.33
ANS	0.258	-0.079	-0.01	0.02
NETD	0.037	-0.273	-0.027	.666*
GEOD	-0.056	-0.302	-0.086	.777*
TPZ	-.828*	-.703*	-.743*	0.248
AQP	0.109	-0.011	-0.004	-0.129
RED	0.02	-0.252	0.127	0.185
AC	0.05	-0.177	0.094	-.536*
DCZ	-0.276	-0.479	-0.355	.685*
DCO	0.105	-0.225	0.045	.698*
TPO	-.768*	-.609*	-.739*	0.156

According to Spearman's correlation, a correlation with a significance value > 0.05 can be considered to be significant, and therefore, in our work, such correlations are considered to be meaningful and are highlighted as shown in Table 5 . As it can be seen, significant correlations can be found between some of the metrics and the four DNS quality attributes. This shows that the structural metrics defined for a DNS model can be used as early indicators for external quality attributes of the DNS. In addition, the correlations can be explained by the following two points:

- Metrics that reflect third party influence (as a result of peering with external organizations for secondary server hosting and placing servers under third party zones) such as AS, TPO and TPZ has clear negative impact on the availability, security and stability of the DNS. Choosing servers with names under other zones (increasing third party zones) provides zone redundancy but may incur security and stability threats to the zone due to increasing the Attack Surface (AS) metric of the model. DNS administrators should try to avoid such practice by reducing the size of their dependency graph (AS metric) by placing authoritative name servers for a certain zone under the same zone.
- Physically distributing the servers (geographical and network wise diversity metrics) ensures a certain degree of resistance against different types of failures and subsequently have positive impact on the resiliency of the whole system. Resiliency of the DNS is positively correlated with those metrics that are directly configured by the system administrator such as (GeoD, NetD, AC, DCZ and DCO). DNS administrators has to pay more attention regarding the deployment of their servers geographically and from a network distribution prospective. Also coordination with peer hosting organisations is vital in case of failures and the necessity to reduce this metric and consequently reduce zone complexity is clear to guarantee a higher level of resiliency of the system.

6.6 Prediction Models

In this section, we will apply some machine learning algorithms in order to construct prediction models of the quality attributes of the DNS system out of the structural metrics of the dependency model (H2).

Predictive models are created to best predict the probability of an outcome based on some prior observations. We built predictive models based on the methodology outlined in Figure 4. The models that are developed are based on the Random Forest (RF) decision tree, Simple Logistic (SL) functions, Lazy LWL and Rule-based PART [14]. These models take the structural metrics of a DNS model instance as input and try to find the most relevant value of the quality attribute for the given model. We employ WEKA [15] to train and test our predictive models. For each of the four quality attributes, one instance of each of the mentioned predictive models is developed (4 model types and 4 characteristics = 16 predictive models).

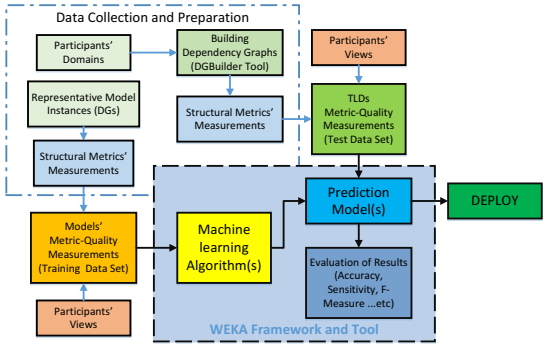

Figure 4: Methodology of Building DNS Quality Prediction Models

We used the measured structural metrics of the 9 models with the perceived quality attributes as keyed in by the participants as the training dataset for the prediction models. As far as the test dataset is concerned, we constructed the Dependency Graphs of the 15 participants' TLDs using our *DGBuilder* tool and then measured the various structural metrics on these models. We combined this data with the perceived quality attributes from the participants concerning their own TLDs to construct the test dataset. Figure 4

shows the methodology we used in order to build the prediction models for the DNS quality attributes out of a set of structural metrics defined over the DNS model. The two data sets used in this experiment are totally independent and they can be effectively used to train the models and test their performance.

Table 6 Performance of the Predictive Models in terms of the correctly classified instances out of the test dataset

Classifier Name	Availability	Security	Stability	Resiliency
RF	73%	47%	53%	40%
LWL	53%	53%	67%	33%
SL	7%	53%	20%	27%
PART	20%	73%	20%	73%

Figure 5 shows the different parameters used to evaluate the performance of the different prediction models on the test dataset. In order to evaluate the developed predictive models, we employ two strategies, namely the percentage of correctly classified instances within the test dataset and the receiver operating characteristic (ROC), or ROC curve. ROC Curve is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. The curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings.

Model accuracy is measured by the area under the ROC curve. An area of 1 represents a perfect test; an area of 0.5 represents a worthless test. Table 6 shows the percentage of correctly classified instances using each of the predictive models and Figure 5 shows the performance of the predictive models in terms of the area under the ROC curve and other useful model performance indicators. The results of applying the evaluation strategies on the produced models indicate that the RF classifier outperformed other classifiers in producing the best prediction model for the DNS availability, while LWL is the best for stability. PART outperformed other classifiers in predicting the quality attributes of security and resiliency.

Figure 5: Quality Attributes Prediction Models and Their Performance Indicators.

As it can be seen from Figure 5, the Mean Absolute Error (MAE) is in the worst case less than 0.4 out of 5. We can find an upper and lower bound on the accuracy of the predictive models. Since the values of the quality attributes to be predicted are natural numbers from 1 to 5, the error of around 0.4 can either be rounded up to 1 for the worst case, or considered as is for the best case. If we consider the worst case, the accuracy of the predictive models will be $\frac{(5-1)}{5} = 80\%$; however, for the best case, this is equivalent to $\frac{(5-0.4)}{5} = 92.5\%$ accuracy for the predictive model. Even for the worst case, the accuracy rate of the predictive models is quite high and supports our hypothesis that acceptable predictive models can be built from structural metrics of the DNS model in order to predict the DNS quality attributes of availability, security, stability and resiliency.

6.7 Threats to Validity

Empirical evaluation is always subject to different threats that can influence the validity of the results. We will specifically refer to the aspects of our experiment that may have been affected by these threats.

6.7.1 Conclusion Validity

In our experiment, a limited number of data points were collected due to the limited number of participants amongst the DNS operators. In addition, there were almost no models at our disposal and we have to build customized models using our *DGBuilder* tool. These limitations may pose threats to the drawn conclusions.

6.7.2 Construct Validity

The dependent variables which are the four quality attributes of the DNS model were measured using the subjective opinion of the participants. The threat posed by using subjective measurement mechanisms is that different participants may have different attitudes toward the evaluation of these attributes. In general, the participants of this experiment have a considerable number of years of experience within the DNS administration and their subjective views does capture what we claim to measure. It should also be noted that the used set of metrics may not be comprehensive and other consecutive research could further complete this proposed set by defining new metrics from other perspectives.

6.7.3 Internal Validity

Each of the dependency models represented different DNS system configuration and deployment structure. However, the models were simple enough to be understandable by the participants and they were given enough time (2 weeks) to become familiar with the concepts, structure and components of each model. The use of the 5-point Likert scale could have impacted the internal validity of the experiment due to the discrete nature of this ordinal scale in capturing the participants' views.

6.7.4 External Validity

The following two issues were considered for external validity: (1) the models used in the experiment are representative of wide range of real-world operational configurations and deployment choices. (2) We needed participants with high level of industrial experience to be able to complete the experiment and the target group of TLD operators did the job perfectly. Another threat to validity may be related to the tools that were used; however, since the tools were used to build the models and extract the metrics; we believe that it possibly affected all of the model measurements in the same way.

7. RELATED WORK

The DNS is a complex distributed system, a system of systems composed of a highly interconnected infrastructure, protocols and operations procedures. DNS name dependencies are analysed in [4], [7] and [16], in which the potential for a large number and variety of servers affecting name resolution is demonstrated. Individual operators and independent researchers have measured

various aspects of the DNS and from various prospective such as from the user, resolver or network points of view [18, 19]. In [16] Deccio et al. present a model of DNS name resolution from which the availability of a domain name can be quantified in the context of its deployment. Shulman et al. [2] studied the operational characteristics of the DNS infrastructure and how some factors impact resilience, stability and security of the DNS services. Casalicchio et al. [19] proposed a framework for the evaluation of the health and security levels of operational DNS. To the extent of our knowledge, only very few preliminary studies for defining suitable metrics to measure the quality attributes of the DNS system have been conducted [1, 7,16, 19]. Even within these existing works, not much theoretical or empirical evaluation of the proposed metrics has been done.

8. CONCLUSIONS AND FUTURE WORK

Overall, the results of the study show that we can reasonably claim that our objectives have been accomplished. The main implication of this is that structural metrics can indeed be used as early indicators of some of the external quality attributes of the DNS system. Prediction models based on the proposed structural metrics can be effectively built and utilized to predict the quality attributes of the system with good performance indicators. This work is part of a wider research project that aims at building a Quality Assurance Framework for the DNS. The DNS model developed as part of this research was limited to the static structure (design and deployment). We plan to grow our research by introducing some additional elements to the DNS model to represent the dynamic behavior of the system.

9. REFERENCES

[1] D. Wessels, M. Fomenkov, N. Brownlee and k. claffy, "Measurements and Laboratory Simulations of the Upper DNS Hierarchy," in *Passive and Active Network Measurement*, vol. 3015, C. Barakat and I. Pratt, Eds., Springer Berlin Heidelberg, 2004, pp. 147-157.

[2] H. Shulman and M. Waidner, "Towards Security of Internet Naming Infrastructure," in *European Symposium on Research in Computer Security*, 2015.

[3] E. Robert, B. Randy, B. Scott and P. Michael, "RFC 2182: Selection and Operation of Secondary DNS Servers," *International Engineering Task Force, Status: Standard,* 1997.

[4] V. Ramasubramanian and E. G. Sirer, "Perils of Transitive Trust in the Domain Name System," in *Proceedings of the 5th Conference on Internet Measurement 2005, Berkeley, California, USA, October 19-21, 2005*, 2005.

[5] M. Radwan and R. Heckel, "Detecting and Refactoring Operational Smells within the Domain Name System," *arXiv preprint arXiv:1504.02615,* 2015.

[6] V. Pappas, D. Wessels, D. Massey, S. Lu, A. Terzis and L. Zhang, "Impact of configuration errors on DNS robustness," *Selected Areas in Communications, IEEE Journal on,* vol. 27, pp. 275-290, April 2009.

[7] E. Osterweil, D. McPherson and L. Zhang, "Operational implications of the DNS control plane," *IEEE Reliability Society Newsletter,* 2011.

[8] P. Mockapetris, "RFC 1035: Domain names implementation and specification," *Work in Progress,* 1987.

[9] P. Mockapetris, "RFC 1034: Domain names: concepts and facilities," *Work in Progress,* 1987.

[10] K. Lu, K. Dong, C. Wang and H. Xu, "DNS configuration detection model," in *Systems and Informatics (ICSAI), 2014 2nd International Conference on*, 2014.

[11] M. Lotter, "Rfc 1033: Domain Administrators Operations Guide," *Work in Progress,* 1987.

[12] A. J. Kalafut, C. A. Shue and M. Gupta, "Understanding Implications of DNS Zone Provisioning," in *Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement*, New York, NY, USA, 2008.

[13] A. Herzberg and H. Shulman, "DNSSEC: Security and availability challenges," in *Communications and Network Security (CNS), 2013 IEEE Conference on*, 2013.

[14] J. Han, J. Pei and M. Kamber, Data mining: concepts and techniques, Elsevier, 2011.

[15] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. H. Witten, "The WEKA data mining software: an update," *ACM SIGKDD explorations newsletter,* vol. 11, pp. 10-18, 2009.

[16] C. Deccio, J. Sedayao, K. Kant and P. Mohapatra, "Measuring Availability in the Domain Name System," in *INFOCOM, 2010 Proceedings IEEE*, 2010.

[17] M. Corporation, "Microsoft Responds to DNS Issues," Microsoft Corporation, 2001.

[18] R. Chandramouli and S. Rose, "An integrity verification scheme for DNS zone file based on security impact analysis," in *Computer Security Applications Conference, 21st Annual*, 2005.

[19] E. Casalicchio, M. Caselli, A. Coletta, S. Di Blasi and I. Fovino, "Measuring Name System Health," in *Critical Infrastructure Protection VI*, vol. 390, J. Butts and S. Shenoi, Eds., Springer Berlin Heidelberg, 2012, pp. 155-169.

[20] D. Barr, "RFC 1912: Common DNS operational and configuration errors," *International Engineering Task Force, Status: Standard,* 1996.

[21] E.-H. Alikacem and H. A. Sahraoui, "A Metric Extraction Framework Based on a High-Level Description Language," in *Source Code Analysis and Manipulation, 2009. SCAM '09. Ninth IEEE International Working Conference on*, 2009.

[22] Working Group 4, "Report: DNS Best Practices," The Communications Security, Reliability and Interoperability Council III, 2012.

[23] P. E. Shrout and J. L. Fleiss, "Intraclass Correlations: Uses in Assessing Rater Reliability," *Psychological Bulletin,* vol. 86, no. 2, p. 420–428, 1979.