IEEE Open Journal of the
Computer Society

Received 17 March 2020; accepted 6 May 2020. Date of publication 11 May 2020;
date of current version 18 June 2020. The review of this paper was arranged by Associate Editor Deze Zeng.

Digital Object Identifier 10.1109/OJCS.2020.2994095

Analyzing CSP Trustworthiness and Predicting Cloud Service Performance

ROBERT MAESER (Senior Member, IEEE)

Dell Technologies, ISG office of the CTO

CORRESPONDING AUTHOR: ROBERT MAESER (e-mail: bob_maeser@dell.com)

ABSTRACT Analytics firm Cyence estimated Amazon's four-hour cloud computing outage in 2017 "cost S&P 500 companies at least $150 million" and traffic monitoring firm Apica claimed "54 of the top 100 online retailers saw site performance slump by at least 20 percent". According to Ponemon, 2015 data center outages cost Fortune 1000 companies between $1.25 and $2.5 billion. Despite potential risks, the cloud computing industry continues to grow. For example, Internet of Things, which is projected to grow 266% between 2013 and 2020, will drive increased demand on cloud computing as data across multiple industries is collected and sent back to cloud data centers for processing. RightScale estimates enterprises will continue to increase cloud demand with 85% having multi-cloud strategies. This growth and dependency will influence risk exposure and potential for impact (e.g. availability, performance, security, financial). The research in this paper and proposed solution calculates cloud service provider (CSP) trustworthiness levels and predicts cloud service and cloud service level agreement (SLA) availability performance. Evolving industry standards (e.g. NIST, ISO/IEC) for cloud SLAs and existing work regarding CSP trustworthiness will be leveraged as regression-based predictive models are constructed to analyze CSP cloud computing services, SLA performance and CSP trustworthiness.

INDEX TERMS CSP, cloud service provider, cloud service, cloud security, machine learning, predictive model, SLA, service level agreement, trust model, trustworthiness.

I. INTRODUCTION

A 2017 cloud survey from Skyhigh Networks and Cloud Security Alliance [41] identified the following key drivers for moving applications to cloud infrastructure (i.e. Infrastructure-as-a-Service providing subscription based processors, storage, network, software): increased security of cloud platforms, scalability based on workload, preference for operating expense vs capital, and lower costs. Despite these drivers, key challenges exist. A 2017 cloud survey from Rightscale [32] identified challenges that included complexity and lack of expertise, security, ability to manage cloud spend, governance and ability to manage multiple cloud services. Clearly, there is overlap between the drivers and challenges, with lots of potential for financial risk and missed service level expectations.

A. PROBLEM DESCRIPTION

As of January 2017, 95% of companies [32] are dependent on Infrastructure-as-a-Service (IaaS) CSPs, and as cloud computing usage increases (400% 2013-2020) [25], impact to service levels and financial risk from data center outages increase (81% 2010-2016) [30]. For example, the cost of data center outages in 2015 for Fortune 1000 companies was between $1.25 and $2.5 billion [30]. Of the data center outages between 2010 and 2016, 22% were caused by security [30]. Furthermore, data breaches were up 40% in 2016 [15]. With the increased demand and dependency on cloud computing, risk to expected service levels (e.g. availability, reliability, performance, security), potential for financial impact, and level of trustworthiness of CSPs are all important areas requiring attention. For this research paper and based on previous work reviewed in Section II (Related Work), CSP trustworthiness is analyzed in terms of the following: For historical QoS, did the CSP deliver the quality of service (QoS) they said they would deliver? Did they meet cloud service customer (CSC) expectations? How comprehensive and transparent are the CSP's delivery and security capabilities (e.g. what is their level of Cloud Security Alliance (CSA) Cloud Controls Matrix (CCM) compliance)? With respect to future performance, does the CSP have the required capabilities to meet future

CSC's service level requirements as represented by cloud SLAs?

B. SOLUTION APPROACH

Due to the described cloud computing challenges, risks and impacts, CSCs may be unable to trust CSPs. Trust needs to be based on more than CSP claims. With security being one of the most important factors that can influence trust between the CSPs and CSCs, the CSA designed the CCM [5] and Consensus Assessments Initiative Questionnaire (CAIQ) [4] to assist CSCs when assessing overall security capabilities of the CSP. While security remains a priority and rightly so, measuring and establishing CSP trust should consider criteria based on a comprehensive quantitative and qualitative assessment of CSP capabilities, CSC service level requirements, and CSP cloud service historical performance. An appropriate methodology and model is required to assess the CSP with respect to the broader set of trust criteria, including evaluating cloud computing SLAs between the CSC and CSP, and predicting cloud service and SLA performance (e.g. cloud service Availability).

To address the need, cloud computing service SLAs and a model for assessing, predicting and governing performance of cloud services and service levels are required to mitigate cloud computing service level risks and impact (e.g. availability, reliability, performance, security, financial). Research objectives related to this proposed solution are focused on analyzing CSPs and cloud computing services, and predicting cloud SLA performance. Industry standards and evolving research related to cloud computing SLAs [52], [14], [16], [29], [13] and CSP trustworthiness [11], [49] will be leveraged. A predictive model (using Linear Regression Analysis) was built for calculating SLA performance based on industry standardized cloud SLAs, CSP cloud service performance, CSP trustworthiness and other cloud computing characteristics. With the focus on cloud computing service levels, the related work was organized around the lifecycle of cloud computing SLAs. The lifecycle is comprised of four phases which are reviewed in Section II (Related Work): 1. Cloud Computing SLA specification, 2. CSP Trust Models, 3. Cloud Computing SLA monitoring, and 4. Cloud Computing SLA enforcement.

The outcome from Section III (Methods and Procedures) is the creation of a model that predicts cloud computing SLA availability. The methodologies used include Graph Theory to analyze and model cloud SLAs, Multi-Criteria Decision Analysis/Analytic Hierarchy Process (MCDA/AHP) to calculate CSP trustworthiness, and Linear Regression Analysis to build the predictive model based on multiple variables, including output from the other methodologies.

To assess the research contributions, the following hypothesis was proposed and tested based on the composite outcomes from all methodologies: *Predicting SLA-based cloud service availability has greater accuracy when calculated with more criteria than just historical cloud service downtime (e.g. CSP trustworthiness; global cloud service locations, IT resource capacity and performance).*

II. RELATED WORK

Over the past two decades in the Information Technology (IT) industry, a significant amount of important work occurred related to managing service levels and service level requirements, e.g. IT Service Management Forum Information Technology Infrastructure Library [14]. Related to this work, SLAs have been established as written agreements between service providers and IT customers. SLAs define key service targets and responsibilities, and serve as a basis for managing the relationship and establishing trust with the service provider. Within the cloud computing community, service level management and SLAs are also increasingly being adopted. Over the past few years, the European Commission has established the industry group (Cloud Select Industry Group – Subgroup on SLA: C-SIG-SLA), to provide a set of SLA standardization guidelines [52] for CSPs and CSCs. There is also work at the international level with ISO/IEC 19086 [16] from the ISO Cloud Computing Working Group. With the momentum and need around managing cloud computing service levels, the related work is organized around four phases of the lifecycle of cloud computing SLAs.

The first phase in the lifecycle focuses on Cloud Computing SLA specifications, their importance, benefits, standards and frameworks. A significant amount of work has been done by standards organizations (e.g. ITSMF ITIL, ISO/IEC, NIST, EC) [14], [16], [29], [52] related to defining standardized cloud SLAs. However, the work is not complete. CSP SLAs are diverse and gaps include adoption and standard methods for verifying compliance along with linking SLA compliance to predictive and proactive automation.

The next phase of the lifecycle is Cloud Computing trust models based on CSP capabilities (e.g. security), transparency and trustworthiness. Over the past few years, CSP trust models and CSP trustworthiness specifically related to cloud security have received significant attention. However, gaps exist with respect to considering both historical service level performance and other service level capabilities (vs just security) that can influence future performance. Questions to address when assessing CSP trustworthiness for historical performance include: historically, did the CSP deliver what they said they would deliver …did they meet service level expectations; and regarding future performance, does the CSP have the required capability to deliver against future service level expectations (not just security but all aspects of the SLA)? Existing trust model and trustworthiness research related to CSP assessment, evaluation and selection is organized into the following three groups.

- CSP evaluation and selection with fuzzy logic based trust models [51], [26], [31], [48].
- CSP assessment and trustworthiness based on security capabilities and CSA CCM [20], [23], [49], [24], [1], [34].
- CSP selection and trustworthiness based on SLAs and QoS requirements [27], [11], [33], [2], [35].

The third phase of the lifecycle is Cloud Computing monitoring and CSP trust levels based on QoS and SLAs. Over the last seven years, a significant amount of work has occurred (EC, IEEE, DMTF) [52], [28], [11], [49], [9], [8] related to low level instrumentation and monitoring of cloud services, correlation of low-level events, and measuring quality of service. CSP trust levels have received significant attention with respect to security. However, when looking at CSP trust levels in terms of the complete SLA, gaps exist. These gaps extend into monitoring in terms of a service level context that associates cloud service QoS, CSP trustworthiness and SLA compliance. Gaps also include predictive modeling of service level risks and SLA compliance in terms of taking into account both cloud service level performance and CSP trustworthiness.

The last phase of the lifecycle is Cloud Computing SLA enforcement through proactive QoS management and autonomic computing. There has been valuable work (IEEE, DMTF) [22], [53], [8], [54] related to cloud orchestration, containers, and autonomic self-managing clouds. The gap relates to bringing focus on SLAs and service level risk with governance and automation to proactively drive SLA compliance based on predictive models. To be more specific: based on a predicted probability of service level risk and SLA compliance, what cloud service and property changes could be made to automatically and proactively mitigate the service level risk and ensure SLA compliance.

III. METHODS AND PROCEDURES

The overall goal of the research was to utilize multiple methodologies to analyze, model and predict cloud industry SLA performance and validate the explanatory power and influence of multiple factors (e.g. cloud SLA standards and security, CSP trustworthiness, historical CSP SLA performance, cloud service performance and characteristics).

The methodologies utilized by the research include Graph Theory, MCDA/AHP and Linear Regression Analysis. The composite outcomes and models from the methodologies served to support the hypothesis testing and results. To validate the models, multiple Goodness of Fit measures were evaluated in addition to assessing the value of the models at addressing the hypothesis. The data used by the models is based on actual CSP historical SLA and cloud service performance and configurations from Amazon, Google and Microsoft, in addition to their self-assessments concerning security capabilities from the CSA Security, Trust and Assurance Registry (STAR) and CAIQ [6].

A. GRAPH THEORY TO MODEL CSPS, SLAS, SECURITY CONTROLS

A graph is an efficient way to describe a network structure and represent information about relationships between nodes (e.g. organizations, data structures related to SLAs and security controls). Graph Theory was used for this research to model cloud SLA work driven by industry standards organizations (EC, ENISA, ISO/IEC) [52], [13], [16], its relationship with cloud SLAs provided by the top three CSPs (Amazon, Google, Microsoft) [42], [43], [44], and its relationship with cloud security controls from CSA [5]. The primary goal was to define and analyze the structure that would be input to AHP for calculating CSP trustworthiness.

TABLE 1. Graph Theory Questions

What cloud SLA standards are related to what industry standards organizations and CSPs?
What SLAs exist for each CSP?
What SLAs have interrelationships across CSPs and what are the SLA gaps and overlaps between the CSPs?
What are the relationships between the CSP SLAs (Amazon, Google, Microsoft), industry standards organizations (ISO/IEC, EC, ENISA) and industry cloud SLA standards (i.e. content areas)?
Knowing the CCM control domains have some overlap with the SLA standards, what is the relationship between CSA CCM and cloud SLA standards … where are the gaps and how strong are the overlaps?
What relationships exist between the CSA CCM and CSP's CAIQ (Amazon, Google, Microsoft)?
How strong are CSP's CCM capabilities across each control domain?

The CSA CCM [5] represents 16 governing and operating domains separated into control domains and across the domains, 133 controls and 296 control related questions. The CAIQ [4] provides a survey for CSPs to assess and communicate their capabilities across all domains. The CSP CAIQ survey results are maintained in the CSA STAR [45], [46], [47]. Standards bodies such as EC, ENISA and ISO/IEC are working on SLA standards [52], [13], [16] for the cloud industry. The SLA standards for EC, ENISA and ISO/IEC represent a combined set of 12 content areas and 82 components. Table 1 presents the questions considered when collecting the data and building the graphs.

Once constructed, each graph was assessed based on the following factors [40], [7]:

- Degree centrality - relevance of SLA content area or security control domain to an organization.
- Betweeness centrality - quantifies SLA content area or security control domain strength over an organization.

B. AHP TO MODEL CSP TRUSTWORTHINESS

CSCs want to confirm claims from CSPs regarding the quality and security of their cloud services. Calculating and ranking the trustworthiness of CSPs provides a means for CSCs to rank CSPs and validate their ability to meet service level expectations [52], [13], [16]. AHP was used to model and assess CSP trustworthiness based on CSP SLA and security related capabilities, and CSP historical SLA and cloud service performance. AHP provided an effective methodology for addressing MCDA problems that involve the assessment of multiple quantitative and qualitative criteria using pairwise comparisons organized into a hierarchical structure with weighting and aggregation of scores, and ranking resulting from the comparisons [39], [36], [38], [37]. Related AHP, work focused on security controls [48], [49], [23], [24], was extended to include SLA capabilities, and historical SLA and cloud service performance (e.g. availability, latency and throughput; utilization for CPU, storage, memory and network; provisioning time). With AHP, CSP and cloud service criteria were quantified, weighted and assessed for each CSP. The results

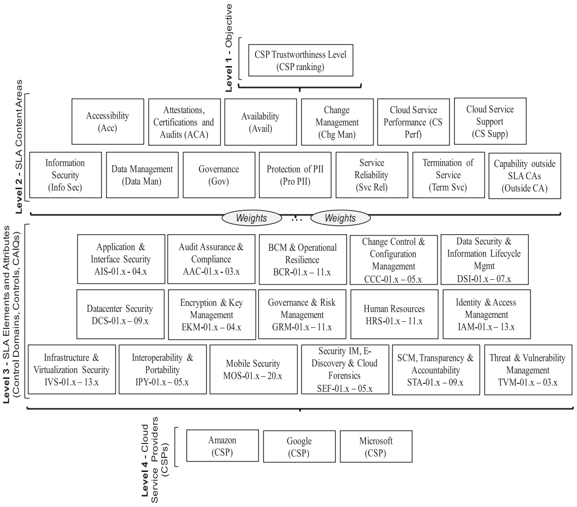

Level 1 - Objective
CSP Trustworthiness Level (CSP ranking)
Level 2 - SLA Content Areas
Accessibility (Acc)
Attestations, Certifications and Audits (ACA)
Availability (Avail)
Change Management (Chg Man)
Cloud Service Performance (CS Perf)
Cloud Service Support (CS Supp)
Information Security (Info Sec)
Data Management (Data Man)
Governance (Gov)
Protection of PII (Pro PII)
Service Reliability (Svc Rel)
Termination of Service (Term Svc)
Capability outside SLA CAs (Outside CA)
Weights
Weights
Level 3 - SLA Elements and Attributes (Control Domains, Controls, CAIQs)
Application & Interface Security AIS-01.x - 04.x
Audit Assurance & Compliance AAC-01.x - 03.x
BCM & Operational Resilience BCR-01.x – 11.x
Change Control & Configuration Management CCC-01.x – 05.x
Data Security & Information Lifecycle Mgmt DSI-01.x – 07.x
Datacenter Security DCS-01.x – 09.x
Encryption & Key Management EKM-01.x – 04.x
Governance & Risk Management GRM-01.x – 11.x
Human Resources HRS-01.x – 11.x
Identity & Access Management IAM-01.x – 13.x
Infrastructure & Virtualization Security IVS-01.x – 13.x
Interoperability & Portability IPY-01.x – 05.x
Mobile Security MOS-01.x – 20.x
Security IM, E-Discovery & Cloud Forensics SEF-01.x – 05.x
SCM, Transparency & Accountability STA-01.x – 09.x
Threat & Vulnerability Management TVM-01.x – 03.x
Level 4 - Cloud Service Providers (CSPs)
Amazon (CSP)
Google (CSP)
Microsoft (CSP)

FIGURE 1. AHP hierarchy structure.

of the comparisons were aggregated and CSPs were ranked. CSP trustworthiness was calculated based on the ranking. The AHP based approach for calculating CSP trustworthiness is organized into the following five steps.

In Step 1, the cloud SLA structure (i.e. Content Areas) was defined based on related industry work from ISO/IEC, EC, and ENISA [52], [13], [16]. The CSA CCM framework [5] was then mapped to the cloud SLA hierarchy. Graph Theory analysis of the cloud SLA specifications and CCM framework established the structure and mapping that was used for the AHP hierarchy.

Figure 1 displays the AHP hierarchy structure based on the following levels.

- Level 1 - Objective: Establish CSP trustworthiness based on the overall AHP-based CSP ranking.
- Level 2 - SLA Content Areas: Identify the high-level criteria, i.e. Content Areas (e.g. Accessibility), of the SLA structure. The Content Areas represent the hierarchies of SLA elements and attributes.
- Level 3 - SLA Elements and Attributes (Control Domains, Controls, CAIQs): Define the CSA CCM Control Domains (e.g. Application & Interface Security) and related hierarchy of Controls and CAIQs (e.g. AIS-01.x-04.x) within a Content Area.
- Level 4 – CSPs: Identify the CSPs whose capabilities are being assessed, compared and ranked with respect to trustworthiness.

Step 2 focuses on mapping evaluation criteria (CSP CAIQ, cloud performance) to the SLA structure defined in Step 1.

The CSP CAIQ evaluations are scored (quantified and normalized) based on the degree of CCM compliance for the CSP (i.e. number of CAIQ questions the CSP answered with a yes). Leveraging the Graph Theory analysis, Cloud SLA Content Areas are mapped to CCM Control Domains, associated controls and CAIQ compliance (i.e. number and percentage of questions answered with yes) for each CSP [4], [45], [46], [47]. The resultant level of CAIQ compliance for each control domain and control as related to the SLA Content Areas serves as input for the Comparison Matrices. The weighting of the AHP criterion is based on the quantity of CCM coverage for each SLA Content Area (i.e. number of CAIQs, Controls and Control Domains as ratios of the totals associated with each SLA Content Area). For SLA Content Area Availability, CSP performance measurements are provided by Gartner's Technology Planner Cloud Module [10]. The measurements span 2015 to 2017 and are normalized and weighted based on the number of service regions in operation during each measurement period. Similar to CCM compliance for each CSP (via CAIQ), the availability performance for each CSP is a proportion (based on measurements and number of service regions) associated with the SLA Content Area (i.e. AHP evaluation criterion) Availability.

TABLE 2. Comparison Matrices and Priority Vector Terms

Term	Definition
ca	SLA Content Area.
C_i	CSP i where i=a for Amazon, i=g for Google, i=m for Microsoft.
W_{ca}	Weight factor for SLA Content Area criteria (ca).
$V_{i,ca}$	Value of content area (ca) capability ratio score for CSP (C_i).
$R_{i/j,ca}$	Relative rank ratio for pairwise comparison $V_{i,ca}$ / $V_{j,ca}$ for two CSPs (C_i and C_j).
C_i / C_j	Relative rank ($R_{i/j,ca}$) of C_i over C_j regarding ca.
PV_{ca}	Priority vector entry for SLA Content Area (ca).
$PV\,C_i$	Priority vector for CSP (C_i).
$WPV\,C_i$	Weighted priority vector for CSP (C_i).
$TL\ \ C_i$	Trustworthiness level for CSP (C_i).

Step 3 focused on pairwise comparisons and building AHP comparison matrices. Two types of comparison matrices (using terms in Table 2) were constructed:

- Comparison Matrix for SLA Content Area capability criteria: represents pairwise comparisons of AHP evaluation criteria utilizing SLA Content Area capability ratios from Step 2.
- CSP Comparison Matrices for SLA Content Areas (one per SLA Content Area): represent pairwise comparisons utilizing CSP capability ratios for each SLA Content Area from Step 2.

When performing the pairwise comparisons in each SLA Content Area comparison matrix, the relationship of CSP capability ratio scores between two CSPs was represented as: $V_{i,ca}$ / $V_{j,ca}$ (for CSPs C_i and C_j). The result of the pairwise comparison calculation was the relative rank ratio ($R_{i/j,ca}$), which indicated the performance of C_i compared to C_j, for the specified SLA Content Area *(ca)*. For each SLA Content Area, the comparison matrix (size 3 x 3 based on a pairwise comparisons of Amazon, Google, Microsoft) would be constructed as shown in equation 1:

$$CM_{ca} = \begin{array}{c} \\ C_a \\ C_g \\ C_m \end{array} \left| \begin{array}{ccc} C_a & C_g & C_m \\ (C_a/C_a) & (C_a/C_g) & (C_a/C_m) \\ (C_g/C_a) & (C_g/C_g) & (C_g/C_m) \\ (C_m/C_a) & (C_m/C_g) & (C_m/C_m) \end{array} \right| \qquad (1)$$

Step 4 creates the normalized, reciprocal matrices and priority vectors (i.e. normalized principle Eigenvector) for the AHP evaluation criteria (i.e. SLA Content Area capability) and for the CSP capability per each SLA Content Area. Two types of priority vectors (PVs) are constructed.

- Priority Vector for SLA Content Area capability: per comparison matrix calculated in Step 3 - to calculate the priorities (weights), a pairwise comparison is performed among all SLA Content Areas using their capability ratios as presented in Section IV.B.1; this PV represents weights (Table 4) for each SLA Content Area; the weights help to calculate overall CSP priority vectors.
- CSP Priority Vectors for each SLA Content Area: per comparison matrices for each SLA Content Area calculated in Step 3 - to calculate the local CSP priorities, pairwise comparisons are performed using the CSP capability ratios of each SLA Content Area presented in Section IV.B.1; for each SLA Content Area, a comparison matrix and priority vector is constructed to present the corresponding CSP pairwise comparisons and priorities; these PVs are used as input to Step 5 for performing CSP AHP based estimation to calculate the overall CSP priority vectors.

Utilizing Equation 2, shown at the bottom of this page, (Normalized, reciprocal SLA Content Area Matrix and Priority Vector), the normalized, reciprocal matrices and priority vectors were built using the SLA Content Area Comparison Matrix from Equation 1, then columns were summed and each element ($C_i\,/\,C_j$) was divided by the sum of its column ($R_{i/j,ca}/\sum_{i=a,g,m} R_{i/j,ca}$) to normalize the relative weight of each element.

Once each matrix element is normalized, the priority vectors (i.e. normalized principle Eigenvectors) were created by averaging each row of the normalized, reciprocal matrix for each SLA Content Area ($\sum_{j=a,g,m}\ R_{i/j,ca}/3$). The sum of all elements in the priority vector for each SLA Content Area is one. The priority vector (PV) in Equation 2 represents relative weights related to each CSP for each SLA Content Area. The priority vector represents numerical values that specify the CSP order of preference for the SLA Content Area (i.e. a representation of which CSP is more capable for each SLA Content Area).

The final phase (Step 5) calculates the AHP based estimation by building the CSP priority vectors based on the Step 4 SLA Content Area Priority Vectors. The weighted factors (Table 4) calculated by Step 4 are then applied to the CSP priority vectors to build the weighted CSP priority vectors. Finally, the Trustworthiness Level for each CSP is calculated by aggregating the weighted CSP Priority Vectors. The following steps were taken when building the CSP Priority Vectors, applying the weighting and calculating CSP Trustworthiness Levels. The SLA Content Area priority vectors (PV_{ca}) introduced by Equation 2 are arranged by CSP to build CSP priority vectors as defined by Equation 3. Refer to Table 4 for *ca* abbreviations.

$$\begin{aligned}&PVC_i = (PV_{ca} \ldots PV_{ca})\\&(i = a, g, m)\\&(ca = Acc,\ ACA,\ Avail,\ Chg\ Man,\ CS\ Perf,\ CS\ Supp,\\&Data\ Man,\ Gov,\ Info\ Sec,\ Pr\ o\ PII,\ Svc\ Rel,\ Term\ Svc,\\&Outside\ CA)\end{aligned} \tag{3}$$

For each CSP Priority Vector ($PV\,C_i$) per Equation 3, each SLA Content Area priority vector entry (PV_{ca}) was multiplied by corresponding SLA content area criteria weight (Table 4). The result is weighted AHP based estimation for each CSP as represented in Equation 4, i.e. weighted priority vectors for each CSP. Refer to Table 4 for *ca* abbreviations.

$$\begin{aligned}&WPVC_i = (PV_{ca}xW_{ca} \ldots PV_{ca}xW_{ca})\\&(i = a, g, m)\\&(ca = Acc,\ ACA,\ Avail,\ Chg\ Man,\ CS\ Perf,\ CS\ Supp,\\&Data\ Man,\ Gov,\ Info\ Sec,\ Pr\ o\ PII,\ Svc\ Rel,\ Term\ Svc,\\&Outside\ CA)\end{aligned} \tag{4}$$

In Equation 4, the weight factor (W_{ca}) was multiplied by each SLA Content Area Priority Vector (PV_{ca}) entry. The calculation of the weight factors from Table 4 were based on the priority vector for SLA content area capability criteria. For each CSP (C_i where $i = a,g,m$), weighted priority vector entries ($PV_{ca}\,x\,W_{ca}$) for all SLA Content Areas (*ca*) from Equation 4 and $WPV\,C_i$, were aggregated to represent the final CSP Trustworthiness Level represented by Equation 5.

$$TLC_i = \sum (PV_{ca}\,xW_{ca}) \tag{5}$$

C. REGRESSION ANALYSIS TO MODEL, PREDICT SLA AVAILABILITY

Regression analysis was used to investigate functional relationships between CSP and cloud service variables and ultimately develop models to predict cloud service SLA availability. The predictive modeling utilized CSP trustworthiness

$$\begin{array}{l|ccc|ll} & C_a & C_g & C_m & \boldsymbol{PV_{ca}} & (j = a, g, m) \\ C_a & (C_a/C_a)\,/sum & (C_a/C_g)\,/sum & (C_a/C_m)\,/sum & \sum (C_a/C_j)\,/3 & \\ C_g & (C_g/C_a)\,/sum & (C_g/C_g)\,/sum & (C_g/C_m)\,/sum & \sum (C_g/C_j)\,/3 & (j = a, g, m) \\ C_m & (C_m/C_a)\,/sum & (C_m/C_g)\,/sum & (C_m/C_m)\,/sum & \sum (C_m/C_j)\,/3 & (j = a, g, m) \\ \text{sum} & \sum C_i/C_a & \sum C_i/C_g & \sum C_i/C_m & & \\ & (i = a, g, m) & (i = a, g, m) & (i = a, g, m) & & \end{array} \tag{2}$$

levels in addition to historical CSP cloud service performance data (e.g. downtime; networking latency and throughput; utilization for CPU, storage, memory and network; provisioning turnaround). Two regression models (simple and multiple variable) were created to relate predictor variables with the response variable and drive predictions that enabled analysis and testing of the stated hypothesis. The following activities guided the regression analysis [3], [21], [12].

- Define the problem statement (Hypothesis): Predicting SLA-based cloud service availability has greater accuracy when calculated with more criteria than just historical cloud service downtime (e.g. CSP trustworthiness; global cloud service locations, IT resource capacity and performance).
- Select variables that explain the response variable: Response variable = Y (Cloud service Availability); Predictor variables = X_1, X_2, X_3, X_4, X_5
    - ○ X_1 = Cloud service downtime
    - ○ X_2 = CSP trust level
    - ○ X_3 = Global service region locations
    - ○ X_4 = Number of cloud service regions
    - ○ X_5 = Cloud service performance
- Identify and collect required data: Data was provided by Gartner, Amazon, Google, Microsoft, CSA, ISO/IEC, EC, ENISA.
- Define linear regression model specification: The relationship between the response (Y) and predictor variables (X_1, X_2, X_3, X_4, X_5) can be approximated by the regression models of $Y = f(X_1) + E$ for simple regression and $Y = f(X_1, X_2, X_3, X_4, X_5) + E$ for multiple regression. "*E* is assumed to be a random error representing the discrepancy in the approximation" and accounts for the "failure of the model to fit the data exactly" [3], [21], [12].
- Select the fitting method: Least squares was used for estimating parameters of the model.
- Fit the model: The model was fit to the data (101 observations, 41 CSP global service regions) to estimate the regression coefficients.
- Validate and critique the model: Analyze linear regression assumptions for both models and related data, and validate the Goodness of Fit measures.
- Utilize the model for the identified problem.

1) SIMPLE REGRESSION MODEL

The simple linear regression model is denoted by the equation $Y = B_0 + B_1X + E$ where B_0 and B_1 represent the regression coefficients and E is the "random disturbance or error" [3], [21]. B_1 is the slope (i.e. changes in Availability for a unit of change in Downtime), and B_0 is the constant coefficient (i.e. intercept, predicted value of Y when $X = 0$). For the model, the goal is to estimate B_0 and B_1, and "find the straight line that provided the best fit" of the response versus predictor variable using least squares [3], [21].

2) MULTIPLE REGRESSION MODEL

The multiple linear regression model is denoted by the equation $Y = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + B_4X_4 + B_5X_5 + E$, where B_0, B_1, B_2, B_3, B_4, B_5 are the regression coefficients, and E is the "random disturbance or error" [3], [21]. Each regression coefficient (B_1, B_2, B_3, B_4, B_5) represents the slope (i.e. changes in Availability for a unit of change in a coefficient as other coefficients are held constant), and B_0 is the constant coefficient (i.e. intercept, predicted value of Y when $X = 0$). For the model, the goal was to estimate parameters $B_0, B_1, B_2, B_3, B_4, B_5$ and "find the straight line that provided the best fit" of the response versus predictor variables using least squares method [3], [21].

3) REGRESSION MODEL ANALYSIS AND TRANSFORMATION

Outliers and linear assumptions were assessed, predictor variable(s) were transformed, and the final models were again fitted and analyzed. A K-fold cross-validation method was applied to calculate measure of errors for polynomial fit with different degrees. The cross-validation analysis was leveraged to build the modified simple linear regression model for a better fit. Multiple methods were applied to select a subset of predictor variables for the multiple regression equation. The methods included stepwise selection (forward, backward, both using AIC), all-subset regression using leaps, and lastly best subset selection. From the model summary the following analysis occurred.

- Is p-value for the overall model < 0.05 (an overall test of the statistical significance for the model)?
- Does the model have a strong R-squared, indicating majority of variance related to the response variable is explained by predictor variable(s) vs "random error" [50]?
- How large was the residual standard error, which measures the average amount the response variable "deviates from the true regression line for any given point" (i.e. how far observed Y values are from the predicted values) [50], [3]?
- Are the p-values for the regression coefficients statistically significant (i.e. < 0.05)?
- After analysis of the model's results, ANOVA and linear regression assumptions were analyzed (Linearity, Normality, Independence, Homoscedasticity), driving any required variable transformation [50], [3].

4) GOODNESS OF FIT MEASURES AND MODEL VALIDATION

Goodness of Fit measures for both simple and multiple linear regression models were calculated and analyzed [50], [3], [12]. Predictions were performed with multiple prediction intervals between 50% and 95% using 5% increments. The total and average number of successful predictions that fell within the prediction intervals were assessed. ANOVA was used to compare the models and measure what the multiple regression added to the prediction vs the simple regression. Predicted vs observed availability was plotted with 95% confidence and prediction interval, and fit for both models was compared.

All data was provided directly from the CSPs and/or validated against published CSP data (e.g. CSP outage notifications and duration cross-checked against Gartner data).

5) HYPOTHESIS TESTING

Testing of the hypothesis assessed whether the additional regression coefficients of the multiple linear regression model provided statistically significant contribution to explaining cloud service availability vs the simple linear regression model with a single coefficient (i.e. did the additional variables of the multiple linear regression model provide explanatory or predictive power to the model)? The following null and alternative hypothesis were tested.

H_0 *: Multiple = Simple LR model, i.e. zero difference in accuracy* $(B_j = 0)$*; no additional variables add explanatory/predictive power to the Simple LR model.*

H_a *: At least one additional variable* (B_j) *has a statistically significant contribution to explaining/predicting* Y $(B_j\ != 0)$*, and therefore should remain in the model.*

Decision to reject or fail to reject the null hypothesis was based on the Goodness of Fit measures, including the ANOVA F-test results comparing the two models, and the validation results concerning successful predictions.

IV. RESULTS

A comprehensive study of the cloud SLA lifecycle was conducted, analyzing SLA and security specifications and frameworks, CSP trustworthiness, and historical cloud service performance (e.g. cloud IT resources, CSP SLAs). The analysis focused on the top three CSPs (Amazon, Google, Microsoft) and how the data influenced cloud service performance and predicting SLA Availability. Predicting SLA performance and hypothesis testing was based on the composite outcomes of the following methodologies. Graph Theory output fed AHP, whose output was input to Regression Analysis. The primary question: *Is there greater accuracy predicting cloud SLA performance with historical SLA and cloud service performance, and CSP trustworthiness, vs just historical cloud service downtime?*

A. GRAPH THEORY TO MODEL CSPS, SLAS, SECURITY CONTROLS

As outlined in Section III, graph theory was used to visualize and analyze gaps, overlaps and relationships between industry SLA standards [16], [29], [52], industry cloud security controls [5], and CSP cloud security controls and SLAs (e.g. Amazon, Google, Microsoft) [45], [46], [47], [42], [43], [44]. Table 1 details the questions that drove graph model development. As an example, Figure 2 depicts the graph for the first question "*What cloud SLA standards are related to what industry standards organizations and CSPs?*". Figure 2 models the contributed work related to cloud SLA standards from standards bodies (triangle node) ISO/IEC, ENISA and EC [16], [13], [52] in addition to how the work relates with SLAs from CSPs (circle node) Amazon, Google and Microsoft.

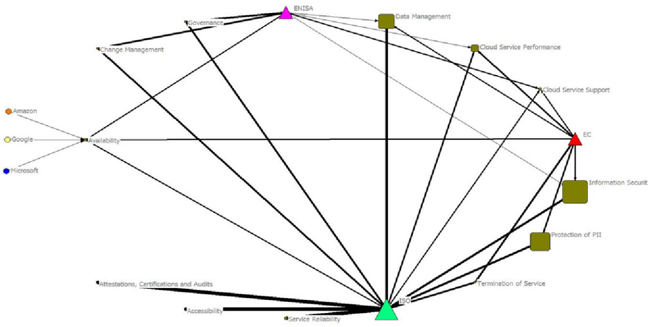


FIGURE 2. **Cloud SLA standards relationships.**

The standardized cloud SLA content has been organized into 12 content areas (square brown node) based on the SLA framework developed by ISO/IEC [17], [18], [19]. For each content area, the node size is based on the number of SLA details (i.e. components) specific to that content area in proportion to the combined total number of SLA details (i.e. components) for all content areas. The size of the node for each standards body is based on the number of relationships across all SLA content areas in proportion to the total number of available SLA content areas. The thickness (strength) of the relationship (tie) between a standards body (e.g. EC, ENISA, ISO/IEC) and a specific content area denotes the quantity of SLA details (i.e. components) the standards body contributed to the content area in proportion to the total number of SLA details that exist for that content area. Based on modeled nodes, ties and relationships, centrality factors were assessed for all graphs. As an example, the following observations were made for Figure 2 [7].

- 1 of 12 SLA content areas had relationship with 3 of 3 CSPs (Amazon, Google, Microsoft): Availability.
- 3 of 12 SLA content areas received most contributions (each > 20% of total components vs others <= 10%): Info Security, Protection of PII, Data Management.
- ISO/IEC contributed 12 of 12 SLA content areas; Primary for 9 of 12 (each contribution >= 50% of content areas' components): Accessibility; Attestations, Certifications and Audits; Change Management; Data Management; Governance; Info Security; Protection of PII; Service Reliability; Termination of Service.
- ENISA contributed 7 of 12 SLA content areas; Primary for 2 of 12 (each contribution >= 50% of the content areas' components): Change Management, Governance.
- EC contributed 7 of 12 SLA content areas; Primary for 1 of 12 (contribution >= 50% of the content areas' components): Termination of Service.

Based on the analysis and findings related to Graph Theory, Table 3 shows how the organizations (standards bodies CSA, EC, ENISA, ISO/IEC; and CSPs Amazon, Google, Microsoft) mapped to the cloud industry standards SLA content areas and compared with each other. Green indicates the organization contributes and aligns with 100% of the content area SLA capabilities. Yellow indicates partial contribution and alignment

TABLE 3. Cloud SLA Standard—Organization Coverage

Cloud SLA Standards	CSA	EC	ENISA	ISO	Amazon	Google	Microsoft
Accessibility	🟥	🟥	🟥	🟩	🟥	🟥	🟥
Attestations, Certifications and Audits	🟧	🟥	🟥	🟩	🟥	🟥	🟥
Availability	🟥	🟨	🟨	🟨	🟧	🟧	🟨
Change Management	🟧	🟥	🟨	🟨	🟥	🟥	🟥
Cloud Service Performance	🟧	🟧	🟧	🟧	🟥	🟥	🟥
Cloud Service Support	🟧	🟧	🟧	🟧	🟥	🟥	🟥
Data Management	🟧	🟧	🟧	🟨	🟥	🟥	🟥
Governance	🟧	🟥	🟨	🟨	🟥	🟥	🟥
Information Security	🟧	🟧	🟧	🟨	🟥	🟥	🟥
Protection of PII	🟧	🟧	🟥	🟨	🟥	🟥	🟥
Service Reliability	🟧	🟥	🟥	🟩	🟥	🟥	🟥
Termination of Service	🟥	🟨	🟥	🟨	🟥	🟥	🟥

TABLE 4. SLA Content Area Criteria Weight

Weight	SLA Content Area Criteria	
0.0655	Acc	Accessibility
0.0236	ACA	Audits
0.0982	Avail	Availability
0.0222	Chg Man	Change Management
0.0653	CS Perf	Cloud Service Performance
0.0264	CS Supp	Cloud Service Support
0.0348	Data Man	Data Management
0.0473	Gov	Governance
0.3198	Info Sec	Information Security
0.1029	Pro PII	Protection of PII
0.0487	Svc Rel	Service Reliability
0.0655	Term Svc	Termination of Service
0.0799	Outside CA	Capability outside SLA CAs

between 50-99%. Orange indicates partial contribution and alignment between 1-49%. Red indicates none of the content area SLA capabilities were addressed.

B. AHP TO MODEL CSP TRUSTWORTHINESS

The objective of using AHP is to analyze and model the trustworthiness of CSPs with respect to multiple factors (e.g. delivery and support capability, security, quality of service, performance). This requires a framework to identify and understand the problem, criteria and their relationships. AHP provides this framework, enabling the problem, CSP capability and performance, and CSC expectations to be structured into a hierarchical order along with judgements and numeric values that reflect their importance and impact. The values are synthesized to determine overall priorities and rankings which ultimately serve to represent the trustworthiness of the CSP. The principle dimensions (criteria) that were evaluated by this research relates to cloud SLA Content Areas based on the Graph Theory Analysis [52], [13], [16]. The content areas are depicted in Figure 1 AHP Hierarchy Structure. To calculate the SLA Content Area capability and CSP capability strength, proportions (ratios) were used based on absolute measurements according to published CSP assessments [6], [4], [45], [46], [47], cloud security capability framework standards and compliance assessments [5], [4], and CSP performance [10].

1) CAPABILITY RATIOS FOR SLA CONTENT AREAS AND CSPS

Capability ratios for SLA Content Areas and CSPs serve as the input based on absolute measurements for calculating a ratio based scale for building the comparison matrices, priority vectors, weighting (based on SLA Content Areas as evaluation criteria) and overall ranking of the AHP alternatives (i.e. CSPs). Level two of the AHP hierarchy consists of criteria (SLA Content Areas) that contribute to trustworthiness and are used to help assess each CSP. For each SLA Content Area, capability ratios are calculated based on the strength and importance value derived during the Graph Theory Analysis of each SLA Content Area. Level three of the AHP hierarchy consists of the factors (e.g. CCM Compliance for a CSP based on CAIQ assessments, CSP cloud service historical availability) related to each level two criterion. Similar to each criterion (i.e. SLA Content Area), a CSP capability ratio was calculated based on the CSPs performance for all factors organized by SLA Content Area.

With pairwise analysis of the CSP capability ratios organized by SLA Content Area, each CSP's relative performance and capability was ultimately prioritized for each corresponding SLA Content Area.

2) SLA CONTENT AREA CAPABILITY—PRIORITY VECTORS, MATRIX

When assessing SLA Content Area capability criteria, we need to identify their relative priorities since not all priorities are the same. The priority (weight) in Table 4 essentially reflects the importance/contribution/influence of a SLA Content Area with respect to the others.

3) CSP SLA CONTENT AREA—PRIORITY VECTORS, MATRIX

For each CSP (AHP alternative), the strength of their performance and capability with respect to each SLA Content Area (AHP criterion) is analyzed. The strength is represented by calculating the CSP relative priorities for each SLA Content Area. Since these priorities are specifically related to each criterion, they are local priorities vs the overall CSP priorities derived in Section IV.B.4 [39], [36], [38], [37]. One comparison matrix and related priority vector for each SLA Content Area is constructed. When performing the comparison, the calculation portrays which CSP has the stronger performance and capability related to the SLA Content Area of the matrix.

4) CSP PRIORITY VECTORS AND RELATIVE TRUSTWORTHINESS LEVEL

In the final AHP process step, the overall priority for each CSP (i.e. each alternative) was calculated. In the previous Section, the priorities were local and reflected the CSP priority with respect to each individual SLA Content Area (i.e. criterion). With the overall priority for each CSP, we take into account not only the CSP priorities for each criterion but also the weighting factor of each criterion (i.e. SLA Content Area). This ties all levels of the hierarchy together and represents "model synthesis" [39], [36], [38], [37]. In this step, calculating CSP trustworthiness performs AHP based estimation by building CSP priority vectors (Table 5) based on the CSP

TABLE 5. CSP Priority Vectors and Weighting Factors

	Acc	ACA	Avail	Chg Man	CS Perf	CS Supp	Data Man	Gov	Info Sec	Pro PII	Svc Rel	Term Svc	Outside CA
Amazon	0.3404	0.3421	0.3334	0.3448	0.3626	0.3514	0.3636	0.3333	0.3525	0.3485	0.3387	0.3404	0.2701
Google	0.3312	0.3158	0.3334	0.3103	0.3077	0.2973	0.3182	0.3333	0.3104	0.3258	0.3226	0.3312	0.4453
Microsoft	0.3284	0.3421	0.3332	0.3448	0.3297	0.3514	0.3182	0.3333	0.3370	0.3258	0.3387	0.3284	0.2847

TABLE 6. CSP Weighted Priority Vectors

	Acc	ACA	Avail	Chg Man	CS Perf	CS Supp	Data Man	Gov	Info Sec	Pro PII	Svc Rel	Term Svc	Outside CA
Amazon	0.0223	0.0081	0.0327	0.0077	0.0237	0.0093	0.0126	0.0158	0.1127	0.0359	0.0165	0.0223	0.0216
Google	0.0217	0.0075	0.0327	0.0069	0.0201	0.0079	0.0111	0.0158	0.0993	0.0335	0.0157	0.0217	0.0356
Microsoft	0.0215	0.0081	0.0327	0.0077	0.0215	0.0093	0.0111	0.0158	0.1078	0.0335	0.0165	0.0215	0.0228

TABLE 7. CSP Relative Trustworthiness Levels

Amazon	0.3411
Google	0.3293
Microsoft	0.3296

Priority Vectors for each SLA Content Area from Section IV.B.3.

The weight factors for each SLA Content Area Criteria (Table 5) are then applied to the CSP priority vectors (i.e. each SLA Content Area weight factor from Table 4 was multiplied by each priority in the CSP priority vector) to build the weighted CSP priority vectors (Table 6).

Finally, the comparative trustworthiness level for each CSP was calculated by aggregating the weighted CSP Priority Vectors. Table 7 shows the relative trustworthiness ranking of the CSPs. The higher the CSP ranking, the higher the level of CSP trustworthiness [39], [36], [38], [37]. As an example, given the importance (weight) of each SLA Content Area (i.e. evaluation criteria), Amazon is the most trusted of the three CSPs.

5) ABSOLUTE CSP TRUSTWORTHINESS LEVEL FOR REGRESSION

The CSP relative trustworthiness levels in Table 7 are effective for comparing the capabilities of two or more specific CSP's. The ultimate objective is to provide Regression Analysis with an overall absolute (vs relative) trustworthiness level for each CSP. This requires comparison matrices and priority vectors for each content area based on a CSP's capabilities vs the maximum potential capabilities. Table 8 presents an example for Amazon and the Accessibility SLA Content Area. The same process was then implemented as outlined for Tables 5, 6, 7. For the Amazon example, the resultant Priority Vector, Weighted Priority Vector (i.e. each SLA Content Area weight factor from Table 4 was multiplied by each priority in the Amazon priority vector) and Absolute Trustworthiness Level are presented. After aggregating Amazon's Weighted Priority Vector, the absolute CSP trustworthiness level was calculated. The absolute trustworthiness level is a ratio (between 0 and 1) of the CSP's capability vs the maximum potential and is leveraged as a predictor variable for Regression Analysis in Section IV.C. Table 8 depicts Amazon's absolute trustworthiness level of 92.13%. Google (88.6%) and Microsoft (89.13%) were calculated in the same manner.

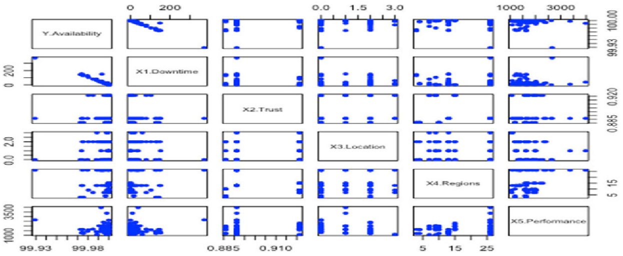


FIGURE 3. Matrix scatterplot of variables.

C. REGRESSION ANALYSIS TO MODEL, PREDICT SLA AVAILABILITY

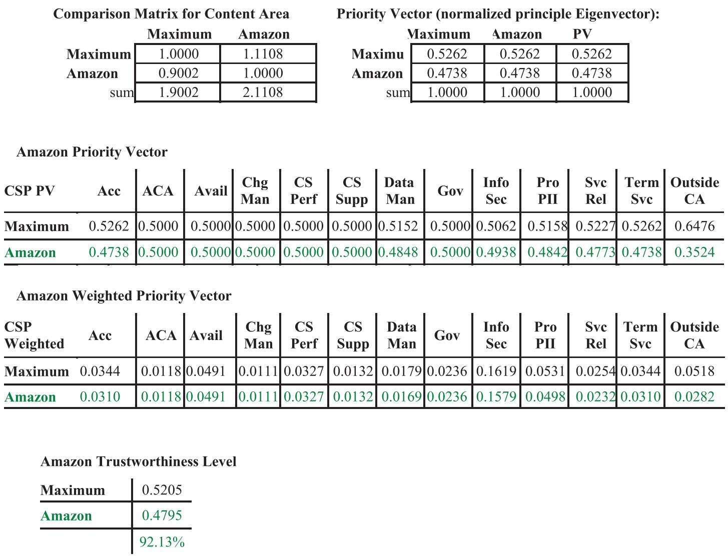
Two linear regression models (simple and multiple variable) were created to relate predictor variables with the response variable and drive predictions that enable analysis and testing of the stated hypothesis. The dataset used for the regression analysis represents 101 observations related to response variable Y (Cloud Service Availability) and 5 predictor variables: X_1 (Cloud service downtime), X_2 (CSP trust level), X_3 (Global service region locations), X_4 (Number of cloud service regions), X_5 (Cloud service performance). While the multiple regression leveraged all predictor variables, the simple regression focused on the single predictor X_1.

With the variable matrix scatterplot (Figure 3), patterns and relationships can be identified among the response and predictor variables (e.g. Y availability, X_4 cloud regions, X_5 performance). As the number of service regions increase so does performance and downtime (i.e. increases in service regions represent increases in consumption of compute services and potential for exceeding capacity which can effect downtime). For both regressions, the same outliers exist related to Y availability and X_1 downtime, so further analysis was required as the model was fit and linear regression assumptions were tested (e.g. linearity, normality, independence, homoscedasticity, and outliers).

Figure 4 provides visualizations with 95% confidence interval (grey) and 95% prediction interval (black dotted line) for both the simple and multiple regression models. The multiple regression provides a better fit which is also reinforced by the Goodness of Fit measures in Table 9.

ANOVA was used to measure what the multiple regression model added to the prediction over the simple regression. With the conventional test, regression sums of squares was compared for the two models (i.e. extra sum of squares test). The null hypothesis that "additional predictors all have zero coefficients" was tested using the F-statistics [50]. The "extra sum of squares" was the "amount by which the residual sum of squares [was] reduced by the additional predictors" [50]. From the test, we have Sum of Squares = 0.00059632, with a p-value = 2.2E-16 (< 0.05 so statistically significant). With

TABLE 8. Example Absolute CSP Trustworthiness Level



Comparison Matrix for Content Area

	Maximum	Amazon
Maximum	1.0000	1.1108
Amazon	0.9002	1.0000
sum	1.9002	2.1108

Priority Vector (normalized principle Eigenvector):

	Maximum	Amazon	PV
Maximu	0.5262	0.5262	0.5262
Amazon	0.4738	0.4738	0.4738
sum	1.0000	1.0000	1.0000

Amazon Priority Vector

CSP PV	Acc	ACA	Avail	Chg Man	CS Perf	CS Supp	Data Man	Gov	Info Sec	Pro PII	Svc Rel	Term Svc	Outside CA
Maximum	0.5262	0.5000	0.5000	0.5000	0.5000	0.5000	0.5152	0.5000	0.5062	0.5158	0.5227	0.5262	0.6476
Amazon	0.4738	0.5000	0.5000	0.5000	0.5000	0.5000	0.4848	0.5000	0.4938	0.4842	0.4773	0.4738	0.3524

Amazon Weighted Priority Vector

CSP Weighted	Acc	ACA	Avail	Chg Man	CS Perf	CS Supp	Data Man	Gov	Info Sec	Pro PII	Svc Rel	Term Svc	Outside CA
Maximum	0.0344	0.0118	0.0491	0.0111	0.0327	0.0132	0.0179	0.0236	0.1619	0.0531	0.0254	0.0344	0.0518
Amazon	0.0310	0.0118	0.0491	0.0111	0.0327	0.0132	0.0169	0.0236	0.1579	0.0498	0.0232	0.0310	0.0282

Amazon Trustworthiness Level

Maximum	0.5205
Amazon	0.4795
	92.13%

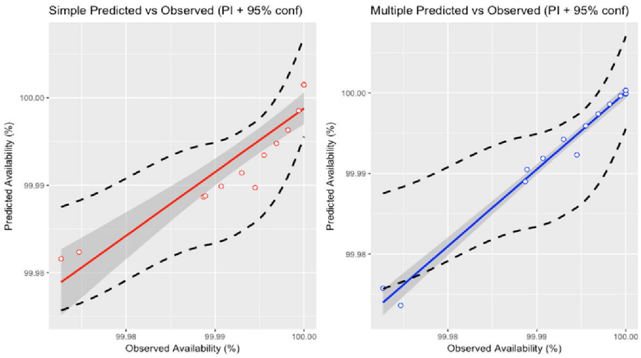
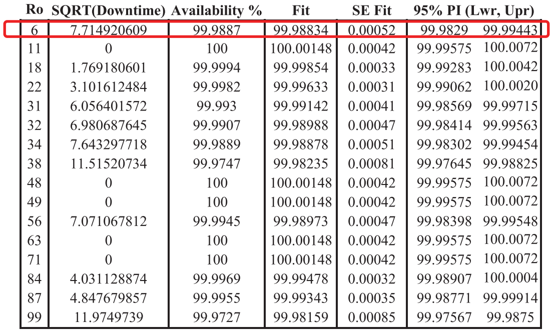


FIGURE 4. Predicted vs observed with 95% CI and PI.

TABLE 9. Goodness of Fit Measurements

Measurement	Simple	Multiple
Residual Standard Error	0.0029	0.0009
R.Squared	0.8113	0.9839
Adjusted.R.Squared	0.8089	0.9831
Predicted.R.Squared	0.7927	0.9806
Predictve Residual Sum of Squares (PRESS)	7.2290e-04	7.7646e-05
Akaike's An Information Criterion (AIC)	-733.29	-920.19
Bayesian Information Criterion (BIC)	-726.03	-905.68
Mallow's Cp	2	5
Root mean squared error (RMSE)	0.0034	0.0010
Mean Absolute Error (MAE)	0.0024	0.0007
Ratio of RMSE (RSR)	0.5031	0.1582

TABLE 10. Simple Regression Predictions 95% PI



Ro	SQRT(Downtime)	Availability %	Fit	SE Fit	95% PI (Lwr, Upr)	
6	7.714920609	99.9887	99.98834	0.00052	99.9829	99.99443
11	0	100	100.00148	0.00042	99.99575	100.0072
18	1.769180601	99.9994	99.99854	0.00033	99.99283	100.0042
22	3.101612484	99.9982	99.99633	0.00031	99.99062	100.0020
31	6.056401572	99.993	99.99142	0.00041	99.98569	99.99715
32	6.980687645	99.9907	99.98988	0.00047	99.98414	99.99563
34	7.643297718	99.9889	99.98878	0.00051	99.98302	99.99454
38	11.51520734	99.9747	99.98235	0.00081	99.97645	99.98825
48	0	100	100.00148	0.00042	99.99575	100.0072
49	0	100	100.00148	0.00042	99.99575	100.0072
56	7.071067812	99.9945	99.98973	0.00047	99.98398	99.99548
63	0	100	100.00148	0.00042	99.99575	100.0072
71	0	100	100.00148	0.00042	99.99575	100.0072
84	4.031128874	99.9969	99.99478	0.00032	99.98907	100.0004
87	4.847679857	99.9955	99.99343	0.00035	99.98771	99.99914
99	11.9749739	99.9727	99.98159	0.00085	99.97567	99.9875

this p-value, we can reject the null hypothesis that the models are the same. Since p-value is < 0.05, the models differ, so the additional predictors in the multiple model do account for enough variance (and have explanatory and predictive power).

Tables 10 and 11 illustrate and validate the results with actual observed availability values contrasted against the 95% prediction interval along with the associated Fitted values and Standard Error. For the simple regression model, Table 10 row 6 (circled in red) illustrates a downtime of 59.52 minutes (transformed with the square root which was 7.7149), with the actual availability of 99.9887%. This is within the 95% PI. The PI lower limit illustrates the minimum availability % which could be as low as 99.98290%. The fitted availability was 99.98834%.

For the multiple regression model, Table 11 row 6 (same observation circled in blue) illustrates a performance of 1066 (which represents latency and throughput values), plus the interaction of downtime and trust level which is 54.83578 (59.92 minutes $*$ 89.13% trust), plus the transformed downtime per location of data centers (using polynomial of degree 2 for 59.92 minutes $*$ 2 which represents APAC location) which is 32062.2336, with the actual availability of 99.9887%. This is within the 95% PI. The PI lower limit depicts the minimum availability % which could be as low as 99.99240%. The fitted availability is 99.98899%.

V. CONCLUSION

Effective management of service levels (e.g. availability, reliability, performance, security) and financial risk to CSCs is dependent on the CSP's capability and trustworthiness. Cloud

TABLE 11. Multiple Regression Predictions 95% PI

Row	Performance	(Downtime * Trust)	Poly(Downtime * Location),2)	Availability %	Fit	SE Fit	95% PI (Lwr, Upr)	
6	1066	54.835776	32062.2336	99.9887	99.98899	0.00035	99.98709	99.9909
11	1386	0	0	100	100.0003	0.00014	99.99851	100.00209
18	1700	2.883669	45.4476	99.9994	99.99962	0.00011	99.99783	100.0014
22	1700	8.862906	389.4176	99.9982	99.99862	0.00012	99.99683	100.00041
31	1736	32.692884	5455.0496	99.993	99.99424	0.00025	99.9924	99.99608
32	2016	43.433049	9595.9116	99.9907	99.99187	0.0003	99.99	99.99374
34	1332	52.069746	13768.4256	99.9889	99.99049	0.0003	99.98863	99.99236
38	1332	118.18638	70596.24	99.9747	99.9736	0.00079	99.97123	99.97596
48	1991	0	0	100	99.99989	0.00014	99.99809	100.00168
49	2002	0	0	100	99.99988	0.00014	99.99809	100.00167
56	1215	44.565	2550	99.9945	99.9923	0.00022	99.99048	99.99413
63	1920	0	0	100	99.99993	0.00013	99.99814	100.00173
71	1391	0	0	100	100.0003	0.00014	99.99851	100.00209
84	1584	14.3975	0	99.9969	99.9974	0.00012	99.99561	99.99919
87	1710	21.65055	0	99.9955	99.99592	0.00012	99.99413	99.99771
99	1632	127.0524	0	99.9727	99.97574	0.00049	99.97371	99.97777

SLAs and a model for assessing, predicting and governing compliance of cloud services and service levels is required. Industry efforts are underway to standardize cloud SLAs [52], [14], [16], [29], [13] and examine frameworks for assessing CSP trustworthiness [11], [49]. Existing research related to the lifecycle of cloud SLAs and CSP trust models was leveraged and extended.

The primary objective of this research was to analyze cloud SLAs and cloud security requirements, assess and define CSP trustworthiness levels, and establish a model for predicting cloud service and SLA availability based on multiple factors (e.g. CSP trust levels, cloud service historic performance and cloud service characteristics). The research hypothesis asserted that cloud service availability could be predicted with greater accuracy when based on more criteria than merely historical downtime. The other criteria considered included CSP trustworthiness, global cloud service locations, cloud service capacity and cloud service performance.

To test the hypothesis, a null hypothesis was defined that proposed a simple linear regression model with a single regression coefficient for cloud service downtime had zero difference in accuracy vs a multiple linear regression model with multiple regression coefficients. In other words, the hypothesis asserted that no additional variables besides cloud service downtime offered explanatory or predictive power to the linear regression model. The alternative hypothesis stated that at least one additional variable provided statistically significant contribution to explaining cloud service availability and therefore should be in the model. Based on multiple Goodness of Fit measures and validation results reviewed in Section IV (Results), the null hypothesis was rejected. Predicting cloud service and SLA availability performance does have greater accuracy when calculated with more predictor variables than just historical cloud service downtime. With Figure 5, we can visualize the improved fit of the multiple regression model (blue solid line) vs the simple regression model (red dotted line), against the actual observed (black line) availability % and downtime minutes.

Below are five scenarios depicting how results of this research study, CSP trustworthiness and the predictive model could be applied. Organizations can utilize the model to:

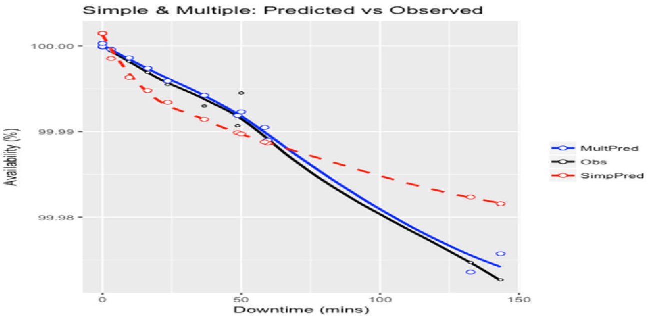


FIGURE 5. Simple vs multiple availability.

- Predict and set expectations concerning quality of cloud computing services.
- Identify required levels of investment to support the required quality of cloud computing service.
- Identify improvement programs to enhance cloud service capabilities and SLA and regulatory compliance.
- Identify opportunities to right size cloud service capabilities based on financial drivers, competitive market conditions, customer requirements.
- Evaluate and select CSPs or compare and contrast private vs public cloud computing services.

Related to the scenarios above, organizations (i.e. CSCs or CSPs) can specify the required prediction intervals to drive required values of corresponding predictor variables which ultimately ensure predictions remain within the expected prediction interval. This reflects a reverse engineering exercise that reconciles the required capabilities and security in support of the required prediction interval. The required predictor variable values can provide visibility and awareness to the required cloud service capabilities and budgets (i.e. traceability and transparency from expectations to the required investment levels and cloud service capabilities that satisfy requirements). The predictive model can apply to both public cloud and/or private cloud scenarios (e.g. organizations that are comparing whether to utilize private vs public vs hybrid; or organizations that are assessing the required investment to their internal private cloud based on required cloud SLA availability levels).

ACKNOWLEDGMENT

The author thanks Dell Technologies ISG Chief Technology Officer for supporting the publication and Gartner (via TechPlanner) for providing CSP data in support of doctoral research.

REFERENCES

[1] M. I. M. Almanea, “Cloud advisor - a framework towards assessing the trustworthiness and transparency of cloud providers,” in *Proc. IEEE/ACM 7th Int. Conf. Utility Cloud Comput.*, 2014, doi: 10.1109/ucc.2014.168.

[2] S. Chakraborty and K. Roy, “An SLA-based framework for estimating trustworthiness of a cloud,” in *Proc. IEEE 11th Int. Conf. Trust, Secur. Privacy Comput. Commun.*, 2012, doi: 10.1109/trustcom.2012.84.

[3] S. Chatterjee and A. S. Hadi, *Regression Analysis by Example*. Somerset: Wiley, 2013.

[4] C. A. I. Q. CSA (2017). “Consensus Assessments Working Group.” Cloud Security Alliance. Accessed: Jul. 02, 2017. https://cloudsecurityalliance.org/group/consensus-assessments/#_overview

[5] C. C. M. CSA (2017). “Cloud Controls Matrix Working Group.” Cloud Security Alliance. Accessed: Jul. 02, 2017. https://cloudsecurityalliance.org/group/cloud-controls-matrix/

[6] S. T. A. R. CSA (2017). “CSA Security, *Trust & Assurance Registry (STAR)*.” Cloud Security Alliance. Accessed: Jul. 15, 2017. https: //cloudsecurityalliance.org/star/

[7] R., Diestel. *Graph Theory*. Berlin: Springer, 2005.

[8] DMTF (2015). “Cloud Infrastructure Management Interface (CIMI) Use Cases.” www.dmtf.org. Feb. 26, 2015. https://www.dmtf.org/sites/default/files/standards/documents/DSP2042_1.0.0.pdf

[9] V. C. Emeakaroha, I. Brandic, M. Maurer, and S. Dustdar, “Low level metrics to high level SLAs - LoM2HiS framework: Bridging the gap between monitored metrics and SLA parameters in cloud environments,” in *Proc. Int. Conf. High Perform. Comput. Simul.*, 2010, doi: 10.1109/hpcs.2010.5547150.

[10] Gartner (2017). “Gartner Technology Planner Cloud Module.” Gartner. Accessed: 2017. http://techplanner.gartner.com/ia.aspx?

[11] N. Ghosh, S. K. Ghosh, and S. K. Das, “SelCSP: A framework to facilitate selection of cloud service providers,” *IEEE Trans. Cloud Comput.*, vol. 3, no. 1, pp. 66–79, Jan.–Mar. 1, 2015.

[12] F. E. Harrell, *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Cham: Springer, 2015.

[13] G., Hogben, and M. Dekker. “Procure Secure: A guide to monitoring of security service levels in cloud contracts” *Enisa.europa.eu*. Apr. 02, 2012. https://www.enisa.europa.eu/publications/procure-secure-a-guide-to-monitoring-of-security-service-levels-in-cloud-contracts/at_download/fullReport

[14] L. Hunnebeck, C. Rudd, S. Lacy, and A. Hanna, *ITIL Service Design*. London: TSO, 2011.

[15] IDTC (2016). "Data Breach Reports - 2016 End of Year Report.” www.idtheftcenter.org. Jan. 18, 2017. http://www.idtheftcenter.org/images/breach/2016/DataBreachReport_2016.pdf

[16] ISO/IEC (2016). “Information Technology. Cloud Computing. Service Level Agreement (SLA) Framework.” doi: 10.3403/bsisoiec19086.

[17] ISO/IEC 19086-1 (2016). "ISO - international organization for standardization.” ISO/IEC 19086-1:2016 - Information Technology – Cloud Computing – Service Level Agreement (SLA) Framework – Part 1: Overview and Concepts. Sep. 21, 2016. https://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=67545

[18] I. S. O. /. I. E. C. 19086-2 (2017). "ISO - International Organization for Standardization.” ISO/IEC DIS 19086-2 - Information Technology – Cloud Computing – Service Level Agreement (SLA) Framework – Part 2: Metric Model. Jul. 17, 2017. https://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=67546

[19] I. S. O. /. I. E. C. 19086-3 (2017). "ISO - International Organization for Standardization.” ISO/IEC 19086-3:2017 - Information Technology – Cloud Computing – Service Level Agreement (SLA) Framework – Part 3: Core Conformance Requirements. Jul. 20, 2017. https://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=67547

[20] J. Kanpariyasoontorn and T. Senivongse, “Cloud service trustworthiness assessment based on cloud controls matrix,” in *Proc. 19th Int. Conf. Adv. Commun. Technol.*, 2017, doi: 10.23919/icact.2017.7890100.

[21] M. H. Kutner, C. Nachtsheim, and J. Neter, *Applied Linear Regression Models*. Boston, MA, USA: McGraw-Hill, 2008.

[22] D. S. Linthicum, “Moving to autonomous and self-migrating containers for cloud applications,” *IEEE Cloud Comput,*, vol. 3, no. 6, pp. 6–9, Nov./Dec. 2016.

[23] J. Luna, R. Langenberg, and N. Suri, “Benchmarking cloud security level agreements using quantitative policy trees,” in *Proc. ACM Workshop Cloud Comput. Secur. Workshop.*, 2012, doi: 10.1145/2381913.2381932.

[24] J. Luna, A. Taha, R. Trapero, and N. Suri, “Quantitative reasoning about cloud security using service level agreements,” *IEEE Trans. Cloud Comput.*, vol. 5, no. 3, pp. 457–471, Jul.-Sep. 1, 2017.

[25] K. Mason and E. Krieger, “Public cloud customer research 1H16,” *Cloud Bus. Quart.*, Tori.com. Jun. 9, 2016. http://www.tbri.com/library/2q16/clbq/pubcloudcustres/tbr_public_cloud_research_clbq_1h16.pdf

[26] J. Mitchell, S. Rizvi, and J. Ryoo, “A fuzzy-logic approach for evaluating a cloud service provider,” in *Proc. 1st Int. Conf. Softw. Secur. Assurance*, 2015, doi: 10.1109/icssa.2015.014.

[27] M. K. Naseer, S. Jabbar, and I. Zafar, “A novel trust model for selection of cloud service provider,” in *Proc. World Sympo. Comput. Appl. Res.*, 2014, doi: 10.1109/wscar.2014.6916772.

[28] M. Natu, R. K. Ghosh, R. K. Shyamsundar, and R. Ranjan, “Holistic performance monitoring of hybrid clouds: Complexities and future directions,” *IEEE Cloud Comput.*, vol. 3, no. 1, pp. 72–81, Jan./Feb. 2016, doi: 10.1109/mcc.2016.13.

[29] NIST (2015). “Cloud Computing Service Metrics Description.” Nist.gov. 2015. http://www.nist.gov/itl/cloud/upload/RATAX-CloudServiceMetricsDescription-DRAFT-20141111.pdf

[30] Ponemon (2017). This Site Produced and Maintained by Byte Productions, www.byte-productions.com. “News & Updates.” Ponemon Institute. Accessed: Jul. 02, 2017. http://www.ponemon.org/blog/2016-cost-of-data-center-outages

[31] L. Qu, Y. Wang, and M. A. Orgun, “Cloud service selection based on the aggregation of user feedback and quantitative performance assessment,” in *Proc. IEEE Int. Conf. Services Comput.*, 2013, doi: 10.1109/scc.2013.92.

[32] RightScale (2017a). "RightScale 2017 State of the Cloud Report.” State of the Cloud Report | RightScale. Accessed: Jul. 02, 2017. https://www.rightscale.com/lp/state-of-the-cloud

[33] S. Ristov, and M. Gusev, “A methodology to evaluate the trustworthiness of cloud service providers’ availability,” in *Proc. IEEE EUROCON - Int. Conf. Comput. as a Tool*, 2015, doi: 10.1109/eurocon.2015.7313734.

[34] S. S. Roy, T. H. Sarker, and M. M. A. Hashem, “A novel trust measurement system for cloud-based marketplace,” in *Proc. 2nd Int. Conf. Elect. Inf. Commun. Technologies.*, 2015, doi: 10.1109/eict.2015.7391921.

[35] A. Ruan, M. Wei, A. Martin, D. Blundell, and D. Wallom, “Breaking down the monarchy: Achieving trustworthy and open cloud ecosystem governance with separation-of-powers,” in *Proc. IEEE 9th Int. Conf. Cloud Comput.*, 2016, doi: 10.1109/cloud.2016.0073.

[36] R. W. Saaty, “The analytic hierarchy process - what it is and how it is used,” *Math. Modelling*, vol. 9, no. 3-5, pp. 161–76, 1987.

[37] T L. Saaty, *Decision Making for Leaders the Analytic Hierarchy Process for Decisions in a Complex World*. Pittsburgh, PA, USA: RWS Publ., 2012.

[38] T L. Saaty, “How to make a decision: The analytic hierarchy process,” *Eur. J. Oper. Res.*, vol. 48, no. 1, pp. 9–26, 1990.

[39] T L. Saaty, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. New York: NY, USA: McGraw-Hill Int. Book, 1980.

[40] B., Satyanarayana and K. S. Prasad, *Discrete Mathematics and Graph Theory*. Delhi, India: PHI Learn. Private Limited, 2014.

[41] Skyhigh (2017). "IaaS and Custom Applications Trends Report 2017.” Skyhigh. Accessed: Jul. 02, 2017. https://www.skyhighnetworks.com/csa-report-4/

[42] A. W. S. SLA (2017). “AWS Service Level Agreements.” Amazon AWS. Accessed: Jul. 15, 2017. https://aws.amazon.com/legal/service-level-agreements/

[43] S. L. A. Google (2017). “Google Cloud Platform Service Level Agreements.” Google. Accessed: Jul. 15, 2017. https://cloud.google.com/terms/sla/

[44] S. L. A. Microsoft (2017). "SLA summary for Azure services” Microsoft Azure. Accessed: Jul. 15, 2017. https://azure.microsoft.com/en-us/support/legal/sla/summary/.

[45] A. W. S. STAR (2017). “STAR Registrant Amazon AWS.” Amazon AWS : Cloud Security Alliance. Accessed: Jul. 15, 2017. https://cloudsecurityalliance.org/star-registrant/amazon-aws/

[46] S. T. A. R. Google (2017). “STAR Registrant Google.” Google : Cloud Security Alliance. Accessed: Jul. 15, 2017. https://cloudsecurityalliance.org/star-registrant/google/

[47] S. T. A. R. Microsoft (2017). “STAR Registrant Microsoft.” Microsoft : Cloud Security Alliance. Accessed: Jul. 15, 2017. https://cloudsecurityalliance.org/star-registrant/microsoft/

[48] M. Supriya, K. Sangeeta, and G. K. Patra, “Comparison of AHP based and fuzzy based mechanisms for ranking cloud computing services,” in *Proc. Int. Conf. Comput., Control, Informat. Its Appl.*, 2015, doi: 10.1109/ic3ina.2015.7377768.

[49] A. Taha, R. Trapero, J. Luna, and N. Suri, “AHP-based quantitative approach for assessing and comparing cloud security,” in *Proc. IEEE 13th Int. Conf. Trust, Secur. Privacy Comput. Commun.*, 2014, doi: 10.1109/trustcom.2014.39.

[50] P. Tattar, S. Ramaiah, and B. G. Manjunath, *A Course in Statistics with R*. Chichester, West Sussex, U.K.: John Wiley Sons, 2016.

[51] Z. Wu and Yu Zhou, “Customized Cloud Service Trustworthiness Evaluation and Comparison Using Fuzzy Neural Networks,” in *Proc. IEEE 40th Annu. Comput. Softw. Appl. Conf.*, 2016, doi: 10.1109/compsac.2016.86.

[52] EC (2014). "Cloud Service Level Agreement Standardisation Guidelines.” Ec.europa.eu. Jun. 24, 2014. http://ec.europa.eu/newsroom/dae/document.cfm?action=display&doc_id=6138

[53] J. Diaz-Montes, I. Petri, O. Rana, and M. Parashar, “Special issue on autonomic clouds,” *IEEE Cloud Comput.*, vol. 3, no. 3, pp. 26–29, May/Jun. 2016.

[54] N. Ferry, A. Rossini, F. Chauvel, B. Morin, and A. Solberg, “Towards model-driven provisioning, deployment, monitoring, and adaptation of multi-cloud systems,” in *Proc. IEEE Sixth Int. Conf. Cloud Comput.*, 2013, doi: 10.1109/cloud.2013.133.

[55] R., Maeser. *A Model-Based Framework for Analyzing Cloud Service Provider Trustworthiness and Predicting Cloud Service Level Agreement Performance*. DEng diss., The George Washington University, 2018, doi: 10785821.

ROBERT MAESER (Senior Member, IEEE) received the bachelor’s degree in computer science from the University of Minnesota, USA, the M.S. degree in software engineering from the University of St. Thomas, USA, and the D.Eng. degree from The George Washington University, USA. He is a member of technical staff for the Office of the CTO at Dell Technologies. He holds active ISC^2 CISSP and CCSP certifications. His engineering and research are focused in the areas of machine learning, cloud computing, security, and software architecture.