

# A Framework for Evaluating Data Quality on Military Enterprise Networks

Lee P. Battle<sup>1,2</sup>, Edward F. Harrington<sup>1,3</sup>

<sup>1</sup>Maritime Patrol and Reconnaissance Aircraft Program Office, NAVAIR, Patuxent River, USA

<sup>2</sup>SAVVEE Consulting, Lexington Park, USA

<sup>3</sup>Defence Science and Technology Group, Canberra, Australia

## Email address:

Lee.battle.ctr@navy.mil (L. P. Battle), edward.harrington.as@navy.mil (E. F. Harrington)

## To cite this article:

Lee P. Battle, Edward F. Harrington. A Framework for Evaluating Data Quality on Military Enterprise Networks. *International Journal on Data Science and Technology*. Vol. 2, No. 6, 2016, pp. 76-83. doi: 10.11648/j.ijdst.20160206.14

**Received:** August 15, 2016; **Accepted:** November 19, 2016; **Published:** December 21, 2016

---

**Abstract:** This paper introduces a framework to determine data quality on enterprise networks for net-centric and net-ready initiatives as introduced by the US Department of Defense (DoD). Traditionally quality of data delivered to an enterprise user focuses on network performance, i.e. *quality of service* (QoS). It is proposed to add two new attributes pertaining to data sharing performance to QoS: *data relevance* (DR) and *quality of data at source* (QDS); and further a method to evaluate these new attributes. The QDS attribute brings distinction to the resultant data quality of the network's quality of service. This distinction is necessary to reflect the separation in procurement and management for sensor systems and network systems for the DoD. The DR attribute is introduced; it is important in enabling enterprise data consumers to sort, filter and prioritize data. There is also a need to assess the quality of data sharing across the enterprise network. One recent method subjectively assess the quality of data is to measure the user satisfaction referred to as *quality of experience* (QoE). The QoE is assessed for each of the framework's attributes using the best practices from survey statistics in sampling and estimation. The overall value of data quality on enterprise networks is decided using a minimax decision model consisting of the three attributes. The resultant minimax value correlates to the lowest performing attributes of the framework. The minimax decision model is chosen to meet the design philosophy that little advantage to the overall enterprise network performance will result from further investment in high performing attributes prior to balancing performance across all three model attributes. The presented framework offers decision support tools to enable agencies to allocate limited resources towards improving the performance of their net-centric service offerings to the enterprise network.

**Keywords:** Data Quality, Data Relevance, Quality of Data at Source, Quality of Service, Net-Centric, Net-Ready, Quality of Experience, Minimax

---

## 1. Introduction

To achieve information dominance often there is focus on mechanisms that increase bandwidth and rates of data transfer; predicated upon the assumption that all data is good data. As a result, an ever-increasing amount of data is being created for military operators and commanders through the evolution of sensor capability and increased situational awareness data sources. The ability to generate data combined with the ability to store data is beginning to outpace the capacity of the bandwidth necessary to simultaneously share *all* data amongst all the operational stakeholders. That same data abundance

challenges the network capacity and overloads the capacity of the human operator. A fundamental shift in paradigm is required to ensure the stressors of conducting military operations are supported through data (of both high-quality and high-relevance) and, not burdened by attempts to manage its excess.

A movement to shift from "platform-centric warfare" to "network-centric warfare" was initiated in the final years of the 20<sup>th</sup> century [1]. The US Department of Defense (DoD) introduced the terms *net-centric* and *net-ready* to describe the mechanisms by which operators can search and discover information within the bounds of data. Net-ready also introduced concepts to improve access through mechanisms

of data entry and network management; all in pursuit of increasing support to military operations.

To implement net-ready, system developers add data tagging, search algorithms, additional communications paths, and a suite of tools to exploit the new forms of data organization to help operators sift more rapidly through data to find relevant information. At the same time, developers are also facing a new host of challenges from increasing cyber threats. All of this is in addition to the historical problems of network management and quality of service.

When discussing the theoretical, it is easy to dismiss the challenges of limited resources to implement new policies such as schedule, budget, manpower, etc. Successful policy implementation amongst other factors requires system analysis methods that assist acquisition agencies in targeting limited developmental resources to areas of greatest impact to the overall mission objectives. This paper introduces such a method in support of achieving maximum data quality for military enterprise networks: a quantitative mechanism by which the value of different net-ready implementation options can be evaluated and graded. Importantly, a model for valuing enterprise data quality is introduced to bridge the gap between measure of technical performance and operational benefit. The discussion of quality of data in a communications network is usually limited to the *Quality of Service* (QoS) which measures user's satisfaction based on network performance metrics like latency and bandwidth. Traditional QoS expects prioritization has occurred prior to data entering the network. The value of data quality can be used to indicate the priority of data delivery to the consumer. Not all data is the same; some is more relevant to the users' needs when compared across all the data. Thus the term *data relevancy* (DR) is introduced into the model for valuing data quality in the context of net-centric / net-ready.

Section 2 of this paper provides a brief overview of the origins of net-centric and net-ready policies. Section 3 defines a new model for valuing the quality of data by measuring user satisfaction. The attributes in the new model represents data quality of the enterprise system within three contexts: net-centric measures, traditional network quality of service, and cyber-security. Section 4 presents a method for applying the new model attributes to evaluate data quality across the enterprise.

## 2. Net-Centric/Net-Ready

### 2.1. The Origins of Net-Centric Warfare

Admiral Jay L. Johnson, Chief of Naval Operations, stated in 1997 that the military is undergoing "a fundamental shift from what we call platform-centric warfare to what we call network-centric warfare" [1]. Vice Admiral Arthur K. Cebrowski, U.S. Navy, and John J. Garstka proposed that adoption of network-centric operations across the military enterprise would result in benefits at the strategic, operational, and structural levels and bring forth "a much faster and more effective warfighting style" [2]. This new warfighting style is

Net-Centric Warfare (NCW): an "information superiority-enabled concept of operations that generates increased combat power by networking sensors, decision makers, and shooters to achieve shared awareness, increased speed of command, higher tempo of operations, greater lethality, increased survivability, and a degree of self-synchronization" [1]. Vital to the value of NCW is "the content, quality, and timeliness of information moving between nodes on the [enterprise] network." [2].

### 2.2. Net-Centric: Attributes and Objectives

The DoD introduced four criteria that must be satisfied for "Data Sharing in a Net-Centric Environment" via DoD Directive 8320.02 in 2004 [4] which later in [5] expanded to the seven listed in this section. This directive and the subsequent series of 8320 series documents identify "the cornerstones of Net-Centric Data sharing"; data shall be visible, accessible, understandable, and trusted [6]. In the years following, the Chairman of the Joint Chiefs of Staff Instruction (CJSCI) 6212.01 was released that set-forth the procedures for development and certification of a Net-Ready (NR) Key Performance Parameter (KPP); the NR KPP process later being subsumed into the Joint Capabilities Integration and Development System (JCIDS) process [3] [7]. The NR KPP specifies the attributes required of data sharing Information Technology (IT) introduced into the net-centric operational environment:

- (1) IT must be able to *support military operations* (SMO),
- (2) IT must be able to be *entered and managed on the network* (EMN), and
- (3) IT must effectively *exchange information* (EI) [3].

#### 2.2.1. Data Shall Be Visible

Making data visible is achieved via deployment of discovery capabilities that access and search data asset catalogs and registries in the enterprise [6]. These enterprise catalogs and registries contain discovery metadata entries for an individual data asset or group of data assets [6].

#### 2.2.2. Data Shall Be Accessible

Making data accessible requires providing authorized users the ability to view, transfer, download, or otherwise consume a data asset once discovered [6]. The access process should be via the "network" using "commonly supported access methods" [6].

#### 2.2.3. Data Shall Be Understandable

Making data understandable requires alignment of terminology, data protocols, data formats, and data meaning between producer and consumer [6]. Alignment can be achieved via direct negotiation or—more practically—via the adoption of commonly referenced standards such as those listed indicated by the *global information grid* (GIG) Technical Guidance Federation [6].

#### 2.2.4. Data Shall Be Trusted

A consumer's trust in a data asset is dependent on multiple facets: assessment of the data asset authority, clear

identification of the data asset source, tagging with appropriate security metadata, and maintaining a full pedigree of the data asset throughout the full process [6]. Untrusted data can introduce error, uncertainty, and delay into the military decision process [8].

### 2.2.5. IT Must Be Able to Support Military Operations (SMO)

To satisfy the attribute of support to military operations, IT deployed to the operational environment must support identifiable net-centric operational tasks and mission objectives [3]. Net-centric operational tasks are those that “produce information, products, or services for or consume information, products, or services from external IT” [3]. The performance of the IT must be quantifiable with threshold and objective values that are traceable to measures of effectiveness (MOEs) [3].

### 2.2.6. IT Must Be Able to Be Entered and Managed on the Network (EMN)

To satisfy the attribute of entered and managed on the network, the IT connections to external networks required to perform net-centric operational tasks must be identified [3]. The identification of the connections must be specific (i.e. non-generic); the required performance of the connections be identified by quantifiable and testable measures of performance (MOPs); and the connectivity must be managed by a structured methodology [3].

### 2.2.7. IT Must Effectively Exchange Information (EI)

The specific data elements and assets exchanged with external networks as part of executing net-centric operational tasks are specified by the exchange information attribute [3]. Each net-centric information exchange defined MOPs that are measurable and each information exchange must also identify how the four criteria for net-centric data sharing (visible, accessible, understandable, and trusted) are satisfied for authorized consumers across the enterprise [3].

## 3. QoE and Enterprise Attributes

Evaluation of the data sharing enterprise requires a holistic view that considers the net-centric attributes of the data simultaneously with the quality of service for the data network. Additionally, the interdependence between cyber security and net-centric principles are indicated in the most recent update to the DoD’s instruction for enterprise data sharing [5]. The combined consideration for each of these areas yields a newly defined model for the data sharing enterprise comprised of three equally important attributes: *data relevance*, *quality of data at source*, and *quality of service* for the enterprise network. Fig. 1 illustrates the mapping relationship for each enterprise attribute to various DoD objectives of net-centric, net-ready and cyber-security.

### 3.1. Survey Method for Evaluating Each Attribute’s QoE

The subjective measure of overall user-satisfaction of a service or application is referred to as *quality of experience*

(QoE) [9]. The most commonly used subjective rating in standards and in conjunction with QoE is the *mean opinion score* (MOS) where score of 1 is bad, 2 is poor, 3 is fair, 4 is good, and 5 is excellent. Sampling users is the preferred and direct method for measuring QoE. Where an accurate model exists for QoE as a function of that attribute’s objective measures then it’s used instead of sampling users’ opinions. For most data the paper assumes that a model is inadequate and sampling the users’ opinion is required.

The question arises how to get the users’ opinions. A novel approach to survey online users through some form of random sampling of enterprise users is proposed. Usually the MOS is formed from arithmetic mean of a population of user opinions. But the arithmetic mean assumes that all user groups are of equal size which could lead to biased estimates of the MOS. Thus to avoid bias consider unequal sampling of the user population by using login authentication to identify users to form strata of homogenous users. One possible split is to consider groups of users based on their shared mission or objectives. The survey is not just restricted to a complete enterprise system but can be performed in the early design phases of prototypes and help analyze operational performance of an enterprise attribute as a function of its objective measures, i.e. develop or enhance the prediction model of QoE for each attribute.

A novel application to networks of a common survey practice is proposed to use the Horvitz-Thompson (HT) estimator [10] for determining the value of the total score of the population of size  $N$ . HT is commonly used because of its versatility as an unbiased estimator of the total for a sampling/sample with or without replacement. For a sample of size  $n$  consider  $v$  strata with sample  $s_k$  in each stratum, i.e.  $\sum_{i=1}^{v} n_i = n$  and  $s = \bigcup_{k=1}^{v} s_k$ , where  $n \leq N$  and  $s = \{1, 2, \dots, n\}$ .

The Horvitz Thompson estimator for stratified sampling for the total  $Y$  is

$$y_{HT} = \sum_{k=1}^{v} \sum_{i \in s_k} y_i / p_i \quad (1)$$

where each independent random sample  $i$ , is denoted by  $y_i$ , the probability of inclusion in the sample is denoted by  $p_i$ .

Where it is not easy to design a stratified sampling plan a stratification post collection of  $n$  samples without replacement can be constructed. The sample is organized into their various strata with the number of elements in each total number of strata  $k$ , in the total population  $N$ . The HT estimator for stratification becomes

$$y_{HT} = \sum_{k=1}^{v} (N_k/n_k) \sum_{i \in s_k} y_i \quad (2)$$

and estimate of the stratified mean is

$$u_{HT} = \sum_{k=1}^{v} (N_k/N) m_k \quad (3)$$

where  $m_k = \sum_{i \in s_k} y_i / n_k$  is the arithmetic mean of strata  $k$ . With each attribute having its MOS value using the stratified estimate in (3).

### 3.2. Quality of Data at Source

An important part of assessing the end-to-end performance of a data system is consideration of the inherent quality of the originating data prior to its entry into the network; to be referred to as the *quality of data at source* (QDS). The originating data can vary from content on webpages; chat room information; geolocation data; audio; and data from electro-optic, infrared, or radar sensors.

The QDS takes into account the effects of environmental conditions on sensor performance for given design parameters. QDS is a subjective rating from the perspective of the end user [11]. Traditionally these QoE MOS ratings were undertaken by panels of experts. But cost and time to use panels of experts to assess MOS has resulted in seeking alternative approaches. One alternative approach is using models formed from objective perceptual measurements to predict subjective ratings of QoE [12]. An example of such an estimate for audio (speech wireless, VOIP, fixed, clean) is use of objective measures to form the *perceptual evaluation of speech quality* (PESQ) model. Full reference PESQ is formed by taking the speech codec output and comparing it with the original signal inputted into the codec [13]. To produce a subjective rating MOS, or QoE, the ITU-T P.862.1 [13] is used to map raw PESQ to the final rating.

There are three levels of reference used in determining the models for estimating the subjecting ratings: full reference (undistorted service is available for comparison with distorted outcome), partial (or reduced) reference, and no reference. In the video standards of [14] and [15] they present a series of methods of modelling with objective measurements to predict the subjective ratings. In [15] it's recommended that selection of the video model for sensor calibration use the full reference to determine the quality. The reason to use the full reference is to capture environmental conditions resulting in the most accurate predictions of ratings. Video and audio have models for QoE but other data types still require development of objective measure models to predict their QoE subjective ratings.

Subjective judgment or QoE is not just about the quality of the video alone, but is also about assessing the usefulness of video to complete a task of identification of human and objects [16]. The prevailing method for assessing the quality of a still image is based on the ability to perform certain levels of object recognition with scoring defined by the *national imagery interpretability rating scale* (NIIRS) [17]. A standard to address the rating of motion imagery (i.e. video) has been recently introduced by the *motion imagery standards board* (MISB) and is based on the ability to identify objects in the video according to seven orders of battle [18]. In [17] and [15] there are equations to estimate the interpretability for still imagery and video, respectively. In [15] equations are formulated for quality in terms of interpretability and MOS; also with or without references.

### 3.3. Quality of Service

Over the last several decades there have been many papers

on the topic of QoS. But just for completeness a brief explanation of QoS is given here.

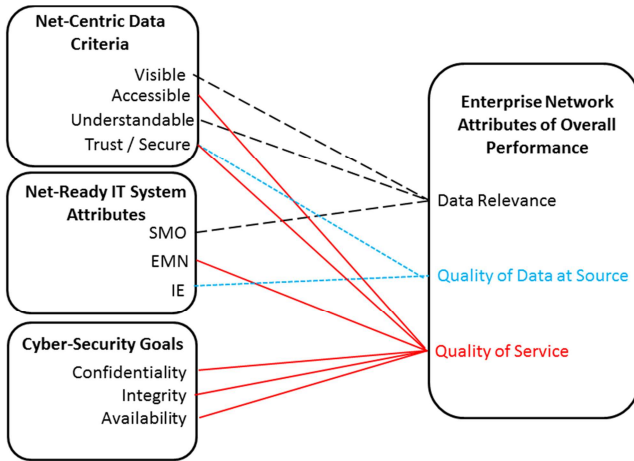
QoS is in essence an engineering optimization problem where the objective is to maximize users' satisfaction while minimizing cost of delivery of the supporting network services. User satisfaction is traditionally associated with network metrics: delay, jitter, throughput, packet loss, order preservation. And the *service level agreement* (SLA) is the users' agreement with network provider(s) on acceptable ranges for the metrics. The most widely deployed QoS architecture used to deliver a SLA on an enterprise IP/MPLS network is referred to as a *differentiated service* (Diffserv) [11].

User-satisfaction provides a true gauge of a network QoS [11] and the subjective assessment of that satisfaction is provided by QoE. However there are a number of challenges to QoE discussed in [19] and [20]. To understand cause and effect it is ideal to have the full reference i.e. data at the source as well as objective measurements of that data on the network that can be used to correlate with the end user's QoE. Examples include understanding relationships between objective measures of QoS like jitter, throughput, and latency to be able to control the QoE. As stated in [20] QoE is likely to be biased by negative responses, tendency to higher responses from unhappy consumers. This type of sample bias is not a concern in this paper because the designed stratification covers all consumers, and the HT estimation technique is used to produce an unbiased estimate of QoE. It is expected that models of QoE can be accurately built as a function of objective measures from network, applications, environment, and terminals. These models of QoE based on QoS objective measures are often referred to as "QoE\QoS" correlators [21] and [12].

Models have been developed to correlate the QoS with QoE for multimedia applications [12]. In [21] QoE for voice was modeled with QoS metric packet loss as the objective measure; and other QoE model for web browsing session times as a function of QoS metric of throughput. For both models [21] showed QoE had an exponential model in terms of their QoS objective metric(s). A number of authors [12, 15] have shown the merit in statistical machine learning classification algorithms being used to analyze network QoS objective measures and to determine their level of contribution to different classification QoE outcomes: *linear discriminant analysis*, *logistic regression*, *decision trees* [22].

### 3.4. Data Relevance

The measure of data relevance expresses the utility provided by the data towards the consumer's objective(s). Highly relevant data must be visible to the consumer, understood by the consumer, and provide support to the consumers' military operations (as depicted in the mapping of Fig. 1).



**Figure 1.** Various goals, objectives, measures of the data DoD focus areas of Net-Centric, Net-Ready, Cyber-Security are mapped to three enterprise attributes of Data Relevance, Quality of Data at Source, Quality of Service.

A scalar measure is required to support evaluating the degree of relevance on the enterprise performance. A simplistic rejection of non-relevant data is insufficient given the negative effect that excess amounts of data can have on human decision makers, even when the data in the network is restricted to only relevant data [23]. The use of the QoE estimate is proposed to provide a subjective rating of the overall relevance of data shared on the enterprise. The three principal elements of data relevance and their respective effects on the QoE for data relevance are now discussed, as depicted in Fig. 2.

### 3.4.1. Intrinsic Data Relevance

Intrinsic data relevance represents the relative value (i.e. QoE) that the data would provide to the consumer assuming perfect discovery and delivery. Intrinsic relevance reflects three properties of the data: form, spatial, and temporal. The data's form properties indicate the suitability for the type of data to convey the required information. Enterprise data systems can offer multiple forms of data to the consumer (e.g. still imagery, motion imagery, acoustic signal, electronic signal, radar) with each revealing different aspects for a given target of interest. The data's spatial properties relate to the location of the data collection sensor relative to the target of interest (e.g. overhead, side-view, rear-view, distant, near). Temporal properties of the data pertain to the time of data collection relative to the actions and conditions of the target of interest (e.g. target while in port, target while in open ocean, target when first detected, target after engagement). One way to analyze these properties and their effect on QoE is to form each stratum of (3) in such a way that we have homogenous groups of consumers with similar intrinsic DR properties.

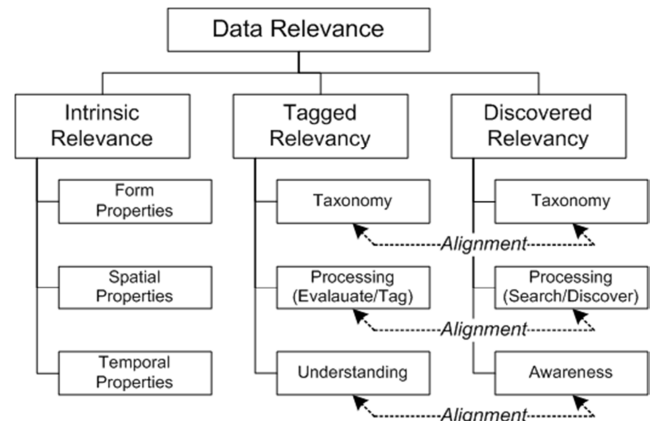
### 3.4.2. Tagged Data Relevance

The measure of tagged data relevance corresponds to how accurate the data producer was in expressing the intrinsic data relevance through the application of metadata that is understandable, relatable, and unambiguous to the data consumer. Tagged data relevance is a function of the quality of the taxonomy, the process performance, and the degree of

understanding the producer has of the potential consumers. The metadata taxonomy needs to be sufficiently diverse to express the essential characteristics of the data product but not so overly detailed that the data tagging approaches the size and complexity of the data itself. The processing of tagged relevance involves the review and analysis of the data product to assess its key features followed by the application of specific metadata values. The goal of the tagging and discovery process is to connect highly relevant data with an authorized consumer. This requires a thorough understanding of the wants, needs, and priorities of the consumer to realize maximum value of tagged relevance, and therefore a higher QoE.

### 3.4.3. Discovered Data Relevance

The measure of discovered data relevance is an indication of how well the enterprise system enables a consumer to differentiate the data product offerings accessible via the network according to the level of relevance and value to the mission objectives. To support a high level of discovered relevance,



**Figure 2.** Decomposition of Data Relevance Measure.

The taxonomy available to the consumer must be sufficient to explicitly discriminate the desired data features from the undesired. A poorly performing taxonomy would be one that prompts extraneous definition of detail or that includes terminology with such subtle variation as to lack the mutual exclusivity necessary to select between one term or the other. The processing element of discovered relevance is a measure of how well the search methods of the system match the descriptive words of the consumer (i.e. metadata) with those of the producer. An identical taxonomy between producer and consumer would simplify this process, but the way of expressing what one needs may not always match the way of expressing what one has to offer [6]. Just as the performance of tagged relevance is improved by an understanding of the consumer by the producer, discovered relevance benefits from awareness that the consumer has of the full range of data offerings of the enterprise and how they may be described. Without a strong mutual awareness, a consumer may prematurely end their discovery process with the first piece of seemingly relevant data believing that it is the best or perhaps

the only product available to them. This is why an overall QoE value for data relevance is important as it indicates the relative success of the discovery process across multiple groups of users at finding data with intrinsic relevancy suitable to meet their needs.

## 4. Data Quality for the Enterprise

We first discuss the design philosophy for the overall assessment of the data quality for the enterprise:

- (1) The enterprise data quality is determined by the attribute with the lowest user satisfaction (QoE).
- (2) A high user satisfaction for the enterprise data quality can only occur when all three attributes have high QoEs.

By focusing on all three attributes we can reduce resources required in one attribute based on the overall value of data quality of the enterprise. For example, if the quality rating is fair for data relevancy then funds to provide an excess bandwidth to have an excellent QoS can be spent elsewhere as the overall data quality at best is going to be fair. Alternatively the data relevance user satisfaction to match the other attributes can be improved.

Thus, aside from determining the value of enterprise data quality for purposes of data prioritization, we also have an analytical tool to identify areas of improvement and to allocate resources more effectively across the overall system. The rest of this section assumes that there is some ability to adjust the designs of the systems associated with the enterprise attributes. We explain here our method to evaluate the overall quality of data of the enterprise based on a minimax decision criterion and its connection with minimax game theory.

### 4.1. Minimax and Game Theory

In game theory, the minimax value of a player is the smallest value that other players can force without knowing the player's actions. The same value is also the largest guaranteed value for that player with knowledge of the other player actions. Formal minimax definition [24] is

$$h_i = \min_{a(-i)} \max_{a(i)} v_i(a(i), a(-i)) \quad (4)$$

where  $a(i)$  denotes the actions of the  $i$ -th player of  $n$  players,  $a(-i)$  is actions of all other players except the  $i$ -th, and  $v_i$  is the value function of player  $i$ .

Consider a simple example in Table 1 to illustrate the minimax in a game theory context. The cells in Table 1 consist of a left value which is the games' pay-off for player one and a right value which is the games' pay-off for player two. Rows in Table 1 represent actions of player one and columns in Table I represent actions of player two. Player one has three action options U, D, N to choose from. Going through each action of Player one with knowledge of player two actions L, R we have maximum payoff for player one for U of 5, for D of 5, for N of 4, making a minimax pay-off action of player one of N. Player two takes each action along columns for action L maximum payoff is 4 and 3 for action R, resulting in minimax action for Player two of R with payoff of 3. Thus, the minimax strategy

is Player one move of N and Player two move of R with a payoff vector (4, 3).

### 4.2. Our Model and Wald's Minimax Criterion

The minimax strategy in game theory inspired the decision theory approach of Abraham Wald's minimax model [25]. The minimax model of Wald produces a decision with an outcome (payoff value) which is the worst chosen amongst the best outcomes of all decisions. The formal definition of Wald minimax model is

$$\min_{d \in D} \max_{s \in S(d)} f(d, s) \quad (5)$$

where  $f(\cdot)$  is pay-off or outcome function,  $d$  is the decision,  $D$  is the decision space,  $S$  is the state space, and  $s$  is the state vector. The model presented for evaluating end-to-end data quality can be written in the form of the Wald minimax model. The first step is to define the decision space  $D$ , consisting of three attributes of quality  $D = \{QDS, QoS, DR\}$ . Each decision  $d$  in  $D$  has a vector state subspace  $S(d)$ , e.g. for  $d = QDS$  have subspace  $S(QDS)$ . For illustration purposes QDS subspace could consist of objective measures SNR, Ground Sample Distance and sensor parameters such as focal length.

Table 1. Example of minimax game.

		Player 2	
		L	R
Player 1	U	5,2	2,1
	D	5,4	0,0
	N	0,2	4,3

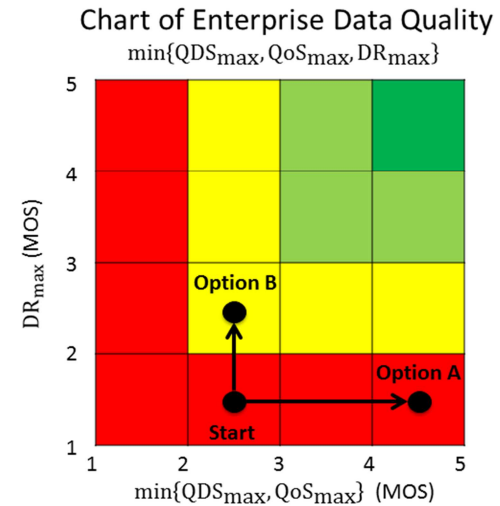


Figure 3. The overall value of the chart is the minimum of the maximums of Data Relevance (DR), Quality of Data at Source (QDS), Quality of Service (QoS). Color map to MOS scaling: red = 1 to less than 2, yellow = 2 to less than 3, light green = 3 to less than 4, green = 4 to less than 5.

And for QoS the objective measures could be latency, packet loss, jitter, and sequence of packets. DR subspace could consist of parameters to measure objectively various properties and elements of data relevance, e.g. percentage of spatial and temporal coverage between consumer desired and actual data, and percentage of ontology alignment user and producer.



Thus, for each  $d$  in  $D$ , starting with  $d=QDS$ , a state vector  $s_{\max}(d)$  is found that produces  $\max_{s \in S(d)} f_d(s)$  where  $f_d(s)$  is a discrete function formed by a predictive model for QoE or in the absence of a model formed from direct user sampled MOS value calculated using (3), i.e.  $f_d(s) = u_{HT}$ . The x-axis of the chart in Fig. 3 is organized to reflect the connection between QDS and QoS. The final enterprise data quality, as shown in Fig. 3, as

$$\min_{d \in D} f_d(s_{\max}(d)) \quad (6)$$

In Fig. 3, a particular value of enterprise data quality is given at the start, but from start one could continue to improve QoS and QDS under Option A with no increase in the value of enterprise data quality, whereas Option B modifying the system delivering DR does increase the overall enterprise data quality.

## 5. Conclusion

This paper introduced a framework to determine data quality on enterprise networks for DoD net-centric and net-ready initiatives. The framework's data quality model consists of three attributes: quality of data at the source, data relevance, and network QoS. The paper described the data quality using minimax decisions based on users' survey ratings for a given enterprise configuration. The final value of data quality for the enterprise network was demonstrated as the attribute with the lowest score. The implemented strategy provides a tool for decision makers to prioritize data and manage their resources without comprising any part of the data sharing system.

The paper notes the definition of the relationships between attribute's objective measures and the final quality of experience. In particular QoS objective measures for audio can be used to predict the QoE as an alternative to conducting surveys. Thus, support is indicated for further research in the development of objective measures using the definition of data relevance elements presented and to determine models for predicting QoE ratings as a function of these objective measures.

## References

- [1] D. S. Alberts, J. J. Garstka and F. P. Stein, "Network Centric Warfare: Developing and Leveraging Information Superiority," *Washington D.C.: DoD CCRP Publication Series*, 2000.
- [2] A. K. Cebrowski and J. J. Garstka, "Network-Centric Warfare: Its Origin and Future," *U.S. Naval Institute Proceedings*, vol. 124, no. 1, pp. 28-35, 1998.
- [3] US Joint Staff Director, "CJCSI 6212.01F Net Ready Key Performance Parameter (NR KPP)," 21 March 2012. [Online]. Available: [http://www.dtic.mil/cjcs\\_directives/cdata/unlimit/6212\\_01.pdf](http://www.dtic.mil/cjcs_directives/cdata/unlimit/6212_01.pdf).
- [4] DoD CIO. Department of Defense Directive 8320.02, "Data Sharing in a Net-Centric Department of Defense," 2 December 2004.
- [5] DoDI 8320.07, "Implementing the Sharing of Data, Information, and Information Technology (IT) Services in the Department of Defense," 3 August 2015. [Online]. Available: <http://www.dtic.mil/whs/directives/corres/832007p.pdf>.
- [6] DoD CIO. U. S. Department of Defense Guide 8320.02., "Guidance for Implementing Net-Centric Data Sharing," 12 April 2006.
- [7] U. S. DoD, "Manual for the Operation of the Joint Capabilities Integration and Development System," 12 February 2015.
- [8] J. G. March and R. and Weissinger-Baylon, *Ambiguity and Control: Organization Perspective on Military Decision Making*, Pitman Publishing, 1986.
- [9] Rec. ITU-T P.10/G.100 Amendment 1 (01/07): New Appendix I, Definition of Quality of Experience (QoE), Geneva: Int. Telecomm. Union, 2007.
- [10] D. G. Horvitz and D. J. Thompson, "A generalization of sampling without replacement from a finite universe," *Journal of the American Statistical Association*, vol. 47, no. 260, pp. 663-685, 1952.
- [11] J. Evans and C. Filsfil, *Deploying IP and MPLS QoS For Multiservice Networks: Theory and Practice*, San Francisco, CA: Morgan Kaufmann, 2010.
- [12] M. Alreshoodi and J. Woods, "Survey on QoE/QoS correlation models for multimedia services," *arXiv preprint arXiv:1306.0221*, 2013.
- [13] Rec. ITU-T P.862.1, Mapping function for transforming P.862 raw results to MOS-LQO, Geneva: Int. telecomm. Union, 2003.
- [14] Rec. ITU-T J.144 Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference, Geneva: Int. Telecomm. Union, 2001.
- [15] MISB RP 1203.3 Video Interpretability and Quality Measurement and Prediction, Motion Imagery Standards Board, 2014.
- [16] Rec. ITU-T P.912: Subjective video quality assessment methods for recognition tasks, Geneva: Int. Telecomm. Union, 2008.
- [17] J. C. Leachtenauer, W. Malila, J. Irvine, L. Colburn and N. Salvaggio, "General image-quality equation: GIQE," *Applied Optics*, vol. 36, no. 32, pp. 8322-8328, 1997.
- [18] MISB ST 901.2 Video-National Imagery Interpretability Rating Scale, Motion Imagery Standards Board, 2014.
- [19] R. Serral-Gracia, E. Cerqueira, M. Curado, M. Yannuzzi, E. Monteiro and X. Masip-Bruin, "An overview of quality of experience measurement challenges for video applications in IP networks," *Wired/Wireless Internet Communications*, pp. 252-263, 2010.
- [20] J. Zhang and N. Ansari, "On assuring end-to-end QoE in next generation networks: challenges and a possible solution," *Communications Magazine, IEEE*, vol. 49, no. 7, pp. 185-191, 7 2011.
- [21] M. Fiedler, T. Hossfeld and P. Tran-Gia, "A generic quantitative relationship between quality of experience and quality of service," *Network, IEEE*, vol. 24, no. 2, pp. 36-41, 2010.

- [22] T. Hastie, Tibshirani and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2009.
- [23] R. Hayes-Roth, "Model-based Communication Networks and VIRT: Filtering Information by Value to Improve Collaborative Decision-Making," in *10th International Command and Control Research and Technology Symposium*, Monterey, CA: Naval Post Graduate School, 2005.
- [24] M. Maschler, E. Solan and S. Zamir, in *Game Theory*, Cambridge University Press, 2013, pp. 176-180.
- [25] A. Wald, Statistical decision functions, Oxford: Wiley, 1950.