12th International Conference on Information Fusion
Seattle, WA, USA, July 6-9, 2009

Building Principles for a Quality of Information Specification for Sensor Information

Chatschik Bisdikian[a], **Lance M. Kaplan**[b], **Mani B. Srivastava**[c], **David J. Thornley**[d],
Dinesh Verma[a], and **Robert I. Young**[e]

[a] IBM T. J. Watson Research Center, Hawthorne, NY, USA; {*bisdik,dverma*}*@us.ibm.com*
[b] U.S. Army Research Laboratory, Adelphi, MD, USA, *lance.m.kaplan@us.army.mil*
[c] Univ. of California, Electrical Engineering Department, Los Angeles, CA, USA, *mbs@ee.ucla.edu*
[d] Imperial College London, Department of Computing, South Kensington, London, UK, *djt@doc.ic.ac.uk*
[e] Robert I. Young, DSTL, Salisbury, Wiltshire, UK, *riyoung@mail.dstl.gov.uk*

Abstract: *In the highly heterogeneous environments of coalition operations, sharing sensor-originated information with desired quality characteristics is key to the effective execution of coalition tasks. A characterization of the quality of information (QoI) is useful in many contexts and can be invaluable in making decisions such as trusting, managing, and using information in particular applications. However, the manner of representing the QoI is highly application-dependent. This leads to divergent QoI characterizations and manifestations hampering the effective and streamlined execution of coalition tasks. An application-agnostic QoI specification can provide consistency in the representation of information and its quality, and enable QoI-aware determinations across many different applications. In this paper, an application-agnostic QoI model which can be readily customized to the needs of specific applications is presented. Object-oriented modelling principles are leveraged to attain a QoI model that can be used in many different contexts.*[1]

Keywords: Quality of Information, Value of Information, QoI, QoI, VoI, QoI metadata, VoI metadata, coalition operations, sensor networks

1 Introduction

Attaining information superiority is the cornerstone of the modern *network-centric operations* (NCO) model in the execution of military operations [1][2] of allied, coalition forces. Distributed sensor systems, comprising a mix of both physical and non-physical (human) sensors, have become essential building blocks for supporting the NCO objectives. The fusion of sensor-originated data and derived information is used to support increased situational awareness and ultimately intelligent decision making and effective action taking.

Sensor systems are deployed in and around the theater of operation, forward operating bases, and so on, and facilitate the collection of relevant or *pertinent* information. These systems may have been designed and deployed for specific purposes or (as a result of their increased mobility, processing and communication capabilities) deployed in *ad hoc* manner to support sensing tasks on demand. No matter what their mode of deployment and operation is, the collection, storage, indexing, search, and dissemination of pertinent, sensor-originated information over computerized systems is crucial to situation analysts and decision makers alike to prepare for or react to situations of concern efficiently and effectively.

Sensors collect data that encode information about objects of interest. Given models about such information encoding, the data from one or multiple sensors can be fused and extract (i.e., infer) the sought after information. From the data collected by a sensor network, information about several objects from the theater of operation, e.g., the battlespace, can be extracted and situation awareness attained, as outlined by the layered JDL fusion model [3]. To achieve the desired situational awareness, this information must be disseminated to the appropriate end-user processes, e.g., human analysts. It is thus becoming important that information is summarized in some fashion, so that end-users can digest it. Furthermore, provenance becomes important so that its users can perform a deep dive into the detailed information (or sensor data) that supports the summary. The overall quality of the information that is presented to the end-user significantly impacts the ability of the user to take appropriate action.

As a simple example of the impact that the quality of the provided information could have on its end-use, consider a deployment of an array of acoustic sensors. If this array were to support an intruder detection functionality then it might need only to operate at a low sampling rate, say 1 Hz or less. However, if the array were to support intruder identification as well, the array might need to reproduce the noise spectrum of the intruder to facilitate necessary pattern matching, thus, requiring it to operate at a much higher (Nyquist) sampling rate.

---

[1] Research was sponsored by US Army Research laboratory and the UK Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the US Army Research Laboratory, the U.S. Government, the UK Ministry of Defense, or the UK Government. The US and UK Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE **JUL 2009**	2. REPORT TYPE	3. DATES COVERED **06-07-2009 to 09-07-2009**

4. TITLE AND SUBTITLE

Building Principles for a Quality of Information Specification for Sensor Information

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

IBM T.J. Watson Center, , ,Hawthrone,NY

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR’S ACRONYM(S)

11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

See also ADM002299. Presented at the International Conference on Information Fusion (12th) (Fusion 2009). Held in Seattle, Washington, on 6-9 July 2009. U.S. Government or Federal Rights License.

14. ABSTRACT

In the highly heterogeneous environments of coalition operations, sharing sensor-originated information with desired quality characteristics is key to the effective execution of coalition tasks. A characterization of the quality of information (QoI) is useful in many contexts and can be invaluable in making decisions such as trusting managing, and using information in particular applications. However, the manner of representing the QoI is highly application-dependent. This leads to divergent QoI characterizations and manifestations hampering the effective and streamlined execution of coalition tasks. An application-agnostic QoI specification can provide consistency in the representation of information and its quality and enable QoI-aware determinations across many different applications. In this paper, an application-agnostic QoI model which can be readily customized to the needs of specific applications is presented. Object-oriented modelling principles are leveraged to attain a QoI model that can be used in many different contexts.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Public Release**	**8**	

In coalition environments, components of the end-to-end information path from the sources to the users including the sensor, communication, and information processing systems may belong to a variety of organizations and coalition members. Hence, for effective coalition operations in support of the accelerated operational tempo expected by NCO operations [2], information collected and managed by one coalition member should be searchable and retrievable (in original or modified form) by another. To achieve this goal, it is necessary that a common, yet flexibly expandable, basis for searching and retrieving information must be established across coalition members. This common basis must be easily communicable (i.e., be interpretable) to and supportable (i.e., be transformable) by information processors (e.g., fusion operators) owned and execute by the various coalition members. With increasing amount of sensor-originated information, to support effective searching for pertinent information, this common basis must facilitate information enrichment using metadata to hasten the search rate for pertinent information and, hence, increase the capacity of the information assessment process. Acknowledging the importance that quality of information plays in achieving desired results (situation awareness, decision making, action taking, etc.), it is a central premise of our research that *quality of information* (QoI) can serve key role in establishing the aforementioned common basis for the purpose of indexing, searching, and exchanging pertinent information between coalition members.

The area of QoI has been extensively studied, typically under the term *information quality* (IQ), in the enterprise (for information collected and stored in data warehouses) and on-line (for information searched through Web-engines) areas [4][5][6]. This paper considers sensor-originated information in the context of military situations. Such information introduces new challenges for the quality of information required to support mission-critical situation awareness and decision making at various scope levels that is persistently updated according to the dynamic and unpredictable nature of coalition operations. The contributions in this paper are: (*a*) the introduction of a QoI framework based on a separation of responsibilities principle between quality and value of information; (*b*) the introduction of a collection of information processor dealing with generation, alteration, and communication of QoI metadata; and (*c*) an application-independent UML-based data model for the QoI metadata.

This paper is organized as follows: In section 2, we introduce QoI in the context of our research, and, in section 3, we propose a QoI framework based on a quality vs. value of information split, along with related vocabulary, and an abstract representation of a usage model of sensor-originated data and derived information. In section 4, we provide the overview of the approach and rationale for the use of the UML representation for modelling QoI and then present our QoI metadata data model. In section 5, we provide an example of QoI-related metadata for a shooter localization application in. We close with a summary in section 6.

2 Quality of information

Information is gathered to build knowledge and, hence, gain an understanding of parts of the real world that are of interest so that appropriate actions can be decided upon. The time horizon for building this understanding may be long (weeks, months, and years) as when studying the purchasing habits of consumers, or the migratory trends of population, or short (minutes, seconds, or subseconds) as when tracking incoming missiles or localizing a shooter's location.

The degree of understanding of the world depends on the degree of *pertinence* of the information gathered to the situation at hand –it depends on the quality of the gathered information. QoI for dynamic, sensor-originated information relates to the ability of using information to draw on, place on, and annotate a map of situation elements of interest accurately and quickly. Examples of such ability include placing the location of a shooter on a geographical map relative to my (the end-user) surroundings, annotating the shooter's location with a foe or friend information, drawing the trajectory of his shooting relative to my position, doing so within 3 seconds of the shooting, and so on.

To enrich the "situational" map with the aforementioned information, data from a number of sensors (acoustic, seismic, infrared, imaging, etc.) may be dynamically summoned and fused. Clearly, the accuracy, speed, precision, and so on, of our ability to enrich the situational map depends on many factors related to the gathering, transport, management, and processing of pertinent information. Unfortunately, not all of these factors are under our complete or even partial control.

The variety of sensor information available to information consumers is important, useful or appropriate at different levels for their particular tasks. For example, a video feed of a track moving through a neighborhood is irrelevant to a shooter localization task but may be relevant to a general surveillance task. Given that there is invariably more information to process than the resources available to process them, one needs to prioritize among the different possible sensor streams for any type of application.

Quality of information is the basis on which the prioritization means of information can be built. A measure associated with the sensor information stream that can measure its quality would be valuable in many such decisions. We can see evidence of this in the effectiveness and efficiency in enterprises that have benefitted from a mature research and implementation base dealing with information quality [7], and this base moves quickly with the times, taking into account the limitations on shared understanding between different parts of a large organization [8][9]. Approaches to formalize the understanding of interplay between disparate information sources are necessary for some level of guarantee of generalization and future-proofing of hard-won implementations of such approaches.

For specific application contexts, there exist quality descriptions and metrics available, such as the Civil NIIRS reference guide for imagery quality [11], the SIAP methodology for evaluating "…the quality of air vehicle portion of the Single Integrated Air Picture (SIAP)…" [12], and the EuroRoadS program for intelligent transportation systems [13]. However, there is no general framework or a common model for QoI which can be used for dealing with QoI especially for dynamic sensor-generated data in a generic manner. There are many benefits in having a common model for characterizing the QoI of sensor information streams. A common model can provide coherency among different types of information streams. Although there will be variations between the quality metrics and models that will be dependent on the specific application context of the information, a common model can provide a common ground for many of the metrics that can be shared. Furthermore, a common quality of information model can provide the context for an application-specific QoI model, and provide a base model from which application-specific models can be developed. Last but by no means least, coalition-based operations involve collections of authorities, processes, doctrines, communication protocols, and so on, that dynamically get together for the execution of tasks of common interest. Establishing common principles and policies for exchanging and sharing not only information but quality metadata about this information is undoubtedly an important facilitator for effective collaboration between the coalition forces.

3 QoI framework

3.1 The QoI/VoI split

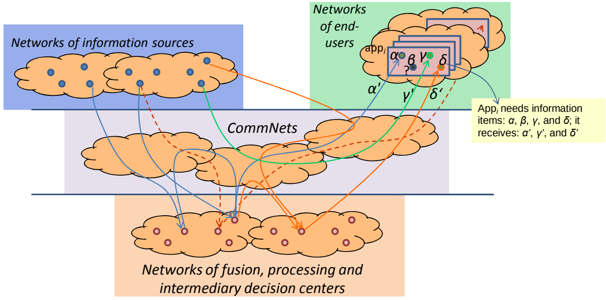

Figure 1: The end-to-end information flow path

Traditionally, sensors are deployed in conjunction with their applications –a *monolithic* relationship. However, the on-demand, just-in-time deployment of sensor-enabled applications expected during NCO coalition operations requires creation of elaborate sensor-based services with on-the-fly, dynamic binding capabilities to available sensor, information management, and communication resources. Figure 1 shows such an example, where end-user applications retrieve needed information from a collection of information sources by engaging any number of communication networks (which may include *ad hoc*, mobile networks that are deployed for missions not specifically designed for) and information processing centers. The example shows an application $\texttt{app}_i$ having information needs for items (α,β,γ,δ) and retrieving only items (α',β',δ') possessing desirable quality characteristics.

The example in Figure 1 can be abstracted to the generic scenario shown in Figure 2, which will serve as our basis for our QoI framework. Specifically, information originates from information sources, which may include sensors of various types and capabilities, as well as human intelligence. The information from a source passes through a channel to reach the receiver of the information. The receiver of the information could be a human analyst or an automated software module performing operations such as information fusion or using the received information for applications such as target tracking, intruder detection or military planning. In general, sources and receivers are functional entities residing in nodes accessible over interconnected physical networks. A node can be a receiver for one information stream, while a source for another. Information transport channels daisy-chain source and receiver functions linking collections of *end-sources* to collections of *end-receivers*.

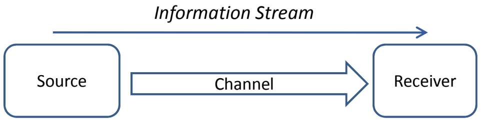

Figure 2: Abstract use scenario for the QoI framework

The quality of information will ultimately reflect upon its end-use. However, the unpredictable variety of end-uses of a piece of information and role that a given piece of information could play in the context of these end-uses presents us with a compelling challenge. To address this issue, we have elected to split the holistic QoI characterization into one that relates to the inherent properties of information, and one that relates to the role of this piece of information in the context of its end-use. We refer to the former as QoI and the later as VoI (for *value of information*), see Figure 3.

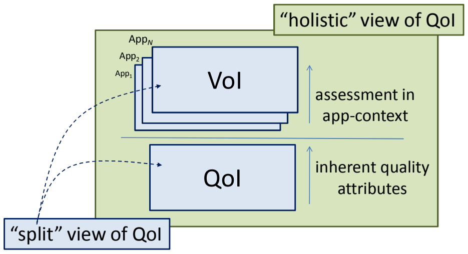

Figure 3: The QoI/VoI "split"

The QoI/VoI split allows us to separate the ability to support assessment of the information, i.e., enabling the process which ends with the passing of a *judgment* or a *verdict* about the information at hand, from the actual judgment reached within the context of particular application. For example, we separate the fact that an image has, say, resolution of 0.5 megapixels from whether this is sufficient resolution for a given application. With such a split of the entire QoI premise in mind, the following terms will be used:

- *Quality of information* (QoI) represents the body of evidence (described by information quality *attributes*) used to make judgments about the *fitness* (or, utility) of the information contained in an information stream.
- *Value of information* (VoI) represents the *utility* of the information in an information stream when used in the specific application context of the receiver.

VoI utility is expressed in an application-specific manner. With regard to the information channel:

- *Quality of service* (QoS) is a characterization of the transport/transmission properties of the channel used by the source to send information to the receiver.

QoI attributes capture the inherent characteristics of information that are independent of the specific application context in which the receiver will use the information. These attributes include such information characteristics as *accuracy*, *latency*, and *provenance* (or *integrity*). Since some of these inherent characteristics could be impacted by the transport of information from a source to a receiver, we can distinguish between QoI *emitted* by a source and QoI *delivered* to a receiver with the channel QoS linking the two. This is shown in Figure 4 which is derived from the abstract scenario in Figure 2 with added emphasis that information streams flow (and combine) from multiple sources to multiple receivers over multiple channels.

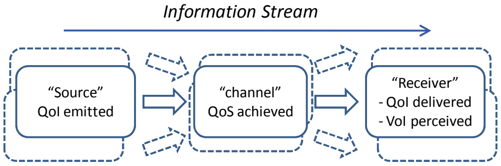

Figure 4: The QoI/QoS/VoI chaining

3.2 The QoI/VoI information processing operators

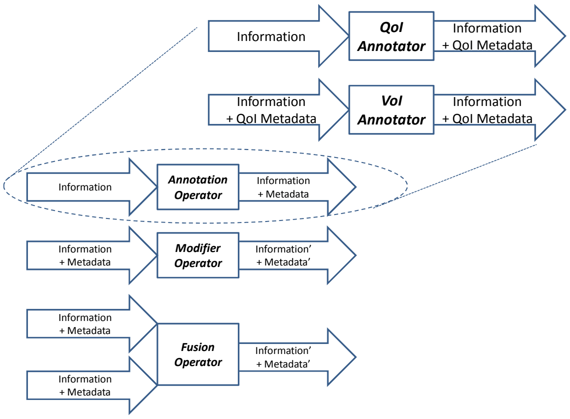

Figure 5: Information processing operators

As information flows toward the receiver of our abstract use scenario, it will be processed through several information processing operators prior to been delivered to its next-user (the end-user, or the next receiver in the end-to-end information chain). We posit the existence of the following three types of such information processing operators, see Figure 5:

- *Annotation operators*, which attach new metadata to an information stream;
- *Modifier operators*, which take an input information stream, with or without metadata and modify either the output information stream or the metadata; and
- *Fusion operators*, which combine two or more information streams.

Depending on the sensor deployment and system design choices, these operators may or may not physically reside and operate within each information receiver or their output is passed from one receiver to the next. For example consider a shooter localization case where *direction of arrival* (DOA) data collected from an acoustic array may pass through a number of receivers prior to being presented as localization information to end-users. DOA metadata may include the error estimate for DOA, the location of the acoustic array, the time of measurements, the ambient temperature, and so on. Information about DOA error may be pre-calculated and stored remotely of the receivers and retrieved upon request as necessary at any stage of the end-to-end path between the sensors and the end-user. Certain metadata may not be available or computable at all or only to a limited degree, e.g., as information crosses coalition domain boundaries, provenance metadata may be constrained for certain pieces of information–in other words, they may be data quality concerns regarding the QoI metadata too!

Of particular interest in this paper are two specific types of annotation operators, the *QoI annotator* and the *VoI annotator*. The QoI annotator takes an incoming input stream and attaches a QoI metadata to the information stream. QoI metadata are metadata attached to a piece of information that describes the quality of information in that piece. The VoI annotator takes a stream with QoI metadata and attaches VoI metadata to the stream. The VoI metadata that are metadata attached to a piece of information describe the value of the information of that piece for the receiver. The receiver uses a combination of these basic information processing operators to perform its functions. Different types of receivers may have different QoI and VoI annotators. However, these annotators will be specializations of a generic QoI and VoI annotators that comply with the QoI and VoI base models described shortly.

For the sensor information itself, we assume that the information characterizes the world of interest that is represented by collection of interrelated objects possessing states in a multidimensional space (e.g., the current location, the direction of movement, the strength, the friend/foe allegiance, etc., of a troop formation). Each piece of information pertains to some range of parameters in this multidimensional space. In some sense QoI relates to how good (accurate, timely, etc.) these parameter estimates are and VoI represents how desirable these parameter estimates are –how desirable the parameters are in the first place, and how acceptable the goodness of these estimates are for the task at hand.

4 QoI metadata model

It is a key objective of our work to develop a structure for expressing and communicating QoI and VoI metadata in an application-agnostic manner. While we realize that the true value of QoI and VoI metadata representation can only be obtained in the context of a specific application, there are many commonalities in the metadata representation across multiple applications, and a common model serves to call out those similarities as well as provide the base from which customizations can be made. While different application contexts have their own instances of metadata that are appropriate, it is our assertion that a large variety of such QoI metadata have common characteristics which can be specified in a base metadata.

In order to specify the QoI and VoI metadata without getting embroiled in the intricacies of defining a metadata format, we have opted to represent the metadata by means of an abstract data model. The data model is represented using the *unified modeling language* (UML) [10]. UML is a method to graphically represent object-oriented data. The structure of the data (types of data, specific attributes) is described visually as a set of classes or objects with shared attributes. Links between the different classes represent relationships such as aggregation, association and inheritance.

UML has proven an effective means for supporting both the communication and sharing of research results in measurement and improvement of information quality [14]. The structures studied in this manner have been used to reason about information quality in tactical systems. For example, in [15] the authors achieve a remarkable clarity of expression of the positive and negative influences on risk in the quality of information in an abstraction of a tactical system based on rates of information provision of various types. The quality of that information is measured in a subset of the essential quality attributes (correctness, consistency, etc.) [4] chosen to fully specify the behaviours of pertinent (time-evolving) information products produced in NCO situations.

4.1 The base QoI data model

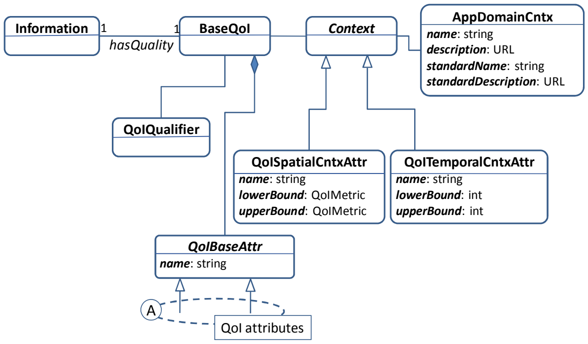

Figure 6: The base QoI data model

The base QoI data model that is shared by all application contexts is shown in Figure 6. The base QoI model comprises a collection of UML classes, which are grouped together into a set of common classes to provide a structure for the model. The specification of all of the classes in the base model provides the metadata representing QoI for a piece of information.

The `BaseQoI` class represents the base class for the QoI metadata. The `BaseQoI` class is associated with a piece of information, which is represented by the `Information` class. The `BaseQoI` class is also associated with a context, which is represented by the `Context` class. The `Context` class is an abstract representation for a collection of several context attribute classes, which includes the physical context of spatial (`QoISpatialCntxAttr`) and temporal (`QoITemporalCntxAttr`) attributes. `QoISpatialCntxAttr` contains the spatial bounds (geographical horizon) for which the QoI metadata are valid for the information. Likewise, `QoITemporalCntxAttr` contains the temporal bounds (time horizon) for which the QoI metadata are valid for the information. These two physical context attributes support the `when` and `where` *relevancy* primitives of the QoI-inspired `5WH` information summarization principle [16]. In addition to the physical context, there is also an application context class (`AppDomainCntx`) that associates the information and its QoI metadata with the specific application domain (`AppDomainCntx`) for which the QoI valid. Any specific instance of QoI metadata would have an instance of `AppDomainCntx` describing the context for which the QoI is being defined. Finally, the `QoIQualifier` class contains metadata about a particular instance of a QoI data model, including its author, time of authoring, publisher, and so on.

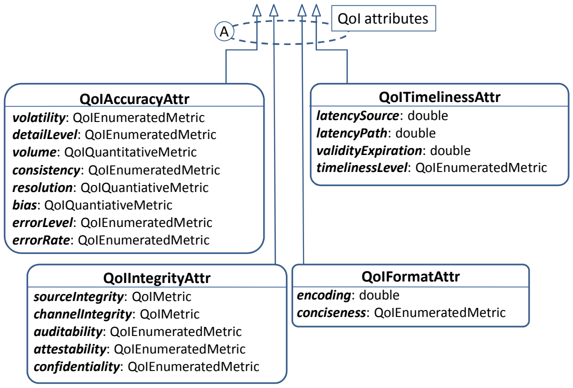

Figure 7: The QoI attribute classes

The `QoIBaseAttr` class represents an abstract placeholder for a collection of QoI attribute classes grouping related QoI attributes. Figure 7 depicts four attribute classes representing the QoI attributes for: accuracy (`QoIAccuracyAttr`), timeliness (`QoITimelinessAttr`), integrity (`QoIIntegrityAttr`), and format (`QoIFormatAttr`). The `QoIFormatAttr` class can be viewed as a representation of the quality of data, which measures quality related to the formatting of the information as data. Each of these classes comprises a collection of pertinent QoI attribute class parameters, such as those shown in the figure. The collection of attribute classes and associated parameters in the figure is not meant to be exhaustive. They can both be augmented or subclassed depending on the needs of the application domain specified by the `AppDomainCntx` class. Also, class parameter may (or may not) be used in a specific application context.

4.2 The QoI metric data model

Most of the class parameters in the QoI attribute classes are of type `QoIMetric`. In strict UML terminology, these should be shown as associations to a class `QoIMetric`. However, for our own application, it is more concise and intuitive to deviate from strict conventions and assume that a base class of type `QoIMetric` is available to characterize individual attribute class parameters. `QoIMetric` is a general class that stands for either a quantitatively computed number or an enumerated value, see Figure 8. The enumeration may rank specific attributes in discrete value. As an example, one may measure the *sourceIntegrity* parameter of `QoIIntegrityAttr` on a scale of 1 to 10 signifying increasing level of integrity. Alternatively, one many define *sourceIntegrity* as belonging to one of four types: untrusted, erratic, trusted, or authoritative. The set of allowed values depends on the context of specific application domain.

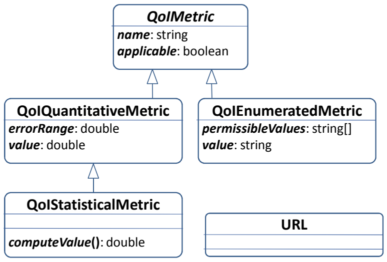

Figure 8: The QoI metric data model

4.3 The base VoI data model

In a manner analogous to QoI, we have also defined a base model for VoI, see Figure 9. The UML diagram in the base model has several similarities to that of the QoI base model: the `Context` class plays similar role to that in QoI and the `BaseVoI` class contains a collection of VoI attributes classes subclassed from the `VoIBaseAttr` class. Since, QoI attributes are used to assess VoI, the `BaseVoI` class is also associated with the corresponding `BaseQoI` class of the particular piece of `Information`.

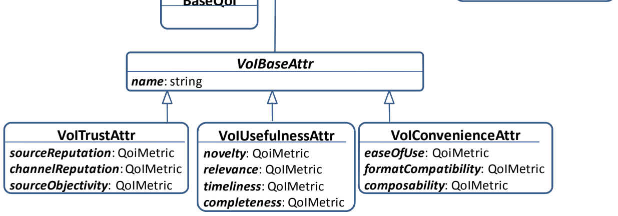

Figure 9: The base VoI data model

The `VoIBaseAttr` represents an abstract placeholder for a collection of VoI attribute classes grouping related VoI attributes. Figure 9 shows three such attribute groupings related to information trust, usefulness and convenience of using the information by the receiver. As in the case with the QoI attributes, the collection of VoI attribute classes and associated parameters in the figure is not meant to be exhaustive.

The `VoITrustAttr` class comprises the reputation of the source of the information, the channel through which the information arrives, and the objectivity of the source all as perceived by the receiver. The objectivity of the information is a measure of the fidelity with which the receiver expects the source to be reporting the information.

The `VoIUsefulnessAttr` class captures the usefulness of information in a specific context as determined by the receiver. The usefulness is assessed along four attributes, one indicating the level of novelty of the information received, a second measuring whether the information achieved is relevant for the needs of the receiver, a third expressing how timely the information is for the purpose of the receiver, and a fourth expressing the level of completeness of the information. The completeness of the information measures the degree by which the information at hand covers all that is needed and sought for by the receiver. This parameter may accept both quantitative or enumerated values, novelty may be an enumerated attribute with values: redundant, corroborative, incremental, new, or surprising!

The `VoIConvenienceAttr` class captures how easy or difficult it is for the receiver to use the information and is assessed along three attributes. The *easeOfUse* attribute assesses whether the information is perceived to be easy to use by the user. The format of the information, whether it is readily usable by the systems of the receiver or requires manipulation is assessed by the *formatCompatibility* attribute. Finally, *composability* measures whether the data can be easily composed with other pieces of information available to the user. The *composability* may be affected by the ability of the user to process the elements in its electronic system.

Taken together, these attribute classes provide an assessment of how much value a piece of information delivers to the user of the information serviced by the particular receiver.

5 A QoI metadata example

In this section, we present an example of an application and show pertinent QoI and VoI attributes.

5.1 The application: Shooter localization

We consider a sensor system comprising a collection of acoustic arrays (the system nodes) deployed to support a shooter localization application. Each node can localize a target by measuring the DOAs of the muzzle blast and shockwave and the *time difference of arrival* (TDOA) between the muzzle blast and shockwave, for further details see [17]. The application produces an *information product* (i.e., its output) comprises two numerical parameters representing the state of the shooter: (*a*) shooter location, and (*b*) direction of shooting (both expressed relative to a reference coordinate system). Optionally, the information product may also include the type of gun used. The value of this product is assessed based on the accuracy in estimating the elements of the shooter's state.

The information product is produced by processing (fusing) measurements from multiple nodes using a localization algorithm. To be able to assess the value of the product, the following must also be known about the measurements: (*a*) errors; (*b*) provenance, e.g., the location of the acoustic arrays; (*c*) timestamps for the DOA measurements; and (*d*) a classification of the raw data used to compute the DOA, e.g., gun type based on gunfire signatures. The last two are useful to properly correlate reports from multiple nodes to the same shooting event, or the same shooter.

Next we summarize the pertinent metadata to support the VoI assessment of the information provided by this application. To specialize the UML base model to this specific instance, we would define extension classes in the UML model to define additional information to augment the base QoI and VoI model.

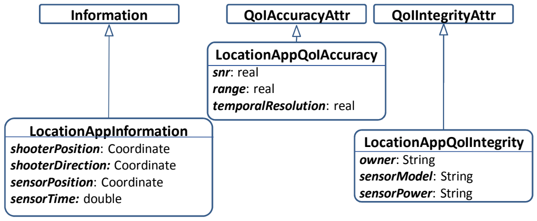

Figure 10: The QoI classes for shooter application

5.2 The quality metadata

In the shooter location example, we will subclass two of the classes (`QoIAccuracyAttr` and `QoIIntegrityAttr`) of the base QoI metadata UML model. The subclasses provide additional information that can be used to provide details for QoI attributes defined in the base case. These extensions for the QoI classes are shown in Figure 10. The figure also shows the model for a single unit of information provided by the sensors in this case. Each unit of information contains the location and direction of the sensor, the position of the sensor, and the time when the measurement was taken.

QoI-related metadata

Extension of Accuracy Attributes

To calculate the QoI attributes related to accuracy, three additional pieces of information are added to the accuracy attribute in the extended class `LocationAppQoIAccuracy`, which is a subclass of `QoIAccuracyAttr`. This information contains the signal-to-noise ratio (SNR), range, and temporal resolution of the sensor. This information could be found in product literature and retrieved on demand by providing a pointer to it.

Extension of Integrity Attributes

The additional information needed to estimate the integrity of the information is the owner of the sensor, because, for example, assets belonging to other partners in a coalition may have a lower integrity; the model of the sensor; and the power level at which the sensor was operating when the measurements were made. Different levels of integrity can be assigned to sensors operating at different levels of power.

On the basis of both these extensions, the different attributes of the QoI of any unit of information can be computed. The QoI metadata fields can be computed and assigned different levels based on the details of the different metrics.

Algorithm requirements (sensor QoI outputs)

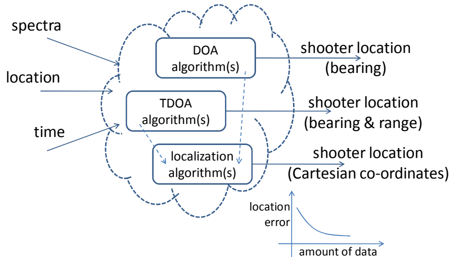

Figure 11: The location estimation algorithms

Figure 11 shows the collection of estimation algorithms for shooter location that could be used, the inputs that they use and the outputs that they produce. As the figure shows, the localization algorithm will typically operate on the combination of the outputs from the DOA- and TDOA-based algorithms prior to producing its own output. Clearly the accuracy of the output results depends on the accuracy of the inputs, their amount, and the characteristics of the specific algorithms employed. Thus, in this case, a QoI annotator operator, see Figure 5, will operate on the incoming information stream (the various measurements) and their QoI metadata (e.g., the *errorLevel* in `QoIAccuracyAttr`). It will produce an outgoing information stream (the information product related to the shooter's state, i.e., his location and direction) each with its own *errorLevel* parameter that is calculated from the corresponding input parameters and capabilities of the location estimation algorithms.

VoI-related metadata

In order to determine the VoI in this example, we do not need to extend the base UML classes, but can use them as is. The value of the particular information product can be assessed along various dimensions captured by the VoI attributes, see Figure 9:

VoI Trust: This is assessed based on the trust level of various entities such as the sources (e.g., the reputation of the sensors, nodes), their owners (e.g., various coalition members), and knowledge of the algorithms used (which impacts the trust on the reported information product and QoI metadata). This can be measured in a 3-point scale of Low, Medium and High.

VoI usefulness: This is assessed based on the proximity to desired values of the reported mean square error (MSE) of the location, shooting orientation, and time of event; the delay in reporting the event (timeliness); and the completeness of the coverage of the reports, e.g., the parentage of spatiotemporal coverage provided by the sensor nodes.

VoI convenience: This is assessed with respect to "compatibility" of information product to the expectations of its recipient. Increasing the VoI convenience may require the syntactic transformation of the information product whenever possible and necessary, e.g., translate between measurement units when the product is provided to other coalition members, or provide a composite product such as "wind-chill factor" from constituent estimates of "ambient temperature" and "wind speed."

It should be noted that coalition members may prefer not to divulge, for example, the exact location or capabilities of their own sensors to certain other members by altering the entries of pertinent Such deliberate modifications of metadata form the basis for risk-based distribution of military information between coalition members [18].

6 Summary

Effective information sharing is a necessary condition for effective coalition operations supporting the NCO objectives. The effective sharing of ever increasing amounts, types, shapes, and forms of dynamically produced, sensor-originated information among coalition members requires a common, flexible indexing, searching, and retrieval means to be established. Motivated by the importance that quality information plays in improved situation awareness, effective decision making and action taking, in this paper we have proposed using QoI and QoI-related metadata for supporting the above.

To cope with the broad reach of QoI, we have proposed its split it into: (*a*) a QoI part that relates to the inherent quality properties of sensor-originated information (including accuracy, latency, and provenance); and (*b*) a VoI part that relates to the value of the information within the context of an application. Based on this split, we have introduced an abstract UML-based data model for QoI and VoI metadata. The data model provides a general template for organizing QoI and VoI metadata in support of QoI-centered annotation operations of information. Using this template, instances of the model can be produced for specific application contexts in a consistent, reproducible, and repeatable manner.

The proposed data model represents only a first step toward QoI-aware NCO operations. It codifies the information that is to be exchanged between QoI-aware systems. To build and operate these systems, though, requires establishing a trusted, auditable infrastructure for communicating the QoI medatada and architecting and building the intelligence for processing them, like the annotators in Figure 5. This is not a small feat and leaves plenty a research questions for future considerations. The aforementioned secure metadata frameworks [18] and policy-based coalition operations [19] are promising paths to consider.

References

[1] D. S. Alberts, J. J. Garstka, F. P. Stein, *Network Centric Warfare, 2nd ed.*, CCRP, 2000.

[2] *The Implementation of Network-Centric Warfare*, US. DoD, Office of Force Transformation, Jan. 2005.

[3] D. L. Hall and S. A. H. McMullen, *Mathematical Techniques in Multisensor Data Fusion, 2nd ed.*, Artech House, 2004.

[4] R. Y. Wang and D. M. Strong, "Beyond accuracy: What data quality means to data consumers," *J. Management Information Systems*, 12, 4; ABI/INFORM Global pg. 5, Spring 1996.

[5] F. Naumann and K.-U. Sattler, "Information Quality: Fundamentals, Techniques, and Use," *EDBT Tutorial*, Munich, March 28, 2006.

[6] S.-A. Knight, "User perceptions of information quality in World Wide Web information retrieval behavior," PhD Thesis, *Edith Cowan University*, Nov. 2007.

[7] The MIT *Total Data Quality Management* (*TDQM*) Program, http://web.mit.edu/tdqm/www/index.shtml.

[8] P. Bertolazzi and M. Scannapieco, "Introducing data quality in a cooperative context," *6th Int'l Conf. on Information Quality* (*IQ-2001*), MIT, Boston, MA, USA, Nov. 2-4, 2001.

[9] M. Mecella, M. Scannapieco, A. Virgillito, R. Baldoni I, T. Catarci, and C. Batini, "Managing data quality in cooperative information systems," *10th Int'l Conf. Cooperative Information Systems* (*CoopIS*), Irvine, CA, USA, Oct. 30-Nov. 1, 2002.

[10] Open Management Group (OMG), *Unified Modeling Language* (UML), http://www.uml.org/.

[11] *National Imagery Interpretability Rating Scale* (NIIRS), Imagery Resolution Assessments and Reporting Standards (IRARS) Committee, http://www.fas.org/irp/imint/niirs.htm.

[12] *Single Integrated Air Picture (SIAP) Attributes*, ver. 2.0, SIAP System Engineering Task Force (SE TE), Technical Report 2003-029, Aug. 2003.

[13] T. Wiltschko and T. Kaufmann, "A quality model for quality assurance of road information," *5th European Congress and Exhibition on Intelligent Transportation Systems and Services*, Hannover, Germany, June 1-3, 2005.

[14] M. Scannapieco, B. Pernici, E. Pierce, "IP-UML: Towards a Methodology for Quality Improvement Based on the IP-MAP Framework," *7th Int'l Conf. on Information Quality* (*ICIQ-02*), MIT, Boston, MA, USA, Nov. 8-10, 2002.

[15] P. J. Driscoll, M. Tortorella, and E. Pohl, "Information Product Quality in Network Centric Operations," Technical Report DSE-TR-0516, DTIC#: ADA435046, *US Military Academy at West Point*, New York, May 2005.

[16] C. Bisdikian, J. Branch, K. K. Leung, and R. I. Young, "A Letter Soup for the Quality of Information in Sensor Networks," *1st Int'l Workshop on Information Quality and Quality of Service for Pervasive Computing* (*IQ2S 2009*), –part of *IEEE MASS 2008*–, Galveston, Texas, March 9, 2009.

[17] L. M. Kaplan, T. Damarla, and T. Pham, "QoI for Passive Acoustic Gunfire Localization," *1st Int'l Workshop on Quality of Information for Sensor Networks* (*QoISN 2008*) –part of *IEEE MASS 2008*–, Atlanta, GA, USA, Sept. 29, 2008.

[18] M. Srivatsa, P. Rohatgi, S. Balfe, and S. Reidt, "Securing Information Flows: A Metadata Model," *1st Int'l Workshop on Quality of Information for Sensor Networks* (*QoISN 2008*) –part of *IEEE MASS 2008*–, Atlanta, GA, USA, Sept. 29, 2008.

[19] A. Schaeffer-Filho, E. Lupu, M. Sloma, S.-L. Keoh, J. Lobo, and Seraphin Calo, "A Role-Based Infrastructure for the Management of Dynamic Communities," *2nd Int'l Conf. Autonomous Infrastructure, Management and Security: Resilient Networks and Services*, Bremen, Germany, July 01 - 03, 2008.