**NIPWG3-30.1**

## Paper for consideration by NIPWG

## Data Quality – Fuzzy Areas – Proposed Types

## 10 August 2016

## Proposal Type

|  |  |
| --- | --- |
| **Type of Change Requested** | **Mark All that Apply** |
| NPUB (S-123 etc.) |  |
| New/Amended Feature | X |
| New/Amended Complex Attribute |  |
| New/Amended Simple Attribute | X |
| New/Amended Information Type |  |
| New/Amended Association/Aggregation/Composition | X |
| New/Amended Enumerate Value |  |

**Introduction - Fuzzy areas model**

The proposed extension for modeling fuzzy areas consists basically of adding a generic feature type that allows datasets to demarcate areas where the cartographer does not have complete confidence in the existence of the concept described by an associated geographic feature. The concept can be the availability of a service, the existence of a natural phenomenon, etc. The level of confidence is described by a limited set of ranked values that correspond roughly to the probability or likelihood that the service will be available, or that the phenomenon will occur.

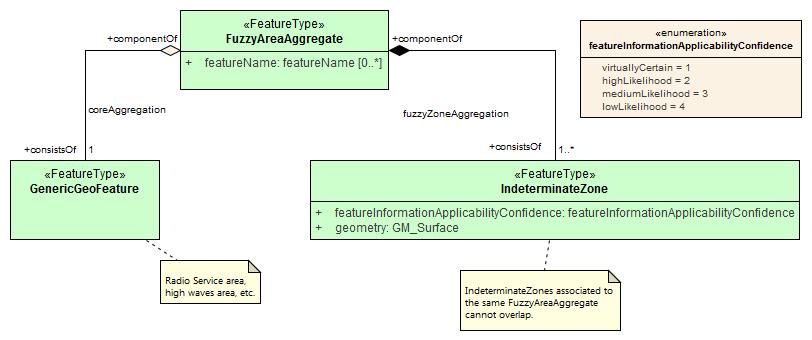
The use of a small set of ranked values for confidence intervals instead of a probability distribution, or intervals described by percentage probability thresholds, was based on the following criteria:

* it should be possible to portray the data on the screen, as well as show pick reports, in ways that mariners can understand without much training and without significant expenditure of attention during route monitoring;
* the model must be acceptable for cartographers and encoders to use, even in the absence of statistical or scientific basis for deciding the likelihood at any specific point;
* it should be consistent with current approaches to S-100 portrayal processing and lend itself to efficient portrayal processing.

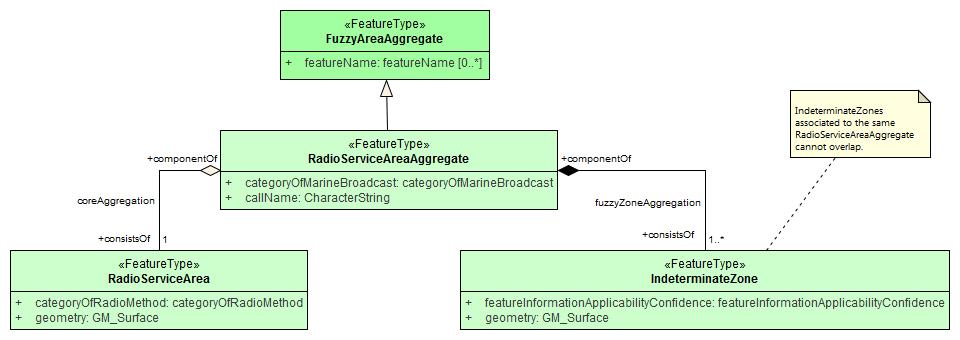
Areas of uncertainty are modeled by a new **Indeterminate Zone** geographic feature. A model of a ‘fuzzy area’ will therefore consist of a ‘core’ feature of the appropriate geographic feature type and the appropriate geometry (e.g., Radio Service Area) and one or more ‘fuzzy’ **Indeterminate Zone** features (with surface geometry). Each **Indeterminate Zone** feature has a thematic attribute expressing the level of confidence that the service described by the core feature will be available (or the natural phenomenon will occur) in the region demarcated by its spatial primitive(s). To provide for application schemas which need to model properties of the collection of core and fuzzy features (e.g., a collective name for the totality of core and fuzzy areas), an aggregation feature is also proposed, associated to the core and indeterminate zone features. This aggregation feature may be sub-classed by application schemas, and bind additional thematic attributes as needed in the domain model (e.g., to avoid duplication of attributes common to related instances).

To reduce complexity, the model provides for associations to the aggregate feature but not between the indeterminate zone and core features, because the latter would be redundant given that it is possible for applications to access core and fuzzy features via the association to the aggregate (though at the cost of an extra object access operation).

The model is depicted in the figure that follows.



The typical application schema will subclass the aggregate feature and bind attributes to the geographic feature type **RadioServiceAggregate** as depicted in the next figure, which shows a hypothetical fragment of an application schema for radio services.



*Recommendations*

1. Define the feature types, attributes, and associations listed below:

* New geographic feature type IndeterminateZone
* New geographic feature type Fuzzy Area Aggregate (aggregation feature without geometry)
* New aggregation association: Core Aggregation
* New composition association: Fuzzy Zone Aggregation

**Annex A. Definitions of the new types**

**New feature: Indeterminate Zone**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| IHO Definition: **INDETERMINATE ZONE:** A region in which there is some uncertainty about the existence of a phenomenon or the availability of a service. (IHO). | | | | | | | | | | | | | |
| **S-??? Geo Feature: Indeterminate Zone** | | | | | | | | | | | | | |
| **Primitives:** GM\_Surface | | | | | | | | | | | | | |
| *Real World* | | | *Paper Chart Symbol* | | | | | *ECDIS Symbol*  *(generally a lighter shade of the associated core area, lightness depending on the confidence attribute)*  *(boundary line style? TBD)*  *(fill for indeterminate zone when the core feature is a curve or point?)*  *(for core areas with no fill, use the line style rule?)* | | | | | |
| **S-??? Attribute** | | | | **S-57 Acronym** | | **Allowable Encoding Value** | | | | **Type** | | **Multiplicity** | |
| Feature information applicability confidence | | | |  | | 1: virtually certain  2: high likelihood  3: medium likelihood  4: low likelihood | | | | E | | 1 | |
| **Feature associations** | | | | | | | | | | | | | |
| **Type** | **Association Name** | **Association Ends** | | | | | | | | | | | |
| **Class** | | | **Role** | | **Mult.** | | **Class** | | **Role** | | **Mult.** |
| Comp | fuzzyZoneAggregation | IndeterminateZone | | | consistsOf | | 1..\* | | FuzzyAreaAggregate or its non-abstract sub-types, as defined in the application schema. | | componentOf | | 1 |
| INT 1 Reference: none  **X.X.X Indeterminate Zones (see ?-? – B-YYY.Y)**  **Indeterminate Zone** features are used with geographic features that encode services or phenomena such as hazardous conditions that are intermittent, transient, occasional, or experienced only under certain conditions.  If availability or occurrence is according to a fixed schedule or between known dates or times, the complex attributes **fixed date range** or **periodic date range** must be used instead.  Uncertainty about the position of a feature must be encoded using the appropriate uncertainty attribute (e.g., **quality of horizontal measurement**, **horizontal position uncertainty, etc.),** meta-feature (Quality **of Bathymetric Data**, Quality **of Non-Bathymetric Data**), or information type (e.g., **Spatial Quality**) as described in the S-101 or other DCEG for those features.  **X.X.X.1 Use of fuzzy regions with feature classes**  Individual product specifications must indicate whether the use of fuzzy areas with any particular feature class is mandatory, prohibited, or optional in datasets conforming to the specification.  **X.X.X.2 Requirement for core feature**  A geographic feature other than **Indeterminate Zone** is required to act as the core feature.  The geometry for the core feature should correspond to a confidence level of 99% or higher, even stronger than ‘virtually certain’. The ENC equivalent would be a confidence level high enough to avoid encoding horizontal measurements with **quality of position** = “approximate,” i.e., “third order accuracy, but within 30.5 metres of its correct geographic location” (S-101 DCEG Baseline 2 § 28.14).  If it is possible that a core feature instance may have no known geometry (i.e., there is no area, curve, or point where confidence is sufficient to encode a spatial primitive for a core feature instance), the product specification must allow core feature instances to have no spatial primitive (or area, point, etc., as appropriate). The cartographer must encode an instance of the core feature without a spatial primitive and **Indeterminate Zone** features with the appropriate confidence levels.  **Example**: A particular radio service has no area where reception is definite or virtually certain, only areas where it is medium or low quality. This must be encoded as a **RadioServiceArea** feature with no spatial primitive and two **IndeterminateZone** features with distinct area geometries and **feature information applicability confidence** = ‘*medium likelihood’* and ‘*low likelihood*’ respectively. The **RadioServiceArea** (which has in this case no geometry) and the two instances of the **IndeterminateZone** feature have a feature association to a **RadioServiceAreaAggregate** feature.  **X.X.X.3 Spatial primitives**  Application schemas may include constraints on the allowed spatial primitives, e.g., to circular areas, sectors, or rectangles.  **X.X.X.4 Statistical confidence**  If confidence values are available, or can be easily computed, the recommended confidence thresholds for determining the confidence scale values are:   * Virtually certain: Virtually certain to be experienced by (or available to) an individual vessel; will be experienced by nearly all vessels. In statistical terms, this definition will correspond to 95% and higher probability of availability of service, or the phenomenon is encountered 95% of the time, or by 95% of vessels in the area. * High likelihood: Frequently experienced by (or available to) an individual vessel; experienced by a majority of vessels. In statistical terms, this definition will correspond to 66% to 95% probability of availability of service, or the phenomenon is encountered 66% to 95% of the time or by 66% to 95% of vessels in the area. * Medium likelihood: Occasionally experienced by (or available to) an individual vessel; experienced by (or available to) about half of all vessels. In statistical terms, this definition will correspond to 33% to 66% probability of availability of service, or the phenomenon is encountered 33% to 66% of the time or by 33% to 66% of vessels in the area. * Low likelihood: Unlikely, or rarely experienced by (or available to) an individual vessel; experienced by (or available to) a minority of vessels). In statistical terms, this definition will correspond to 5% to 33% probability of availability of service, or the phenomenon is encountered 5% to 33% of the time or by 5% to 33% of vessels in the area.   **X.X.X.5 Synonyms for confidence level ranks**  General correspondences between terms in source material and the scale values for the confidence attribute are given below. Some terms are ambiguous and may correspond to multiple ranks – in this case consider the term’s ranking in relation to other terms used in the source.   * Virtually certain: assured; almost always; very high probability; very likely; usually * High likelihood: high probability; likely ; very probable; very likely; frequent; probable; often; generally * Medium likelihood: moderate likelihood / probability; neutral; occasional; as likely as not; sometimes; possible * Low likelihood: low likelihood; low probability ; possible but unlikely; infrequently; improbable; occasional; remote; rare; probably not; sometimes   Remarks:   * **Indeterminate Zone** feature instances associated to the same aggregate feature instance must not overlap. * Portrayal catalogues must define suitable presentations (area fill color, symbol, line style, etc.) for **Indeterminate Zone** areas for cases where the core feature has point, curve, or no geometry.   Distinction: none | | | | | | | | | | | | | |

**New feature: Fuzzy Area Aggregate**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| IHO Definition: **FUZZY AREA AGGREGATE:** Aggregation of a geographic feature describing a service or phenomenon with zones of different confidence about the availability of the service, occurrence of the phenomenon, or applicability of the information described by the geographic feature. (IHO). | | | | | | | | | | | | | |
| **S-??? Geo Feature: Fuzzy Area Aggregate** | | | | | | | | | | | | | |
| **Primitives:** none | | | | | | | | | | | | | |
| *Real World*  . | | | *Paper Chart Symbol* | | | | | *ECDIS Symbol* | | | | | |
| **S-??? Attribute** | | | | **S-57 Acronym** | | **Allowable Encoding Value** | | | | **Type** | | **Multiplicity** | |
| Feature name | | | |  | |  | | | | C | | 0,\* | |
| Display name | | | |  | |  | | | | S (BO) | | 0,1 | |
| Language | | | |  | | (ISO 639-3) | | | | S (TE) | | 0,1 | |
| Name | | | | (OBJNM)  (NOBJNM) | |  | | | | S (TE) | | 1,1 | |
| **Feature associations** | | | | | | | | | | | | | |
| **Type** | **Association Name** | **Association Ends** | | | | | | | | | | | |
| **Class** | | | **Role** | | **Mult.** | | **Class** | | **Role** | | **Mult.** |
| Aggr | coreAggregation | FuzzyAreaAggregate (or sub-type defined in application schema) | | | componentOf | | 1 | | (geographic feature type) | | consistsOf | | 1 |
| Comp | fuzzyZoneAggregation | FuzzyAreaAggregate (or sub-type defined in application schema) | | | componentOf | | 1 | | IndeterminateZone | | consistsOf | | 1..\* |
| INT 1 Reference: none  **X.X.X Fuzzy Area Aggregate (see ?-? – B-YYY.Y)**  Fuzzy Area Aggregates are composed of one core geographic feature and one or more indeterminate zones, and are used for linking regions where the confidence in the information described by the geographic feature varies from certainty to remote possibility. **Fuzzy area aggregate** features may also encode thematic attributes common to all the associated features.  The totality of the area where the information may apply consists of the union of the location geometries encoded in the core and indeterminate zone features.  Product specifications must indicate whether the use of fuzzy area aggregates with any particular feature class is mandatory, prohibited, or optional in datasets conforming to the specification (see X.X.X.1), but every instance of **Fuzzy Area Aggregate** must be associated to exactly one geographic feature instance and at least one **Indeterminate Zone** instance.  Remarks:  None.  Distinction: none | | | | | | | | | | | | | |

**New Attribute**

Name: **Feature information applicability confidence**

**Alternatives:**

**confidence**

**likelihood**

**presenceConfidence**

**informationApplicabilityConfidence**

**confidenceLevel**

IHO Definition: The likelihood or probability that a vessel will experience the phenomenon described by a feature, or that the service described by the feature will be available.

1. **Virtually certain**

Virtually certain to be experienced by (or available to) an individual vessel; will be experienced by nearly all vessels. (FAA, adapted.)

1. **High likelihood**

Frequently experienced by (or available to) an individual vessel; experienced by a majority of vessels. (FAA, adapted.)

1. **Medium likelihood**

Occasionally experienced by (or available to) an individual vessel; experienced by (or available to) about half of all vessels. (FAA, adapted.)

1. **Low likelihood**

Unlikely, or rarely experienced by (or available to) an individual vessel; experienced by (or available to) a minority of vessels). (FAA, adapted.)

Remarks:

No remarks

**New Associations**

|  |  |  |  |
| --- | --- | --- | --- |
| **Core Aggregation:** IHO Definition: A feature association for the binding between an aggregation feature that describes areas of varying uncertainty about a service or phenomenon and a geographic feature describing the service or phenomenon. (IHO).  Remarks:   * None | | | |
| **Role Type** | **Role** | **Features** | **Multiplicity** |
| Aggregation | componentOf | Fuzzy Area Aggregate or a sub-type of it | 1 |
| consistsOf | (geographic feature) | 1 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Fuzzy Zone Aggregation:** IHO Definition: A feature association for the binding between an aggregation feature that describes areas of varying uncertainty about a service or phenomenon and zones of uncertainty about the service or phenomenon. (IHO).  Remarks:   * None | | | |
| **Role Type** | **Role** | **Features** | **Multiplicity** |
| Composition | componentOf | Fuzzy Area Aggregate or a sub-type of it | 1 |
| consistsOf | Indeterminate Zone | 1..\* |

**Roles:**

(**component of** and **consists of** are defined in the S-101 DCEG; the definitions below will be discussed with the S-101 PT.)

**component of:**

IHO definition: A pointer to the aggregate in a whole-part relationship.

**consists of:**

IHO definition: A pointer to a part in a whole-part relationship.