             Representation of Uncertainty and Confidence in
        the Presence Information Data Format Location Object (PIDF-LO)

Abstract

   This document defines key concepts of uncertainty and confidence as
   they pertain to location information. Methods for the manipulation
   of location estimates that include uncertainty information are
   outlined.

   This document normatively updates the definition of location
   information representations defined in RFCs 4119 and 5491. It also
   deprecates related terminology defined in RFC 3693.

Status of This Memo

   This is an Internet Standards Track document.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG). Further information on
   Internet Standards is available in Section 2 of RFC 5741.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   http://www.rfc-editor.org/info/rfc7459.

Table of Contents

1. Introduction

   Location information represents an estimation of the position of a
   Target [RFC6280]. Under ideal circumstances, a location estimate
   precisely reflects the actual location of the Target. For automated
   systems that determine location, there are many factors that
   introduce errors into the measurements that are used to determine
   location estimates.

   The process by which measurements are combined to generate a location
   estimate is outside of the scope of work within the IETF. However,
   the results of such a process are carried in IETF data formats and
   protocols. This document outlines how uncertainty, and its
   associated datum, confidence, are expressed and interpreted.

   This document provides a common nomenclature for discussing
   uncertainty and confidence as they relate to location information.

   This document also provides guidance on how to manage location
   information that includes uncertainty. Methods for expanding or
   reducing uncertainty to obtain a required level of confidence are
   described. Methods for determining the probability that a Target is
   within a specified region based on its location estimate are
   described. These methods are simplified by making certain
   assumptions about the location estimate and are designed to be
   applicable to location estimates in a relatively small geographic
   area.

   A confidence extension for the Presence Information Data Format -
   Location Object (PIDF-LO) [RFC4119] is described.

   This document describes methods that can be used in combination with
   automatically determined location information. These are
   statistically based methods.

1.1. Conventions and Terminology

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document are to be interpreted as described in [RFC2119].

This document assumes a basic understanding of the principles of
mathematics, particularly statistics and geometry.

Some terminology is borrowed from [RFC3693] and [RFC6280], in
particular "Target".

Mathematical formulae are presented using the following notation: add
"+", subtract "-", multiply "*", divide "/", power "^", and absolute
value "|x|". Precedence follows established conventions: power
operations precede multiply and divide, multiply and divide precede
add and subtract, and parentheses are used to indicate operations
that are applied together. Mathematical functions are represented by
common abbreviations: square root "sqrt(x)", sine "sin(x)", cosine
"cos(x)", inverse cosine "acos(x)", tangent "tan(x)", inverse tangent
"atan(x)", two-argument inverse tangent "atan2(y,x)", error function
"erf(x)", and inverse error function "erfinv(x)".

2. A General Definition of Uncertainty

Uncertainty results from the limitations of measurement. In
measuring any observable quantity, errors from a range of sources
affect the result. Uncertainty is a quantification of what is known
about the observed quantity, either through the limitations of
measurement or through inherent variability of the quantity.

Uncertainty is most completely described by a probability
distribution. A probability distribution assigns a probability to
possible values for the quantity.

A probability distribution describing a measured quantity can be
arbitrarily complex, so it is desirable to find a simplified model.
One approach commonly taken is to reduce the probability distribution
to a confidence interval. Many alternative models are used in other
areas, but study of those is not the focus of this document.

In addition to the central estimate of the observed quantity, a
confidence interval is succinctly described by two values: an error
range and a confidence. The error range describes an interval and
the confidence describes an estimated upper bound on the probability
that a "true" value is found within the extents defined by the error.

In the following example, a measurement result for a length is shown
as a nominal value with additional information on error range (0.0043
meters) and confidence (95%).

    e.g., x = 1.00742 +/- 0.0043 meters at 95% confidence

This measurement result indicates that the value of "x" is between
1.00312 and 1.01172 meters with 95% probability. No other assertion
is made: in particular, this does not assert that x is 1.00742.

Uncertainty and confidence for location estimates can be derived in a
number of ways. This document does not attempt to enumerate the many
methods for determining uncertainty. [ISO.GUM] and [NIST.TN1297]
provide a set of general guidelines for determining and manipulating
measurement uncertainty. This document applies that general guidance
for consumers of location information.

As a statistical measure, values determined for uncertainty are found
based on information in the aggregate, across numerous individual
estimates. An individual estimate might be determined to be
"correct" -- for example, by using a survey to validate the result --
without invalidating the statistical assertion.

This understanding of estimates in the statistical sense explains why
asserting a confidence of 100%, which might seem intuitively correct,
is rarely advisable.

2.1. Uncertainty as a Probability Distribution

The Probability Density Function (PDF) that is described by
uncertainty indicates the probability that the "true" value lies at
any one point. The shape of the probability distribution can vary
depending on the method that is used to determine the result. The
two probability density functions most generally applicable to
location information are considered in this document:

o The normal PDF (also referred to as a Gaussian PDF) is used where
   a large number of small random factors contribute to errors. The
   value used for the error range in a normal PDF is related to the
   standard deviation of the distribution.

o A rectangular PDF is used where the errors are known to be
   consistent across a limited range. A rectangular PDF can occur
   where a single error source, such as a rounding error, is
   significantly larger than other errors. A rectangular PDF is
   often described by the half-width of the distribution; that is,
   half the width of the distribution.

Each of these probability density functions can be characterized by
its center point, or mean, and its width. For a normal distribution,
uncertainty and confidence together are related to the standard
deviation of the function (see Section 5.4). For a rectangular
distribution, the half-width of the distribution is used.

Figure 1 shows a normal and rectangular probability density function
with the mean (m) and standard deviation (s) labeled. The half-width
(h) of the rectangular distribution is also indicated.

```
                           *****              *** Normal PDF
                         **  :  **            --- Rectangular PDF
                        **     :    **
                        **     :     **
             .---------*---------------*---------.
             |      **        :        **        |
             |      **        :        **        |
             |    * <-- s -->:             *     |
             |     * :        :        : *       |
             |     **         :          **      |
             |    *  :        :        :   *     |
             |    *          :          *        |
             |**       :        :        :  **   |
           **          :        :           **
          *** |        :        :        | ***
         ***** |        :        :        |  *****
       .****-------+.......:.........:.........:.......+-------*****.
                           m
```

   Figure 1: Normal and Rectangular Probability Density Functions

For a given PDF, the value of the PDF describes the probability that
the "true" value is found at that point. Confidence for any given
interval is the total probability of the "true" value being in that
range, defined as the integral of the PDF over the interval.

   The probability of the "true" value falling between two points is
   found by finding the area under the curve between the points (that
   is, the integral of the curve between the points). For any given
   PDF, the area under the curve for the entire range from negative
   infinity to positive infinity is 1 or (100%). Therefore, the
   confidence over any interval of uncertainty is always less than
   100%.

Figure 2 shows how confidence is determined for a normal
distribution. The area of the shaded region gives the confidence (c)
for the interval between "m-u" and "m+u".

```
                            *****
                          **:::::**
                        **:::::::::::**
                      **:::::::::::::**
                    *:::::::::::::::::*
                   **:::::::::::::::::**
                  **:::::::::::::::::::::**
                 *:::::::::::::::::::::::::*
                *:::::::::::::::::::::::::::*
               **:::::::::::::::::::::::::::::**
              *:::::::::::: c :::::::::::::*
             *:::::::::::::::::::::::::::::::*
           **|:::::::::::::::::::::::::::::::|**
          **  |:::::::::::::::::::::::::::::|  **
        ***    |:::::::::::::::::::::::::::::|    ***
      *****     |:::::::::::::::::::::::::::::|     *****
    .****.........!:::::::::::::::::::::::::::::!..........*****.
            |                   |                 |
          (m-u)                 m               (m+u)
```
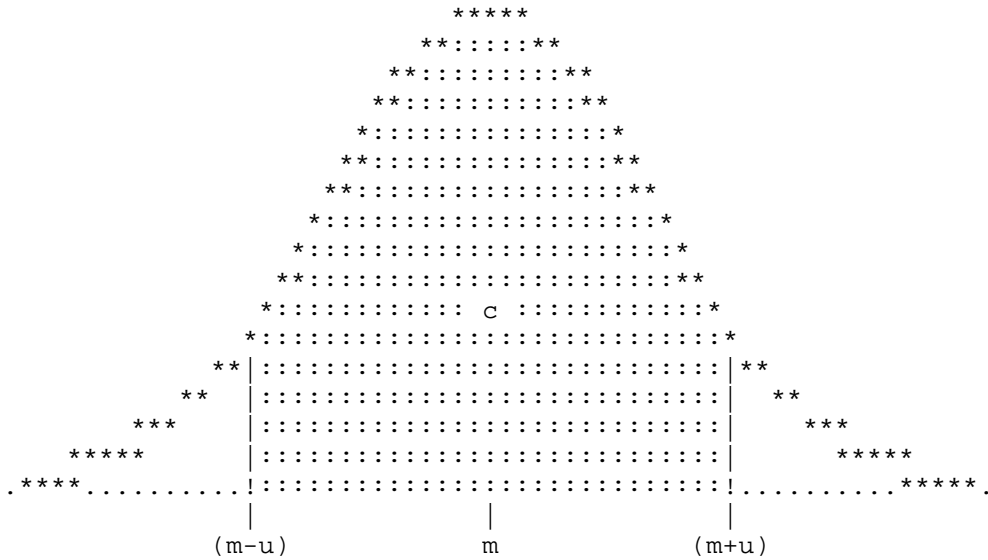
                Figure 2: Confidence as the Integral of a PDF

In Section 5.4, methods are described for manipulating uncertainty if
the shape of the PDF is known.

2.2. Deprecation of the Terms "Precision" and "Resolution"

The terms "Precision" and "Resolution" are defined in RFC 3693
[RFC3693]. These definitions were intended to provide a common
nomenclature for discussing uncertainty; however, these particular
terms have many different uses in other fields, and their definitions
are not sufficient to avoid confusion about their meaning. These
terms are unsuitable for use in relation to quantitative concepts
when discussing uncertainty and confidence in relation to location
information.

2.3. Accuracy as a Qualitative Concept

   Uncertainty is a quantitative concept. The term "accuracy" is useful
   in describing, qualitatively, the general concepts of location
   information. Accuracy is generally useful when describing
   qualitative aspects of location estimates. Accuracy is not a
   suitable term for use in a quantitative context.

   For instance, it could be appropriate to say that a location estimate
   with uncertainty "X" is more accurate than a location estimate with
   uncertainty "2X" at the same confidence. It is not appropriate to
   assign a number to "accuracy", nor is it appropriate to refer to any
   component of uncertainty or confidence as "accuracy". That is,
   saying the "accuracy" for the first location estimate is "X" would be
   an erroneous use of this term.

3. Uncertainty in Location

   A "location estimate" is the result of location determination. A
   location estimate is subject to uncertainty like any other
   observation. However, unlike a simple measure of a one dimensional
   property like length, a location estimate is specified in two or
   three dimensions.

   Uncertainty in two- or three-dimensional locations can be described
   using confidence intervals. The confidence interval for a location
   estimate in two- or three-dimensional space is expressed as a subset
   of that space. This document uses the term "region of uncertainty"
   to refer to the area or volume that describes the confidence
   interval.

   Areas or volumes that describe regions of uncertainty can be formed
   by the combination of two or three one-dimensional ranges, or more
   complex shapes could be described (for example, the shapes in
   [RFC5491]).

3.1. Targets as Points in Space

   This document makes a simplifying assumption that the Target of the
   PIDF-LO occupies just a single point in space. While this is clearly
   false in virtually all scenarios with any practical application, it
   is often a reasonable simplifying assumption to make.

   To a large extent, whether this simplification is valid depends on
   the size of the Target relative to the size of the uncertainty
   region. When locating a personal device using contemporary location
   determination techniques, the space the device occupies relative to

the uncertainty is proportionally quite small. Even where that
device is used as a proxy for a person, the proportions change
little.

This assumption is less useful as uncertainty becomes small relative
to the size of the Target of the PIDF-LO (or conversely, as
uncertainty becomes small relative to the Target). For instance,
describing the location of a football stadium or small country would
include a region of uncertainty that is only slightly larger than the
Target itself. In these cases, much of the guidance in this document
is not applicable. Indeed, as the accuracy of location determination
technology improves, it could be that the advice this document
contains becomes less relevant by the same measure.

3.2. Representation of Uncertainty and Confidence in PIDF-LO

A set of shapes suitable for the expression of uncertainty in
location estimates in the PIDF-LO are described in [GeoShape]. These
shapes are the recommended form for the representation of uncertainty
in PIDF-LO [RFC4119] documents.

The PIDF-LO can contain uncertainty, but it does not include an
indication of confidence. [RFC5491] defines a fixed value of 95%.
Similarly, the PIDF-LO format does not provide an indication of the
shape of the PDF. Section 4 defines elements to convey this
information in PIDF-LO.

Absence of uncertainty information in a PIDF-LO document does not
indicate that there is no uncertainty in the location estimate.
Uncertainty might not have been calculated for the estimate, or it
may be withheld for privacy purposes.

If the Point shape is used, confidence and uncertainty are unknown; a
receiver can either assume a confidence of 0% or infinite
uncertainty. The same principle applies on the altitude axis for
two-dimensional shapes like the Circle.

3.3. Uncertainty and Confidence for Civic Addresses

Automatically determined civic addresses [RFC5139] inherently include
uncertainty, based on the area of the most precise element that is
specified. In this case, uncertainty is effectively described by the
presence or absence of elements. To the recipient of location
information, elements that are not present are uncertain.

To apply the concept of uncertainty to civic addresses, it is helpful
to unify the conceptual models of civic address with geodetic
location information. This is particularly useful when considering

civic addresses that are determined using reverse geocoding (that is,
the process of translating geodetic information into civic
addresses).

In the unified view, a civic address defines a series of (sometimes
non-orthogonal) spatial partitions. The first is the implicit
partition that identifies the surface of the earth and the space near
the surface. The second is the country. Each label that is included
in a civic address provides information about a different set of
spatial partitions. Some partitions require slight adjustments from
a standard interpretation: for instance, a road includes all
properties that adjoin the street. Each label might need to be
interpreted with other values to provide context.

As a value at each level is interpreted, one or more spatial
partitions at that level are selected, and all other partitions of
that type are excluded. For non-orthogonal partitions, only the
portion of the partition that fits within the existing space is
selected. This is what distinguishes King Street in Sydney from King
Street in Melbourne. Each defined element selects a partition of
space. The resulting location is the intersection of all selected
spaces.

The resulting spatial partition can be considered as a region of
uncertainty.

Note: This view is a potential perspective on the process of
   geocoding -- the translation of a civic address to a geodetic
   location.

Uncertainty in civic addresses can be increased by removing elements.
This does not increase confidence unless additional information is
used. Similarly, arbitrarily increasing uncertainty in a geodetic
location does not increase confidence.

3.4. DHCP Location Configuration Information and Uncertainty

Location information is often measured in two or three dimensions;
expressions of uncertainty in one dimension only are rare. The
"resolution" parameters in [RFC6225] provide an indication of how
many bits of a number are valid, which could be interpreted as an
expression of uncertainty in one dimension.

[RFC6225] defines a means for representing uncertainty, but a value
for confidence is not specified. A default value of 95% confidence
should be assumed for the combination of the uncertainty on each
axis. This is consistent with the transformation of those forms into

the uncertainty representations from [RFC5491]. That is, the
confidence of the resultant rectangular Polygon or Prism is assumed
to be 95%.

4. Representation of Confidence in PIDF-LO

On the whole, a fixed definition for confidence is preferable,
primarily because it ensures consistency between implementations.
Location generators that are aware of this constraint can generate
location information at the required confidence. Location recipients
are able to make sensible assumptions about the quality of the
information that they receive.

In some circumstances -- particularly with preexisting systems --
location generators might be unable to provide location information
with consistent confidence. Existing systems sometimes specify
confidence at 38%, 67%, or 90%. Existing forms of expressing
location information, such as that defined in [TS-3GPP-23_032],
contain elements that express the confidence in the result.

The addition of a confidence element provides information that was
previously unavailable to recipients of location information.
Without this information, a location server or generator that has
access to location information with a confidence lower than 95% has
two options:

o The location server can scale regions of uncertainty in an attempt
   to achieve 95% confidence. This scaling process significantly
   degrades the quality of the information, because the location
   server might not have the necessary information to scale
   appropriately; the location server is forced to make assumptions
   that are likely to result in either an overly conservative
   estimate with high uncertainty or an overestimate of confidence.

o The location server can ignore the confidence entirely, which
   results in giving the recipient a false impression of its quality.

Both of these choices degrade the quality of the information
provided.

The addition of a confidence element avoids this problem entirely if
a location recipient supports and understands the element. A
recipient that does not understand -- and, hence, ignores -- the
confidence element is in no worse a position than if the location
server ignored confidence.

4.1. The "confidence" Element

   The "confidence" element MAY be added to the "location-info" element
   of the PIDF-LO [RFC4119] document. This element expresses the
   confidence in the associated location information as a percentage. A
   special "unknown" value is reserved to indicate that confidence is
   supported, but not known to the Location Generator.

   The "confidence" element optionally includes an attribute that
   indicates the shape of the PDF of the associated region of
   uncertainty. Three values are possible: unknown, normal, and
   rectangular.

   Indicating a particular PDF only indicates that the distribution
   approximately fits the given shape based on the methods used to
   generate the location information. The PDF is normal if there are a
   large number of small, independent sources of error. It is
   rectangular if all points within the area have roughly equal
   probability of being the actual location of the Target. Otherwise,
   the PDF MUST either be set to unknown or omitted.

   If a PIDF-LO does not include the confidence element, the confidence
   of the location estimate is 95%, as defined in [RFC5491].

   A Point shape does not have uncertainty (or it has infinite
   uncertainty), so confidence is meaningless for a Point; therefore,
   this element MUST be omitted if only a Point is provided.

4.2. Generating Locations with Confidence

   Location generators SHOULD attempt to ensure that confidence is equal
   in each dimension when generating location information. This
   restriction, while not always practical, allows for more accurate
   scaling, if scaling is necessary.

   A confidence element MUST be included with all location information
   that includes uncertainty (that is, all forms other than a Point). A
   special "unknown" is used if confidence is not known.

4.3. Consuming and Presenting Confidence

   The inclusion of confidence that is anything other than 95% presents
   a potentially difficult usability problem for applications that use
   location information. Effectively communicating the probability that
   a location is incorrect to a user can be difficult.

It is inadvisable to simply display locations of any confidence, or
to display confidence in a separate or non-obvious fashion. If
locations with different confidence levels are displayed such that
the distinction is subtle or easy to overlook -- such as using fine
graduations of color or transparency for graphical uncertainty
regions or displaying uncertainty graphically, but providing
confidence as supplementary text -- a user could fail to notice a
difference in the quality of the location information that might be
significant.

Depending on the circumstances, different ways of handling confidence
might be appropriate. Section 5 describes techniques that could be
appropriate for consumers that use automated processing.

Providing that the full implications of any choice for the
application are understood, some amount of automated processing could
be appropriate. In a simple example, applications could choose to
discard or suppress the display of location information if confidence
does not meet a predetermined threshold.

In settings where there is an opportunity for user training, some of
these problems might be mitigated by defining different operational
procedures for handling location information at different confidence
levels.

5. Manipulation of Uncertainty

   This section deals with manipulation of location information that
   contains uncertainty.

   The following rules generally apply when manipulating location
   information:

   o Where calculations are performed on coordinate information, these
      should be performed in Cartesian space and the results converted
      back to latitude, longitude, and altitude. A method for
      converting to and from Cartesian coordinates is included in
      Appendix A.

         While some approximation methods are useful in simplifying
         calculations, treating latitude and longitude as Cartesian axes
         is never advisable. The two axes are not orthogonal. Errors
         can arise from the curvature of the earth and from the
         convergence of longitude lines.

o Normal rounding rules do not apply when rounding uncertainty.
   When rounding, the region of uncertainty always increases (that
   is, errors are rounded up) and confidence is always rounded down
   (see [NIST.TN1297]). This means that any manipulation of
   uncertainty is a non-reversible operation; each manipulation can
   result in the loss of some information.

5.1. Reduction of a Location Estimate to a Point

   Manipulating location estimates that include uncertainty information
   requires additional complexity in systems. In some cases, systems
   only operate on definitive values, that is, a single point.

   This section describes algorithms for reducing location estimates to
   a simple form without uncertainty information. Having a consistent
   means for reducing location estimates allows for interaction between
   applications that are able to use uncertainty information and those
   that cannot.

   Note: Reduction of a location estimate to a point constitutes a
      reduction in information. Removing uncertainty information can
      degrade results in some applications. Also, there is a natural
      tendency to misinterpret a Point location as representing a
      location without uncertainty. This could lead to more serious
      errors. Therefore, these algorithms should only be applied where
      necessary.

   Several different approaches can be taken when reducing a location
   estimate to a point. Different methods each make a set of
   assumptions about the properties of the PDF and the selected point;
   no one method is more "correct" than any other. For any given region
   of uncertainty, selecting an arbitrary point within the area could be
   considered valid; however, given the aforementioned problems with
   Point locations, a more rigorous approach is appropriate.

   Given a result with a known distribution, selecting the point within
   the area that has the highest probability is a more rigorous method.
   Alternatively, a point could be selected that minimizes the overall
   error; that is, it minimizes the expected value of the difference
   between the selected point and the "true" value.

   If a rectangular distribution is assumed, the centroid of the area or
   volume minimizes the overall error. Minimizing the error for a
   normal distribution is mathematically complex. Therefore, this
   document opts to select the centroid of the region of uncertainty
   when selecting a point.

5.1.1. Centroid Calculation

   For regular shapes, such as Circle, Sphere, Ellipse, and Ellipsoid,
   this approach equates to the center point of the region. For regions
   of uncertainty that are expressed as regular Polygons and Prisms, the
   center point is also the most appropriate selection.

   For the Arc-Band shape and non-regular Polygons and Prisms, selecting
   the centroid of the area or volume minimizes the overall error. This
   assumes that the PDF is rectangular.

   Note: The centroid of a concave Polygon or Arc-Band shape is not
      necessarily within the region of uncertainty.

5.1.1.1. Arc-Band Centroid

   The centroid of the Arc-Band shape is found along a line that bisects
   the arc. The centroid can be found at the following distance from
   the starting point of the arc-band (assuming an arc-band with an
   inner radius of "r", outer radius "R", start angle "a", and opening
   angle "o"):

      d = 4 * sin(o/2) * (R*R + R*r + r*r) / (3*o*(R + r))

   This point can be found along the line that bisects the arc; that is,
   the line at an angle of "a + (o/2)".

5.1.1.2. Polygon Centroid

   Calculating a centroid for the Polygon and Prism shapes is more
   complex. Polygons that are specified using geodetic coordinates are
   not necessarily coplanar. For Polygons that are specified without an
   altitude, choose a value for altitude before attempting this process;
   an altitude of 0 is acceptable.

      The method described in this section is simplified by assuming
      that the surface of the earth is locally flat. This method
      degrades as polygons become larger; see [GeoShape] for
      recommendations on polygon size.

   The polygon is translated to a new coordinate system that has an x-y
   plane roughly parallel to the polygon. This enables the elimination
   of z-axis values and calculating a centroid can be done using only x
   and y coordinates. This requires that the upward normal for the
   polygon be known.

To translate the polygon coordinates, apply the process described in
Appendix B to find the normal vector "N = [Nx,Ny,Nz]". This value
should be made a unit vector to ensure that the transformation matrix
is a special orthogonal matrix. From this vector, select two vectors
that are perpendicular to this vector and combine these into a
transformation matrix.

If "Nx" and "Ny" are non-zero, the matrices in Figure 3 can be used,
given "p = sqrt(Nx^2 + Ny^2)". More transformations are provided
later in this section for cases where "Nx" or "Ny" are zero.

```
        [   -Ny/p      Nx/p      0  ]          [ -Ny/p  -Nx*Nz/p  Nx ]
    T = [ -Nx*Nz/p  -Ny*Nz/p     p  ]    T' = [  Nx/p  -Ny*Nz/p  Ny ]
        [    Nx         Ny       Nz ]          [   0        p     Nz ]
                (Transform)                         (Reverse Transform)
```

                Figure 3: Recommended Transformation Matrices

To apply a transform to each point in the polygon, form a matrix from
the Cartesian Earth-Centered, Earth-Fixed (ECEF) coordinates and use
matrix multiplication to determine the translated coordinates.

```
    [   -Ny/p      Nx/p     0  ]   [ x[1]  x[2]  x[3]  ...  x[n] ]
    [ -Nx*Nz/p  -Ny*Nz/p    p  ] * [ y[1]  y[2]  y[3]  ...  y[n] ]
    [    Nx        Ny       Nz ]   [ z[1]  z[2]  z[3]  ...  z[n] ]

        [ x'[1]  x'[2]  x'[3]  ... x'[n] ]
      = [ y'[1]  y'[2]  y'[3]  ... y'[n] ]
        [ z'[1]  z'[2]  z'[3]  ... z'[n] ]
```

                        Figure 4: Transformation

Alternatively, direct multiplication can be used to achieve the same
result:

    x'[i] = -Ny * x[i] / p + Nx * y[i] / p

    y'[i] = -Nx * Nz * x[i] / p - Ny * Nz * y[i] / p + p * z[i]

    z'[i] = Nx * x[i] + Ny * y[i] + Nz * z[i]

The first and second rows of this matrix ("x'" and "y'") contain the
values that are used to calculate the centroid of the polygon. To
find the centroid of this polygon, first find the area using:

    A = sum from i=1..n of (x'[i]*y'[i+1]-x'[i+1]*y'[i]) / 2

For these formulae, treat each set of coordinates as circular, that
is "x'[0] == x'[n]" and "x'[n+1] == x'[1]". Based on the area, the
centroid along each axis can be determined by:

    Cx' = sum (x'[i]+x'[i+1]) * (x'[i]*y'[i+1]-x'[i+1]*y'[i]) / (6*A)

    Cy' = sum (y'[i]+y'[i+1]) * (x'[i]*y'[i+1]-x'[i+1]*y'[i]) / (6*A)

Note: The formula for the area of a polygon will return a negative
    value if the polygon is specified in a clockwise direction. This
    can be used to determine the orientation of the polygon.

The third row contains a distance from a plane parallel to the
polygon. If the polygon is coplanar, then the values for "z'" are
identical; however, the constraints recommended in [RFC5491] mean
that this is rarely the case. To determine "Cz'", average these
values:

    Cz' = sum z'[i] / n

Once the centroid is known in the transformed coordinates, these can
be transformed back to the original coordinate system. The reverse
transformation is shown in Figure 5.

    [-Ny/p   -Nx*Nz/p Nx]   [Cx'   ]   [Cx]
    [Nx/p   -Ny*Nz/p Ny] * [Cy'   ] = [Cy]
    [0   p   Nz]   [sum of z'[i] / n]   [Cz]

                    Figure 5: Reverse Transformation

The reverse transformation can be applied directly as follows:

    Cx = -Ny * Cx' / p - Nx * Nz * Cy' / p + Nx * Cz'

    Cy = Nx * Cx' / p - Ny * Nz * Cy' / p + Ny * Cz'

    Cz = p * Cy' + Nz * Cz'

The ECEF value "[Cx,Cy,Cz]" can then be converted back to geodetic
coordinates. Given a polygon that is defined with no altitude or
equal altitudes for each point, the altitude of the result can be
either ignored or reset after converting back to a geodetic value.

The centroid of the Prism shape is found by finding the centroid of
the base polygon and raising the point by half the height of the
prism. This can be added to altitude of the final result;
alternatively, this can be added to "Cz'", which ensures that
negative height is correctly applied to polygons that are defined in
a clockwise direction.

The recommended transforms only apply if "Nx" and "Ny" are non-zero.
If the normal vector is "[0,0,1]" (that is, along the z-axis), then
no transform is necessary. Similarly, if the normal vector is
"[0,1,0]" or "[1,0,0]", avoid the transformation and use the x and z
coordinates or y and z coordinates (respectively) in the centroid
calculation phase. If either "Nx" or "Ny" are zero, the alternative
transform matrices in Figure 6 can be used. The reverse transform is
the transpose of this matrix.

```
 if Nx == 0:                                    | if Ny == 0:
    [ 0  -Nz  Ny ]          [ 0   1   0 ]       |                [ -Nz   0   Nx ]
 T = [ 1   0   0 ]   T' = [ -Nz  0   Ny ]       |   T = T' = [  0    1   0  ]
    [ 0   Ny  Nz ]          [ Ny  0   Nz ]      |                [ Nx   0   Nz ]
```

              Figure 6: Alternative Transformation Matrices

5.2. Conversion to Circle or Sphere

   The circle or sphere are simple shapes that suit a range of
   applications. A circle or sphere contains fewer units of data to
   manipulate, which simplifies operations on location estimates.

   The simplest method for converting a location estimate to a Circle or
   Sphere shape is to determine the centroid and then find the longest
   distance to any point in the region of uncertainty to that point.
   This distance can be determined based on the shape type:

   Circle/Sphere: No conversion necessary.

   Ellipse/Ellipsoid: The greater of either semi-major axis or altitude
      uncertainty.

   Polygon/Prism: The distance to the farthest vertex of the Polygon
      (for a Prism, it is only necessary to check points on the base).

   Arc-Band: The farthest length from the centroid to the points where
      the inner and outer arc end. This distance can be calculated by
      finding the larger of the two following formulae:

         X = sqrt(d*d + R*R - 2*d*R*cos(o/2))

         x = sqrt(d*d + r*r - 2*d*r*cos(o/2))

   Once the Circle or Sphere shape is found, the associated confidence
   can be increased if the result is known to follow a normal
   distribution. However, this is a complicated process and provides
   limited benefit. In many cases, it also violates the constraint that
   confidence in each dimension be the same. Confidence should be
   unchanged when performing this conversion.

   Two-dimensional shapes are converted to a Circle; three-dimensional
   shapes are converted to a Sphere.

5.3. Conversion from Three-Dimensional to Two-Dimensional

   A three-dimensional shape can be easily converted to a two-
   dimensional shape by removing the altitude component. A Sphere
   becomes a Circle; a Prism becomes a Polygon; an Ellipsoid becomes an
   Ellipse. Each conversion is simple, requiring only the removal of
   those elements relating to altitude.

   The altitude is unspecified for a two-dimensional shape and therefore
   has unlimited uncertainty along the vertical axis. The confidence
   for the two-dimensional shape is thus higher than the three-
   dimensional shape. Assuming equal confidence on each axis, the
   confidence of the Circle can be increased using the following
   approximate formula:

      C[2d] >= C[3d] ^ (2/3)

   "C[2d]" is the confidence of the two-dimensional shape and "C[3d]" is
   the confidence of the three-dimensional shape. For example, a Sphere
   with a confidence of 95% can be simplified to a Circle of equal
   radius with confidence of 96.6%.

5.4. Increasing and Decreasing Uncertainty and Confidence

   The combination of uncertainty and confidence provide a great deal of
   information about the nature of the data that is being measured. If
   uncertainty, confidence, and PDF are known, certain information can
   be extrapolated. In particular, the uncertainty can be scaled to
   meet a desired confidence or the confidence for a particular region
   of uncertainty can be found.

In general, confidence decreases as the region of uncertainty
decreases in size, and confidence increases as the region of
uncertainty increases in size. However, this depends on the PDF;
expanding the region of uncertainty for a rectangular distribution
has no effect on confidence without additional information. If the
region of uncertainty is increased during the process of obfuscation
(see [RFC6772]), then the confidence cannot be increased.

A region of uncertainty that is reduced in size always has a lower
confidence.

A region of uncertainty that has an unknown PDF shape cannot be
reduced in size reliably. The region of uncertainty can be expanded,
but only if confidence is not increased.

This section makes the simplifying assumption that location
information is symmetrically and evenly distributed in each
dimension. This is not necessarily true in practice. If better
information is available, alternative methods might produce better
results.

5.4.1. Rectangular Distributions

Uncertainty that follows a rectangular distribution can only be
decreased in size. Increasing uncertainty has no value, since it has
no effect on confidence. Since the PDF is constant over the region
of uncertainty, the resulting confidence is determined by the
following formula:

   $Cr = Co * Ur / Uo$

Where "Uo" and "Ur" are the sizes of the original and reduced regions
of uncertainty (either the area or the volume of the region); "Co"
and "Cr" are the confidence values associated with each region.

Information is lost by decreasing the region of uncertainty for a
rectangular distribution. Once reduced in size, the uncertainty
region cannot subsequently be increased in size.

5.4.2. Normal Distributions

Uncertainty and confidence can be both increased and decreased for a
normal distribution. This calculation depends on the number of
dimensions of the uncertainty region.

For a normal distribution, uncertainty and confidence are related to
the standard deviation of the function. The following function
defines the relationship between standard deviation, uncertainty, and
confidence along a single axis:

    S[x] = U[x] / (sqrt(2) * erfinv(C[x]))

Where "S[x]" is the standard deviation, "U[x]" is the uncertainty,
and "C[x]" is the confidence along a single axis. "erfinv" is the
inverse error function.

Scaling a normal distribution in two dimensions requires several
assumptions. Firstly, it is assumed that the distribution along each
axis is independent. Secondly, the confidence for each axis is
assumed to be the same. Therefore, the confidence along each axis
can be assumed to be:

    C[x] = Co ^ (1/n)

Where "C[x]" is the confidence along a single axis and "Co" is the
overall confidence and "n" is the number of dimensions in the
uncertainty.

Therefore, to find the uncertainty for each axis at a desired
confidence, "Cd", apply the following formula:

    Ud[x] <= U[x] * (erfinv(Cd ^ (1/n)) / erfinv(Co ^ (1/n)))

For regular shapes, this formula can be applied as a scaling factor
in each dimension to reach a required confidence.

5.5. Determining Whether a Location Is within a Given Region

   A number of applications require that a judgment be made about
   whether a Target is within a given region of interest. Given a
   location estimate with uncertainty, this judgment can be difficult.
   A location estimate represents a probability distribution, and the
   true location of the Target cannot be definitively known. Therefore,
   the judgment relies on determining the probability that the Target is
   within the region.

   The probability that the Target is within a particular region is
   found by integrating the PDF over the region. For a normal
   distribution, there are no analytical methods that can be used to
   determine the integral of the two- or three-dimensional PDF over an
   arbitrary region. The complexity of numerical methods is also too
   great to be useful in many applications; for example, finding the
   integral of the PDF in two or three dimensions across the overlap

between the uncertainty region and the target region. If the PDF is
unknown, no determination can be made without a simplifying
assumption.

When judging whether a location is within a given region, this
document assumes that uncertainties are rectangular. This introduces
errors, but simplifies the calculations significantly. Prior to
applying this assumption, confidence should be scaled to 95%.

Note: The selection of confidence has a significant impact on the
   final result. Only use a different confidence if an uncertainty
   value for 95% confidence cannot be found.

Given the assumption of a rectangular distribution, the probability
that a Target is found within a given region is found by first
finding the area (or volume) of overlap between the uncertainty
region and the region of interest. This is multiplied by the
confidence of the location estimate to determine the probability.
Figure 7 shows an example of finding the area of overlap between the
region of uncertainty and the region of interest.

```
                    _.-"""""-._
                  .'           '.      _ Region of
                 /               \    /  Uncertainty
            ..+-"""--..           |
          .-'  | :::::::  `-.      |
        ,'     | :: Ao :::  `.     |
       /        \ :::::::::::: \  /
      /          `._ ::::::  _.X
     |              `-....-'   |
     |                         |
     |                         |
      \                       /
       `.                   .'  \_ Region of
        `._             _.'        Interest
           `--.._____..--'
```
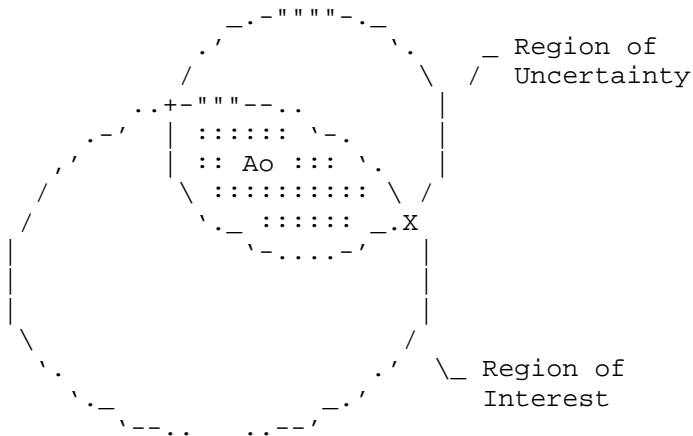
           Figure 7: Area of Overlap between Two Circular Regions

Once the area of overlap, "Ao", is known, the probability that the
Target is within the region of interest, "Pi", is:

    Pi = Co * Ao / Au

Given that the area of the region of uncertainty is "Au" and the
confidence is "Co".

This probability is often input to a decision process that has a
limited set of outcomes; therefore, a threshold value needs to be
selected. Depending on the application, different threshold
probabilities might be selected. A probability of 50% or greater is
recommended before deciding that an uncertain value is within a given
region. If the decision process selects between two or more regions,
as is required by [RFC5222], then the region with the highest
probability can be selected.

5.5.1. Determining the Area of Overlap for Two Circles

Determining the area of overlap between two arbitrary shapes is a
non-trivial process. Reducing areas to circles (see Section 5.2)
enables the application of the following process.

Given the radius of the first circle "r", the radius of the second
circle "R", and the distance between their center points "d", the
following set of formulae provide the area of overlap "Ao".

o If the circles don't overlap, that is "d >= r+R", "Ao" is zero.

o If one of the two circles is entirely within the other, that is
   "d <= |r-R|", the area of overlap is the area of the smaller
   circle.

o Otherwise, if the circles partially overlap, that is "d < r+R" and
   "d > |r-R|", find "Ao" using:

    a = (r^2 - R^2 + d^2)/(2*d)

    Ao = r^2*acos(a/r) + R^2*acos((d - a)/R) - d*sqrt(r^2 - a^2)

A value for "d" can be determined by converting the center points to
Cartesian coordinates and calculating the distance between the two
center points:

    d = sqrt((x1-x2)^2 + (y1-y2)^2 + (z1-z2)^2)

5.5.2. Determining the Area of Overlap for Two Polygons

   A calculation of overlap based on polygons can give better results
   than the circle-based method. However, efficient calculation of
   overlapping area is non-trivial. Algorithms such as Vatti's clipping
   algorithm [Vatti92] can be used.

   For large polygonal areas, it might be that geodesic interpolation is
   used. In these cases, altitude is also frequently omitted in
   describing the polygon. For such shapes, a planar projection can
   still give a good approximation of the area of overlap if the larger
   area polygon is projected onto the local tangent plane of the
   smaller. This is only possible if the only area of interest is that
   contained within the smaller polygon. Where the entire area of the
   larger polygon is of interest, geodesic interpolation is necessary.

6. Examples

   This section presents some examples of how to apply the methods
   described in Section 5.

6.1. Reduction to a Point or Circle

   Alice receives a location estimate from her Location Information
   Server (LIS) that contains an ellipsoidal region of uncertainty.
   This information is provided at 19% confidence with a normal PDF. A
   PIDF-LO extract for this information is shown in Figure 8.

```
   <gp:geopriv>
     <gp:location-info>
       <gs:Ellipsoid srsName="urn:ogc:def:crs:EPSG::4979">
         <gml:pos>-34.407242 150.882518 34</gml:pos>
         <gs:semiMajorAxis uom="urn:ogc:def:uom:EPSG::9001">
           7.7156
         </gs:semiMajorAxis>
         <gs:semiMinorAxis uom="urn:ogc:def:uom:EPSG::9001">
           3.31
         </gs:semiMinorAxis>
         <gs:verticalAxis uom="urn:ogc:def:uom:EPSG::9001">
           28.7
         </gs:verticalAxis>
         <gs:orientation uom="urn:ogc:def:uom:EPSG::9102">
           43
         </gs:orientation>
       </gs:Ellipsoid>
       <con:confidence pdf="normal">95</con:confidence>
     </gp:location-info>
     <gp:usage-rules/>
   </gp:geopriv>
```

                   Figure 8: Alice's Ellipsoid Location

   This information can be reduced to a point simply by extracting the
   center point, that is [-34.407242, 150.882518, 34].

   If some limited uncertainty were required, the estimate could be
   converted into a circle or sphere. To convert to a sphere, the
   radius is the largest of the semi-major, semi-minor and vertical
   axes; in this case, 28.7 meters.

   However, if only a circle is required, the altitude can be dropped as
   can the altitude uncertainty (the vertical axis of the ellipsoid),
   resulting in a circle at [-34.407242, 150.882518] of radius 7.7156
   meters.

   Bob receives a location estimate with a Polygon shape (which roughly
   corresponds to the location of the Sydney Opera House). This
   information is shown in Figure 9.

```
<gml:Polygon srsName="urn:ogc:def:crs:EPSG::4326">
  <gml:exterior>
    <gml:LinearRing>
      <gml:posList>
        -33.856625 151.215906 -33.856299 151.215343
        -33.856326 151.214731 -33.857533 151.214495
        -33.857720 151.214613 -33.857369 151.215375
        -33.856625 151.215906
      </gml:posList>
    </gml:LinearRing>
  </gml:exterior>
</gml:Polygon>
```

                     Figure 9: Bob's Polygon Location

   To convert this to a polygon, each point is firstly assigned an
   altitude of zero and converted to ECEF coordinates (see Appendix A).
   Then, a normal vector for this polygon is found (see Appendix B).
   The result of each of these stages is shown in Figure 10. Note that
   the numbers shown in this document are rounded only for formatting
   reasons; the actual calculations do not include rounding, which would
   generate significant errors in the final values.

```
Polygon in ECEF coordinate space
   (repeated point omitted and transposed to fit):
         [ -4.6470e+06   2.5530e+06   -3.5333e+06 ]
         [ -4.6470e+06   2.5531e+06   -3.5332e+06 ]
pecef = [ -4.6470e+06   2.5531e+06   -3.5332e+06 ]
         [ -4.6469e+06   2.5531e+06   -3.5333e+06 ]
         [ -4.6469e+06   2.5531e+06   -3.5334e+06 ]
         [ -4.6469e+06   2.5531e+06   -3.5333e+06 ]
```

Normal Vector: n = [-0.72782 0.39987 -0.55712]

```
Transformation Matrix:
    [ -0.48152  -0.87643   0.00000 ]
t = [ -0.48828   0.26827   0.83043 ]
    [ -0.72782   0.39987  -0.55712 ]
```

```
Transformed Coordinates:
          [   8.3206e+01  1.9809e+04  6.3715e+06 ]
          [   3.1107e+01  1.9845e+04  6.3715e+06 ]
pecef' = [ -2.5528e+01  1.9842e+04  6.3715e+06 ]
          [ -4.7367e+01  1.9708e+04  6.3715e+06 ]
          [ -3.6447e+01  1.9687e+04  6.3715e+06 ]
          [   3.4068e+01  1.9726e+04  6.3715e+06 ]
```

Two dimensional polygon area: A = 12600 m^2
Two-dimensional polygon centroid: C' = [8.8184e+00 1.9775e+04]

Average of pecef' z coordinates: 6.3715e+06

```
Reverse Transformation Matrix:
     [ -0.48152  -0.48828  -0.72782 ]
t' = [ -0.87643   0.26827   0.39987 ]
     [  0.00000   0.83043  -0.55712 ]
```

Polygon centroid (ECEF): C = [-4.6470e+06 2.5531e+06 -3.5333e+06]
Polygon centroid (Geo): Cg = [-33.856926 151.215102 -4.9537e-04]

                  Figure 10: Calculation of Polygon Centroid

The point conversion for the polygon uses the final result, "Cg",
ignoring the altitude since the original shape did not include
altitude.

To convert this to a circle, take the maximum distance in ECEF
coordinates from the center point to each of the points. This
results in a radius of 99.1 meters. Confidence is unchanged.

6.2. Increasing and Decreasing Confidence

   Assume that confidence is known to be 19% for Alice's location
   information. This is a typical value for a three-dimensional
   ellipsoid uncertainty of normal distribution where the standard
   deviation is used directly for uncertainty in each dimension. The
   confidence associated with Alice's location estimate is quite low for
   many applications. Since the estimate is known to follow a normal
   distribution, the method in Section 5.4.2 can be used. Each axis can
   be scaled by:

      scale = erfinv(0.95^(1/3)) / erfinv(0.19^(1/3)) = 2.9937

   Ensuring that rounding always increases uncertainty, the location
   estimate at 95% includes a semi-major axis of 23.1, a semi-minor axis
   of 10 and a vertical axis of 86.

   Bob's location estimate (from the previous example) covers an area of
   approximately 12600 square meters. If the estimate follows a
   rectangular distribution, the region of uncertainty can be reduced in
   size. Here we find the confidence that Bob is within the smaller
   area of the Concert Hall. For the Concert Hall, the polygon
   [-33.856473, 151.215257; -33.856322, 151.214973;
   -33.856424, 151.21471; -33.857248, 151.214753;
   -33.857413, 151.214941; -33.857311, 151.215128] is used. To use this
   new region of uncertainty, find its area using the same translation
   method described in Section 5.1.1.2, which produces 4566.2 square
   meters. Given that the Concert Hall is entirely within Bob's
   original location estimate, the confidence associated with the
   smaller area is therefore 95% * 4566.2 / 12600 = 34%.

6.3. Matching Location Estimates to Regions of Interest

   Suppose that a circular area is defined centered at
   [-33.872754, 151.20683] with a radius of 1950 meters. To determine
   whether Bob is found within this area -- given that Bob is at
   [-34.407242, 150.882518] with an uncertainty radius 7.7156 meters --
   we apply the method in Section 5.5. Using the converted Circle shape
   for Bob's location, the distance between these points is found to be
   1915.26 meters. The area of overlap between Bob's location estimate
   and the region of interest is therefore 2209 square meters and the
   area of Bob's location estimate is 30853 square meters. This gives
   the estimated probability that Bob is less than 1950 meters from the
   selected point as 67.8%.

   Note that if 1920 meters were chosen for the distance from the
   selected point, the area of overlap is only 16196 square meters and
   the confidence is 49.8%. Therefore, it is marginally more likely
   that Bob is outside the region of interest, despite the center point
   of his location estimate being within the region.

6.4. PIDF-LO with Confidence Example

   The PIDF-LO document in Figure 11 includes a representation of
   uncertainty as a circular area. The confidence element (on the line
   marked with a comment) indicates that the confidence is 67% and that
   it follows a normal distribution.

```
   <pidf:presence
       xmlns:pidf="urn:ietf:params:xml:ns:pidf"
       xmlns:dm="urn:ietf:params:xml:ns:pidf:data-model"
       xmlns:gp="urn:ietf:params:xml:ns:pidf:geopriv10"
       xmlns:gs="http://www.opengis.net/pidflo/1.0"
       xmlns:gml="http://www.opengis.net/gml"
       xmlns:con="urn:ietf:params:xml:ns:geopriv:conf"
       entity="pres:alice@example.com">
     <dm:device id="sg89ab">
       <gp:geopriv>
         <gp:location-info>
           <gs:Circle srsName="urn:ogc:def:crs:EPSG::4326">
             <gml:pos>42.5463 -73.2512</gml:pos>
             <gs:radius uom="urn:ogc:def:uom:EPSG::9001">
               850.24
             </gs:radius>
           </gs:Circle>
  <!--c--> <con:confidence pdf="normal">67</con:confidence>
         </gp:location-info>
         <gp:usage-rules/>
       </gp:geopriv>
       <dm:deviceID>mac:010203040506</dm:deviceID>
     </dm:device>
   </pidf:presence>
```

                  Figure 11: Example PIDF-LO with Confidence

7. Confidence Schema

```xml
<?xml version="1.0"?>
<xs:schema
    xmlns:conf="urn:ietf:params:xml:ns:geopriv:conf"
    xmlns:xs="http://www.w3.org/2001/XMLSchema"
    targetNamespace="urn:ietf:params:xml:ns:geopriv:conf"
    elementFormDefault="qualified"
    attributeFormDefault="unqualified">

  <xs:annotation>
    <xs:appinfo
        source="urn:ietf:params:xml:schema:geopriv:conf">
      PIDF-LO Confidence
    </xs:appinfo>
    <xs:documentation
        source="http://www.rfc-editor.org/rfc/rfc7459.txt">
      This schema defines an element that is used for indicating
      confidence in PIDF-LO documents.
    </xs:documentation>
  </xs:annotation>

  <xs:element name="confidence" type="conf:confidenceType"/>

  <xs:complexType name="confidenceType">
    <xs:simpleContent>
      <xs:extension base="conf:confidenceBase">
        <xs:attribute name="pdf" type="conf:pdfType"
                      default="unknown"/>
      </xs:extension>
    </xs:simpleContent>
  </xs:complexType>

  <xs:simpleType name="confidenceBase">
    <xs:union>
      <xs:simpleType>
        <xs:restriction base="xs:decimal">
          <xs:minExclusive value="0.0"/>
          <xs:maxExclusive value="100.0"/>
        </xs:restriction>
      </xs:simpleType>
      <xs:simpleType>
        <xs:restriction base="xs:token">
          <xs:enumeration value="unknown"/>
        </xs:restriction>
      </xs:simpleType>
    </xs:union>
  </xs:simpleType>
```

```
      <xs:simpleType name="pdfType">
        <xs:restriction base="xs:token">
          <xs:enumeration value="unknown"/>
          <xs:enumeration value="normal"/>
          <xs:enumeration value="rectangular"/>
        </xs:restriction>
      </xs:simpleType>

    </xs:schema>
```

8. IANA Considerations

8.1. URN Sub-Namespace Registration for
      urn:ietf:params:xml:ns:geopriv:conf

   A new XML namespace, "urn:ietf:params:xml:ns:geopriv:conf", has been
   registered, as per the guidelines in [RFC3688].

   URI: urn:ietf:params:xml:ns:geopriv:conf

   Registrant Contact: IETF GEOPRIV working group (geopriv@ietf.org),
      Martin Thomson (martin.thomson@gmail.com).

   XML:

```
      BEGIN
        <?xml version="1.0"?>
        <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
          "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
        <html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en">
          <head>
            <title>PIDF-LO Confidence Attribute</title>
          </head>
          <body>
            <h1>Namespace for PIDF-LO Confidence Attribute</h1>
            <h2>urn:ietf:params:xml:ns:geopriv:conf</h2>
            <p>See <a href="http://www.rfc-editor.org/rfc/rfc7459.txt">
              RFC 7459</a>.</p>
          </body>
        </html>
      END
```

8.2. XML Schema Registration

   An XML schema has been registered, as per the guidelines in
   [RFC3688].

   URI: urn:ietf:params:xml:schema:geopriv:conf

   Registrant Contact: IETF GEOPRIV working group (geopriv@ietf.org),
      Martin Thomson (martin.thomson@gmail.com).

   Schema: The XML for this schema can be found as the entirety of
      Section 7 of this document.

9. Security Considerations

   This document describes methods for managing and manipulating
   uncertainty in location. No specific security concerns arise from
   most of the information provided. The considerations of [RFC4119]
   all apply.

   A thorough treatment of the privacy implications of describing
   location information are discussed in [RFC6280]. Including
   uncertainty information increases the amount of information
   available; and altering uncertainty is not an effective privacy
   mechanism.

   Providing uncertainty and confidence information can reveal
   information about the process by which location information is
   generated. For instance, it might reveal information that could be
   used to infer that a user is using a mobile device with a GPS, or
   that a user is acquiring location information from a particular
   network-based service. A Rule Maker might choose to remove
   uncertainty-related fields from a location object in order to protect
   this information. Note however that information might not be
   perfectly protected due to difficulties associated with location
   obfuscation, as described in Section 13.5 of [RFC6772]. In
   particular, increasing uncertainty does not necessarily result in a
   reduction of the information conveyed by the location object.

   Adding confidence to location information risks misinterpretation by
   consumers of location that do not understand the element. This could
   be exploited, particularly when reducing confidence, since the
   resulting uncertainty region might include locations that are less
   likely to contain the Target than the recipient expects. Since this
   sort of error is always a possibility, the impact of this is low.

10. References

10.1. Normative References

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119, March 1997,
              <http://www.rfc-editor.org/info/rfc2119>.

   [RFC3688] Mealling, M., "The IETF XML Registry", BCP 81, RFC 3688,
              January 2004, <http://www.rfc-editor.org/info/rfc3688>.

   [RFC3693] Cuellar, J., Morris, J., Mulligan, D., Peterson, J., and
              J. Polk, "Geopriv Requirements", RFC 3693, February 2004,
              <http://www.rfc-editor.org/info/rfc3693>.

   [RFC4119] Peterson, J., "A Presence-based GEOPRIV Location Object
              Format", RFC 4119, December 2005,
              <http://www.rfc-editor.org/info/rfc4119>.

   [RFC5139] Thomson, M. and J. Winterbottom, "Revised Civic Location
              Format for Presence Information Data Format Location
              Object (PIDF-LO)", RFC 5139, February 2008,
              <http://www.rfc-editor.org/info/rfc5139>.

   [RFC5491] Winterbottom, J., Thomson, M., and H. Tschofenig, "GEOPRIV
              Presence Information Data Format Location Object (PIDF-LO)
              Usage Clarification, Considerations, and Recommendations",
              RFC 5491, March 2009,
              <http://www.rfc-editor.org/info/rfc5491>.

   [RFC6225] Polk, J., Linsner, M., Thomson, M., and B. Aboba, Ed.,
              "Dynamic Host Configuration Protocol Options for
              Coordinate-Based Location Configuration Information", RFC
              6225, July 2011, <http://www.rfc-editor.org/info/rfc6225>.

   [RFC6280] Barnes, R., Lepinski, M., Cooper, A., Morris, J.,
              Tschofenig, H., and H. Schulzrinne, "An Architecture for
              Location and Location Privacy in Internet Applications",
              BCP 160, RFC 6280, July 2011,
              <http://www.rfc-editor.org/info/rfc6280>.

10.2. Informative References

   [Convert]   Burtch, R., "A Comparison of Methods Used in Rectangular
                to Geodetic Coordinate Transformations", April 2006.

   [GeoShape] Thomson, M. and C. Reed, "GML 3.1.1 PIDF-LO Shape
                Application Schema for use by the Internet Engineering
                Task Force (IETF)", Candidate OpenGIS Implementation
                Specification 06-142r1, Version: 1.0, April 2007.

   [ISO.GUM]   ISO/IEC, "Guide to the expression of uncertainty in
                measurement (GUM)", Guide 98:1995, 1995.

   [NIST.TN1297]
                Taylor, B. and C. Kuyatt, "Guidelines for Evaluating and
                Expressing the Uncertainty of NIST Measurement Results",
                Technical Note 1297, September 1994.

   [RFC5222]   Hardie, T., Newton, A., Schulzrinne, H., and H.
                Tschofenig, "LoST: A Location-to-Service Translation
                Protocol", RFC 5222, August 2008,
                <http://www.rfc-editor.org/info/rfc5222>.

   [RFC6772]   Schulzrinne, H., Ed., Tschofenig, H., Ed., Cuellar, J.,
                Polk, J., Morris, J., and M. Thomson, "Geolocation Policy:
                A Document Format for Expressing Privacy Preferences for
                Location Information", RFC 6772, January 2013,
                <http://www.rfc-editor.org/info/rfc6772>.

   [Sunday02] Sunday, D., "Fast polygon area and Newell normal
                computation", Journal of Graphics Tools JGT, 7(2):9-13,
                2002.

   [TS-3GPP-23_032]
                3GPP, "Universal Geographical Area Description (GAD)",
                3GPP TS 23.032 12.0.0, September 2014.

   [Vatti92]   Vatti, B., "A generic solution to polygon clipping",
                Communications of the ACM Volume 35, Issue 7, pages 56-63,
                July 1992,
                <http://portal.acm.org/citation.cfm?id=129906>.

   [WGS84]   US National Imagery and Mapping Agency, "Department of
                Defense (DoD) World Geodetic System 1984 (WGS 84), Third
                Edition", NIMA TR8350.2, January 2000.

Appendix A. Conversion between Cartesian and Geodetic Coordinates in
             WGS84

   The process of conversion from geodetic (latitude, longitude, and
   altitude) to ECEF Cartesian coordinates is relatively simple.

   In this appendix, the following constants and derived values are used
   from the definition of WGS84 [WGS84]:

      {radius of ellipsoid} R = 6378137 meters

      {inverse flattening} 1/f = 298.257223563

      {first eccentricity squared} e^2 = f * (2 - f)

      {second eccentricity squared} e'^2 = e^2 * (1 - e^2)

   To convert geodetic coordinates (latitude, longitude, altitude) to
   ECEF coordinates (X, Y, Z), use the following relationships:

      N = R / sqrt(1 - e^2 * sin(latitude)^2)

      X = (N + altitude) * cos(latitude) * cos(longitude)

      Y = (N + altitude) * cos(latitude) * sin(longitude)

      Z = (N*(1 - e^2) + altitude) * sin(latitude)

   The reverse conversion requires more complex computation, and most
   methods introduce some error in latitude and altitude. A range of
   techniques are described in [Convert]. A variant on the method
   originally proposed by Bowring, which results in an acceptably small
   error, is described by the following:

      p = sqrt(X^2 + Y^2)

      r = sqrt(X^2 + Y^2 + Z^2)

      u = atan((1-f) * Z * (1 + e'^2 * (1-f) * R / r) / p)

      latitude = atan((Z + e'^2 * (1-f) * R * sin(u)^3)
      / (p - e^2 * R * cos(u)^3))

      longitude = atan2(Y, X)

      altitude = sqrt((p - R * cos(u))^2 + (Z - (1-f) * R * sin(u))^2)

   If the point is near the poles, that is, "p < 1", the value for
   altitude that this method produces is unstable. A simpler method for
   determining the altitude of a point near the poles is:

      altitude = |Z| - R * (1 - f)

Appendix B. Calculating the Upward Normal of a Polygon

   For a polygon that is guaranteed to be convex and coplanar, the
   upward normal can be found by finding the vector cross product of
   adjacent edges.

   For more general cases, the Newell method of approximation described
   in [Sunday02] may be applied. In particular, this method can be used
   if the points are only approximately coplanar, and for non-convex
   polygons.

   This process requires a Cartesian coordinate system. Therefore,
   convert the geodetic coordinates of the polygon to Cartesian, ECEF
   coordinates (Appendix A). If no altitude is specified, assume an
   altitude of zero.

   This method can be condensed to the following set of equations:

      Nx = sum from i=1..n of (y[i] * (z[i+1] - z[i-1]))

      Ny = sum from i=1..n of (z[i] * (x[i+1] - x[i-1]))

      Nz = sum from i=1..n of (x[i] * (y[i+1] - y[i-1]))

   For these formulae, the polygon is made of points
   "(x[1], y[1], z[1])" through "(x[n], y[n], x[n])". Each array is
   treated as circular, that is, "x[0] == x[n]" and "x[n+1] == x[1]".

   To translate this into a unit-vector; divide each component by the
   length of the vector:

      Nx' = Nx / sqrt(Nx^2 + Ny^2 + Nz^2)

      Ny' = Ny / sqrt(Nx^2 + Ny^2 + Nz^2)

      Nz' = Nz / sqrt(Nx^2 + Ny^2 + Nz^2)

B.1. Checking That a Polygon Upward Normal Points Up

   RFC 5491 [RFC5491] stipulates that the Polygon shape be presented in
   counterclockwise direction so that the upward normal is in an upward
   direction. Accidental reversal of points can invert this vector.
   This error can be hard to detect just by looking at the series of
   coordinates that form the polygon.

   Calculate the dot product of the upward normal of the polygon
   (Appendix B) and any vector that points away from the center of the
   earth from the location of polygon. If this product is positive,
   then the polygon upward normal also points away from the center of
   the earth.

      The inverse cosine of this value indicates the angle between the
      horizontal plane and the approximate plane of the polygon.

   A unit vector for the upward direction at any point can be found
   based on the latitude (lat) and longitude (lng) of the point, as
   follows:

      Up = [cos(lat) * cos(lng) ; cos(lat) * sin(lng) ; sin(lat)]

   For polygons that span less than half the globe, any point in the
   polygon -- including the centroid -- can be selected to generate an
   approximate up vector for comparison with the upward normal.

Authors' Addresses

   Martin Thomson
   Mozilla
   331 E Evelyn Street
   Mountain View, CA 94041
   United States

   EMail: martin.thomson@gmail.com

   James Winterbottom
   Unaffiliated
   Australia

   EMail: a.james.winterbottom@gmail.com