Expressing Confidence in a Location Object
draft-thomson-geopriv-confidence-04

Abstract

A confidence element is described that expresses the estimated probability that the associated location information is correct. This element conveys information that might otherwise be lost about the probability distribution represented by a region of uncertainty.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on February 28, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.
1. Introduction

Location information is often less than perfect. Two measures are used to quantify how imperfect the location information is: uncertainty and confidence. These terms, and their relationship with location information are explored in detail in [I-D.thomson-geopriv-uncertainty]. Standard forms for the expression of uncertainty are included in [RFC5491], but confidence is fixed to a value of 95%.

On the whole, a fixed definition for confidence 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 pre-existing systems - location generators might provide location information with some other confidence. Common values include 38%, 67% and 90%; all of which are prevalent in current systems. 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:
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 result in either an overly conservative estimate with high uncertainty or a overestimate of confidence.

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.

1.1. Conventions used in this document

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 relies on the definitions in [I-D.thomson-geopriv-uncertainty] and [RFC3693].

2. Representation of Confidence in PIDF-LO

The confidence element MAY be added to the "location-info" element of the Presence Information Data Format - Location Object (PIDF-LO) [RFC4119] document. This element expresses the confidence in the associated location information as a percentage.

The confidence element optionally includes an attribute that indicates the shape of the probability density function (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; 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, confidence is 95% [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.

2.1. 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.

Confidence MUST NOT be included unless location information cannot be acquired with 95% confidence.

2.2. Consuming and Presenting Confidence

The inclusion of confidence that is anything other than 95% presents a potentially difficult usability 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. [I-D.thomson-geopriv-uncertainty] describes techniques that could be appropriate for consumers that use automated processing as well as background on the issue.

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 pre-determined 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.
3. Example

The PIDF-LO document in Figure 1 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.

```xml
<pidf:presence
 xmlns:pidf="urn:ietf:params:xml:ns:pidf"
 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"
 entity="pres:alice@example.com">
   <dm:device id="sg89ab">
     <pidf:status>
       <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>
         </gp:location-info>
         <gp:usage-rules/>
       </gp:geopriv>
       <pidf:status>
         <dm:deviceID>mac:010203040506</dm:deviceID>
       </pidf:status>
     </pidf:status>
   </dm:device>
</pidf:presence>
```

Figure 1: Example PIDF-LO with Confidence

4. 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>
   </xs:appinfo>
```
source="urn:ietf:params:xml:schema:geopriv:conf">

PIDF-LO Confidence
</xs:appinfo>
</xs:documentation source="http://www.ietf.org/rfc/rfcXXXX.txt">
<!-- [[NOTE TO RFC-EDITOR: Please replace above URL with URL of published RFC and remove this note.]] -->
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:restriction base="xs:decimal">
    <xs:minExclusive value="0.0"/>
    <xs:maxExclusive value="100.0"/>
  </xs:restriction>
</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>

5. IANA Considerations

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

This section registers a new XML namespace, "urn:ietf:params:xml:ns:geopriv:conf", as per the guidelines in [RFC3688].

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

[[NOTE TO IANA/RFC-EDITOR: Please update RFC URL and replace XXXX with the RFC number for this specification.]]

<p>See <a href="[[RFC URL]]">RFCXXXX</a>.</p>
</body>
</html>
END
```

5.2. XML Schema Registration

This section registers an XML schema as per the guidelines in [RFC3688].


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

Schema: The XML for this schema can be found as the entirety of Section 4 of this document.

6. Security Considerations

The security (and privacy) implications related to adding this information are not significant.

7. References

7.1. Normative References


7.2. Informative References


[TS-3GPP-23_032] 3GPP, "Universal Geographic Area Description (GAD)", 3GPP TS 23.032 11.0.0, September 2012.

Author’s Address

Martin Thomson
Microsoft
Mountain View, VA  94043
US
EMail: martin.thomson@gmail.com