Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 27, 2007

R. Barnes
M. Lepinski
R. Watro
BBN Technologies

October 24, 2006

Secure Location Objects
draft-barnes-geopriv-secure-location-object-00

Status of this Memo

Copyright Notice

Abstract

Protection of location information is an essential requirement of the GEOPRIV architecture. Since using protocols cannot be relied upon to provide adequate protections to the location objects they carry, the location objects themselves must be secured. This document examines several candidates for a Secure Location Object format in the context of GEOPRIV and ECRIT security requirements, including both locations by value and by reference.

Requirements Language

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 RFC 2119 [RFC2119].

Table of Contents

1. Introduction

The security of location objects as they are stored and transmitted over the Internet is an essential enabler of the GEOPRIV architecture. Without the ability to guarantee the confidentiality of transmitted location information, an eavesdropper can circumvent GEOPRIV privacy rules; without authenticity and integrity, a third party can forge location information and degrade the reliability of the entire architecture. Even though such security guarantees may at times need to be relaxed -- for instance, in an emergency calling -- it is essential that the components of the GEOPRIV architecture make available a suite of security services.

The fundamental unit of location information in the GEOPRIV architecture is the Location Object (LO): Location Objects are constructed by Location Generators; stored, transformed, and forwarded by Location Servers; and consumed by Location Recipients. In practice, these different entities are often under separate administrative domains, joined by various relationships and separated by diverse networks. The sensitive location information contained in these LOs is thus stored and transmitted in a wide variety of circumstances, and faces the risk of unauthorized access at several points, even in the presence of privacy rules.

While the GEOPRIV architecture does make certain requirements of "using protocols" -- protocols that read or modify LI -- it explicitly makes no specifications with regard to protocols used only for transport of LOs. There are several transport mechanisms currently defined for conveying LOs from one point to another: Extensions have been defined for SIP, DHCP, and RADIUS, and others are being developed. Of these, only SIP has any claim of secure transport, and even the mechanisms that SIP offers are either seldom implemented (S/MIME) or widely regarded as ineffective (SIPS).

The security of LOs used in GEOPRIV thus cannot rely on the security of the underlying transport, but must instead by enforced by security features inherent in the location object itself. We call location objects with such features Secure Location Objects, or SLOs. Below, we summarize the threats to location information that have been identified by the ECRIT and GEOPRIV working groups to date, identify some additional design constraints that must be taken into account by an SLO design, and examine the feasibility of several possible SLO formats in light of these conditions.

2. Security Threats

Prior work by both the ECRIT and GEOPRIV working groups has

identified a set of risks facing location information, which are summarized here as the risks of unauthorized disclosure, forgery, and denial of location-based services. While these risks have special importance in the context of emergency calling, they apply broadly to the GEOPRIV architecture as a whole.

2.1. Disclosure of Location Information

The most immediate threat to location information is the disclosure of sensitive location information to unauthorized parties. In particular, this allows an eavesdropper to circumvent any privacy rules that are supposed to govern the use of a LO. Such disclosure could occur in either of the following ways:

o An attacker located on the path of communication to a legitimate location recipient may intercept a LO and extract sensitive location information. For example, an attacker could be a compromised router on the public internet who scans incoming packets for insecure LOs. Alternatively, such an attacker could be a compromised proxy server for some location-using protocol such as SIP

o An attacker may impersonate a legitimate location recipient in order to obtain a LO from location server. For example, an attacker could attack the LoST service to cause legitimate users to send LOs to an attacker-controlled URI. Alternatively, an attacker could supply forged credentials to a location server to request a LO stored on the server.

This threat is discussed in [Geopriv-Threats], Section 4.1.1 as well as [Geopriv-L7], Section 10.2. In the setting of emergency services, this threat is especially pertinent because the requester of emergency services is often in a vulnerable position and hence location information is particularly sensitive. For example, a motorist injured in an accident on a deserted highway is an easy target for robbery. The disclosure of location information in the event of an emergency is discussed in Security Threats and Requirements for Emergency Call Marking and Mapping [Ecrit-Threats] Section 5.2.3.

2.2. Use of Fake Location Information

A second threat to location information is the ability of an attacker to create LOs that contain false location information, i.e., a location that does not correspond to a target's true location. There are several scenarios in which this might be accomplished:

- o An attacker could replay location information corresponding to the attacker's true location at an earlier point in time. For example, an attacker who visits New York City in January could obtain a legitimate LO indicating his location in New York. The attacker could then use this previously obtained LO to claim he is in New York the following summer.

- o An attacker could replay location information corresponding to another target's location at an earlier point in time. For example, an attacker who has never been to New York City may obtain a legitimate LO that indicates that Alice is in New York City. The attacker could then later use this LO to claim that he is in New York. Note that this attack may or not require the attacker to be able to extract information from the LO.

- o An attacker could from scratch generate a forged LO corresponding to an arbitrary location. For example, an attacker could forge a LO indicating that he is in New York City. The attacker could then pass this LO off as legitimate and claim he is in New York even though he doesn't know of anyone who has been there.

An extensive discussion of this threat appears in [Geopriv-L7], Section 8. The threat is also discussed in also discussed in [Geopriv-Threats] Section 4.1.2. In the setting of emergency services, this threat is especially pertinent because fake location information can be used to force scarce emergency-response resources to be improperly allocated. For example, in the event of a disaster, when the demand for emergency services exceeds supply, an attacker using a fake location could cause an ambulance to be sent to a location where no one is present. Even worse, an attacker could generate huge numbers of emergency phone calls from all over the world that all claim to be from a particular city. Such an attack could easily overwhelm the public safety access point and deny emergency services to legitimate residents of the city. The use of fake location information as it pertains to emergency services is discussed briefly in [Ecrit-Threats] Section 5.2.3.

2.3. Denial of Location-based Services

In a similar vein, an attacker could deny location-based services to an individual with legitimate need for these services. In the current, unsecured architecture, this could occur in several ways:

- o As discussed in Section 2.2, an attacker can deny access to location-based services by using forged location information to overwhelm location-based services for a particular location.

- An attacker could alter a location object to indicate an incorrect location of a target. For example, when a legitimate user attempts to access a location-based service, an attacker on the path of communication may alter the location object to cause the request to be routed to the wrong service provider.

- An attacker could prevent a location recipient from obtaining a location information. For example, an attacker might prevent a service provider from dereferencing a reference to a location object making it difficult for the user to receive location-appropriate services.

This threat is discussed in the context of emergency services in [Ecrit-Threats] Section 5.2.2, but is equally applicable to other uses of the GEOPRIV architecture.

3. Design Considerations

In addition to the security objectives discussed in Section 2 above, there are several other constraints on the design of SLOs. In general, use of SLOs must be compatible with other GEOPRIV and ECRIT protocols and architectures, with minimal modifications to either. One particular issue that is likely to arise is that in several possible SLO designs, the user (e.g., a UAC in the SIP model) may not have access to his own location (by value).

3.1. Integration with LoST

Currently, many of the proposed use cases for the Location-to-Service Translation Protocol (LoST [LoST]) assume that the UAC has access to its own location. For instance:

- As currently defined, a LoST query includes the location of the target in GML. In the case that the UAC is both the target of the query and the originator of the query, this means that the UAC must have access to its own location.

- LoST responses typically contain a service boundary so that a UAC can determine when it has left the region in which its current PSAP URI is valid, and thus must query the LoST server again. Such a boundary is useful only if the UAC is aware of its own location.

A security architecture in which a UAC may not know its own location, may necessitate revisions to the way that LoST is used. For instance, access networks or VoIP providers may need to provide LoST proxies that have access to location information.

3.2. GEOPRIV Considerations

It is as yet unclear which communications in the GEOPRIV architecture can or should make use of SLOs when exchanging location data. The areas of application likely will be determined by the number of each GEOPRIV entity and the contractual, regulatory, or other trust relationships among them; these considerations also will affect the trust model underlying a SLO design. In addition, the functions required of each entity in the GEOPRIV architecture will dictate certain levels of access, and SLOs must accommodate these requirements. For instance, since the GEOPRIV architecture specifies that privacy rules are applied by a Location Server, use of SLOs must not preclude the Location Server from having sufficient knowledge of location information to apply these rules.

3.3. ECRIT Considerations

A critical usage of GEOPRIV location objects is in emergency services, both for routing emergency calls to the correct PSAP and for directing emergency services to the location of the emergency. In such situations, it is essential that all relevant location information be available to all emergency responders that require it. When adding confidentiality features to a location object, therefore, appropriate failover mechanisms must be available.

3.4. Technical Considerations

User devices that are expected to handle location objects are becoming increasingly mobile. In the context of SIP, this will be especially true as SIP is applied within cellular wireless networks. In order to facilitate the use of SLOs by these devices, an SLO design should be adaptable for use in an environment where there are constraints on both the processing power and bandwidth available to user devices. SLOs are generally amenable to such environments, since they require no cryptographic operations to be performed in order to store or transmit them securely, and, at least when expressed as locations by reference, can consume very little bandwidth and storage space.

4. Candidate Secure Location Object Formats

Following the convention of SIP Location Conveyance [SIPLocation], we broadly divide the category of SLOs (objects that can be transmitted while still maintaining certain security properties) into those based on location by value and those based on location by reference. We then discuss candidate SLOs in both categories and discuss the properties of a trust model relied upon by any SLO. Note, however,

that in spite of this division, these two types of SLO can be naturally combined, for instance multi-layer access control could be achieved by using a secured location reference to refer to a secured location value.

4.1. Location By Value

Conveyance of location by value is the act of transmitting an entire LO, rather than a reference to it. Security properties can be added to such a location object by either signing the location object, encrypting it, or both, using a technology such as S/MIME or XMLDsig. Each type of object -- signed, encrypted, or both -- has a different set of security properties, discussed below.

Although we separately discuss the signing and encrypting of LOs, it is natural to consider combining the two approaches. This raises the question of whether a LO should be first signed and then encrypted, or vice-versa. We therefore briefly discuss the advantages of both approaches.

- In settings where denial of service attacks are likely, signing an already encrypted LO is advantageous because a recipient of such LOs can quickly discard LOs with invalid signatures without needing to spend resources decrypting the object. Additionally, in settings where some fields of the LO should be encrypted and other fields should be left unencrypted, it is advantageous to sign the entire LO after the private fields have been encrypted.

- The use of objects that are first signed and then encrypted requires less work on the part of the location producer. Indeed, a location producer may produce a single signed object which can then be encrypted, by a separate party, for delivery to multiple recipients. Similarly, the recipient of such an LO can re-encrypt the signed object for delivery to a new recipient without the involvement of the location producer. Additionally, in settings where different portions of an LO should be signed by different entities, it is advantageous to first sign and then encrypt the LO.

4.1.1. Signed Location

A "signed location" SLO consists of a LO together with a signature of some or all of the LO by a recognized authority, likely a Location Server or Location Generator. Use of a signed location SLO has the following security implications:

- Perhaps most importantly, use of a signed location SLO mitigates all of the threats in Section 2.2 arising from the use of forged

location information. An attacker who is not an authorized signer in the underlying trust model is unable to create fake location objects. In particular, this prevents an attacker from claiming he is at a distant location. Additionally, if a timestamp is included in the signed object, an attacker cannot replay a previously obtained location object (either his own or someone else's).

- Use of a signed location SLO partially mitigates the threats in Section 2.3 regarding denial of service. An attacker on the communication path between a user and a location-based service provider cannot alter the location object in transit to make it appear that the user is at a distant location. Obviously, certain attackers on the communication path can always deny service to an individual by dropping the individual's request. However, an attack that drops the entire request is simpler to detect and respond to an attack that alters the location, since when a request is dropped the user finds out immediately (as soon as the time-out fails) but if a location is altered it is unclear how long it will take to determine that a problem has occurred.

Naturally, the security properties granted by use of signed location SLOs fundamentally rely on a suitable trust model; as discussed in section 4.3, development of this trust model is a nontrivial but tractable problem. In order for signed location to be useful, it must be difficult for an attacker to compromise an authorized signer of location information. When signed location SLOs are used, it is the responsibility of the using protocol to take appropriate action when the signature fails to verify. For example, in most cases, a signed SLO with an invalid signature might be discarded altogether, but in the special case of emergency services, a call with a location signature that fails to verify might be answered but given lower priority than calls with valid SLOs.

4.1.2. Encrypted Location

An "encrypted location" SLO is a LO encrypted in such a way that it is readable only by its intended recipient(s). Use of an encrypted location SLO has the following security implications:

- Use of an encrypted location SLO mitigates all of the threats in Section 2.1 arising from improper disclosure of location information. Unless the attacker is able to compromise the secret decryption key of the intended location recipient, it is infeasible for him to extract information from any encrypted location SLO he might obtain. Therefore, even if the attacker is able, for example, to compromise a proxy on the communication path to a location recipient the sensitive location information

contained in the SLO remains private.

- Use of an encrypted location SLO only partially mitigates the threats in Section 2.2 regarding forgery of location information. Depending on the key distribution architecture, it may be possible for an attacker to obtain the encryption key of a legitimate location recipient and forge an encrypted location SLO. Of course, these threats can be mitigated (as described in Section 4.1.1) by combining signing and encrypting of location objects. On the other hand, because an eavesdropper does not have access to the information contained in an encrypted location SLO, it is very difficult for him to modify the location in transit.

- Use of an encrypted location SLO can pose additional risks regarding the denial of service threats discussed in Section 2.3. In particular, use of an encrypted location SLO introduces the possibility that a user is denied a service because the service provider cannot decrypt the SLO to extract LI. This could occur because of a key-management error, or because of an attack on the mechanism used to distribute public keys.

The use of encrypted location SLOs relies fundamentally on a reliable mechanism to distribute the keys belonging to legitimate service providers; the difficulty of this task will derive from the underlying trust model. In the context of emergency services, for example, one might use the LoST protocol to return a certificate for a PSAP in addition to the PSAP's URI. This reduces the incremental risk of using encryption, since an attacker who is able to use LoST to distribute incorrect public keys can surely disrupt emergency services in other ways.

One often-mentioned advantage of location-by-reference is that the required dereference operation creates an opportunity for location providers to enforce a scheme in which the party dereferencing the URL pays the provider for the location. A secondary advantage of encrypted location SLOs is that they can be used to extend this model to location by value: The encrypted location object can be transmitted to the location recipient, but encrypted in such a way that the SLO cannot be used by the recipient until he performs a second decryption or key exchange transaction with the location provider. However, just as the by-reference payment scheme is viable only if a user cannot dereference a URL to obtain his own location, this model forces a transaction only if a user cannot decrypt an encrypted location SLO containing his location. While this may force some adaptation of existing protocols (as discussed in Section 3), it seems that use of encrypted location SLOs for this purpose is still consistent with broader usage. For example, LoST servers could be operated by entities that maintain business relationships with

location providers, so that encrypted location SLOs included in LoST queries could be decrypted.

4.2. Location By Reference

Conveyance of location by reference is the act of transmitting not an object containing LI, but rather a URL (or other pointer) that can be dereferenced to obtain a LO. Location URLs have several important security implications:

- Perhaps most importantly, use of location by reference forces a location recipient to conduct a separate transaction in order to obtain the desired LI, which has the effect of allowing any security decisions to be delayed until the time when a location URL is dereferenced. This property allows much more complicated security and privacy policies to be enforced at the location server (such as rules about location expiration and retransmission), rather than delegating trust to using protocols. At the same time, however, it also lends itself very naturally to failover, since the location server can make a decision to grant access to parties that can demonstrate a need and authority for access, such as emergency service providers.

- Because a location URL references a resource held by a third party (commonly, a location server), not by the location target or location recipient, location references cannot be constructed by a user, but rather must be obtained from an location server. This yields very powerful anti-forgery (hence anti-spam) properties, since a user cannot forge a location URL that references LI indicating that he is elsewhere than he is, and likewise, a third party (e.g., a man in the middle) cannot modify a URL to deny location-based services.

We call a location reference that employs one or more security protocols in its dereference a secured location reference. Any security protocols used in conjunction with location references will be reliant on a suitable trust model; as discussed in section 4.3, development of this trust model seems to be a nontrivial, but tractable problem. In order for secured location references to be suitable for use in emergency services, the dereferencing protocol and any security protocols employed between the recipient and the location server must be made sufficiently reliable for use in an emergency. As is the case with normal, unsecured location references, the most significant risk is introduced by the dereferencing protocol, since the location server is capable of granting access to LI independent of security policies and protocols.

4.3. Trust Models

Any SLO system will be based on an underlying trust model. The structure of this model deeply influences the nature of the security guarantees that the SLO system can provide. Such guarantees include:

- Authentication of location recipients: Use of SLOs offers another mechanism for authenticating identities referenced by privacy rules. Using secure location by value, objects can be encrypted for a specific recipient, and using secure location by reference, a location server can interpret a cryptographic credential to grant or deny access to specific recipients. In particular, emergency service providers could be unambiguously identified by their credentials to be assured access to the LI they require.

- Authentication and integrity of location information: In the PSTN, location information is provided by wireline or wireless operators, and thus assumed by all using parties to be reliable. Use of location signing in secure location objects provides a mechanism to translate these assurances to IP-based telephony and other location-based Internet services.

- Non-repudiation of location information: By the same token as above, the current PSTN architecture allows failures of the location architecture to be clearly attributed to the provider of faulty LI, for purposes of determining regulatory or civil liability. For location-based services over IP, particularly emergency services, this is an important function that can be enabled by secure location objects.

These features require the identification and issuance of credentials for two classes of entities: Location producers and location consumers. The case of location consumers seems to be the simpler, if we envision two major use cases, (1) emergency services calling, and (2) client-server style location-based services. In the former case, the set of PSAPs and emergency service providers is small and stable enough to be manageable. In some cases, even further simplification will be possible: In the NENA i3 architecture, for example, one might need to manage credentials only for gateways between emergency services networks and the Internet. For other client-server location-based services, there are several current trust models that could be adapted, such as the broad, flat PKI model used by HTTPS or the more flexible model used in DNSSEC.

Constructing a system for authenticating location producers is more difficult. For example, an organization that administers a corporate network of SIP-based desk phones might provision these phones with fine-grained location information, such as their floor and room

number. By some calculations [ref:Henning], in New York City alone there are thousands of organizations that might be expected to do become location producers in this way, several of which appear and disappear each day. On the other hand, such location producers need only be included in a trust model if the goal of this trust model is to provide guarantees stronger than are offered by the PSTN. An initial capability to provide PSTN-equivalent security would require only the inclusion of telecommunications and internet service providers; construction of a PKI on the scale of the former is already being under taken by the SIDR working group and the Regional Internet Registries. In addition, many VoIP providers currently outsource location determination functions to other entities, which further consolidates the set of location producers. Note also that although we have treated here the specific example of location for VoIP, the same access networks that provide location for these services will be used to access other location-based services, so a trust model for location producers in a VoIP setting would be extensible to a model for more general location-based services.

5. Conclusions

The purpose of this document is to start a discussion about the requirements of GEOPRIV and ECRIT for security in location objects. Clearly, it is tempting to push responsibility for security onto the protocols that carry LOs and the using protocols that process them. Currently, however, these protocols are unable to provide the end-to-end security guarantees necessary to mitigate threats to the privacy of location information and the integrity of location-based services. Without securing the location object itself, entities that generate LOs have no assurances that the LOs will not be misused, and critical applications such as emergency services have no assurances that the LOs they receive have not been forged or otherwise tampered with.

In this document, we considered three approaches to securing LOs. Signing a location object can prevent forgery and mitigate resulting denial of service attacks. Encrypting location objects can prevent the improper disclosure of location information, but encryption results in an opaque location object that may require adaptation of using protocols. Using location by reference in conjunction with a secure dereferencing transaction can prevent both forgery and improper disclosure of location information. However, obtaining location information from a location-by-reference object requires an additional transaction that could introduce additional risk in time-critical applications such as emergency services. Naturally, these techniques for securing location objects can be combined to obtain stronger security guarantees or increased robustness. For example, a location reference could be appended to a signed and encrypted

location object to obtain the security guarantees of location-by-reference and yet require a separate de-referencing transaction only in the event that decryption fails. Achieving security through any of these mechanisms will require an appropriate trust model.

6. IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

7. Security Considerations

The focus of this document is security; hence security considerations permeate this specification.

8. Acknowledgements

9. References

9.1. Normative References

[RFC2119] "", 2005.

9.2. Informative References

[Ecrit-Threats]
Nortel, Siemens, Columbia University, and Siemens, "Security Threats and Requirements for Emergency Call Marking and Mapping", July 2006.

[Geopriv-L7]
Siemens Networks and Columbia University, "Geopriv Layer 7 Location Configuration Protocol; Problem Statement and Requirements", October 2006.

[Geopriv-Threats]
Technology and Public Policy Clinic, Technology and Public Policy Clinic, Center for Democracy and Technology, and NeuStar, "Threat Analysis of the Geopriv Protocol", February 2004.

[LoST] Qualcomm, Inc., SunRocket, Columbia University, and

Siemens, "LoST: A Location-to-Service Translation Protocol", September 2006.

[RFC3693] Siemens AG, Center for Democracy and Technology, Technology and Public Policy Clinic, NeuStar, and Cisco, "Geopriv Requirements", February 2004.

[SIPLocation] Qualcomm, Inc. and SunRocket, "Session Initiation Protocol Location Conveyance", October 2006.

Authors' Addresses

Richard Barnes
BBN Technologies
9861 Broken Land Pkwy
Columbia, Maryland 21046
USA

Phone: +1-410-290-6169
Email: rbarnes@bbn.com

Matt Lepinski
BBN Technologies
10 Moulton St.
Cambridge, Massachusetts 02138
USA

Phone: +1-617-873-5939
Email: mlepinsk@bbn.com

Ron Watro
BBN Technologies
10 Moulton St.
Cambridge, Massachusetts 02138
USA

Phone: +1-617-873-2551
Email: rwatro@bbn.com

Full Copyright Statement

Intellectual Property

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).