ECRIT
Internet-Draft
Expires: November 10, 2005

H. Tschofenig
Siemens
H. Schulzrinne
Columbia U.
M. Shanmugam
TUHH

May 9, 2005

Security Threats and Requirements for Emergency Calling
draft-tschofenig-ecrit-security-threats-00.txt

Status of this Memo

Copyright Notice

Abstract

With the increasing interest to replace parts of the public switched telephone network (PSTN) with its IP-based counterpart the functionality of emergency services also needs to be offered using IP-based technologies. Since the PSTN and the Internet follow different design principles, their architecture is quite different.

This fact has to be considered and security threats for an IP-based emergency environment have to be re-evaluated. This document investigates the potential threats for the end hosts and the infrastructure that aims to support emergency services.

Table of Contents

1. Introduction

This document provides an overview of security mechanisms and motivations for using them in the VoIP-based emergency services. PSTN users can summon help for emergency services such as ambulance, fire and police using a well known unique number (e.g., 911 in North America, 112 in in Europe). With the introduction of IP-based telephony support for emergency service also has to be provided. A number of protocols and protocol extensions need to interwork in order to provide emergency functionality.

Since the Internet is hostile place, it is important to understand the security threats for emergency services. Otherwise, an adversary can use the infrastructure to place fraudulent calls, mount denial of service attacks, etc.

This document focuses on the security threats and security requirements for the IP-based emergency service infrastructure only without interaction with PSTN infrastructure elements.

A few discussions within this document are related to emergency handling but solutions will not be developed as part of the ECRIT working group. Hence, the are included mainly for completeness and to point to the need to investigate additional aspects. Depending on the chosen protocols (for the emergency call itself, for directory access related to emergency call routing, for obtaining location information from the network, etc.) various solutions might also already be available to fulfill these security requirements and to address the threats appropriately.

This document is organized as follows: Section 2 describes basic terminology, Section 4 illustrates security threats and Section 5 lists security requirements.

2. 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].

Emergency Caller, Public Safety Answering Point (PSAP), Access Infrastructure Provider, Application (Voice) Service Provider, Emergency Call Taker, etc. is taken from [I-D.schulzrinne-ecrit-requirements].

Additionally, we use the following terms throughout the document:

Emergency Call Routing Support: This term refers to entities that route the emergency call to the appropriate PSAP based on information like location information, language, etc. If SIP is used as a protocol for session setup and call routing, for example, then this entity would correspond to a SIP proxy.

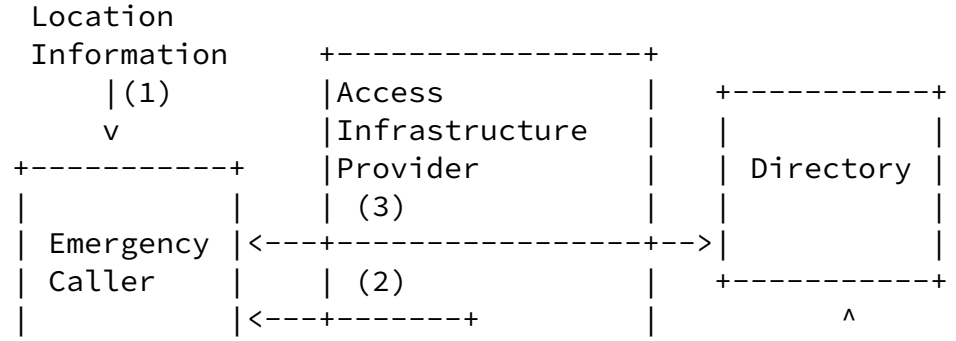
Directory: This entity refers to a distributed directory protocol. DNS is one example of such as distributed directory but there are other protocols that might fulfill the requirements listed in [I-D.schulzrinne-ecrit-requirements] for such a protocol.

Asserted Location Information: The term asserted location information refers to the property that the recipient of such an object is able to verify that it was generated by a particular party that is authorized todo so.

3. Basic Actors

In order to support emergency services covering a large physical area various infrastructure elements are necessary: Access Infrastructure Providers, Application (Voice) Service Provider, PSAPs as endpoints for emergency calls, directory services or other infrastructure elements that assist in during the call routing and potentially many other entities.

This section outlines which entities will be considered in the threat analysis and shows the high level architecture.



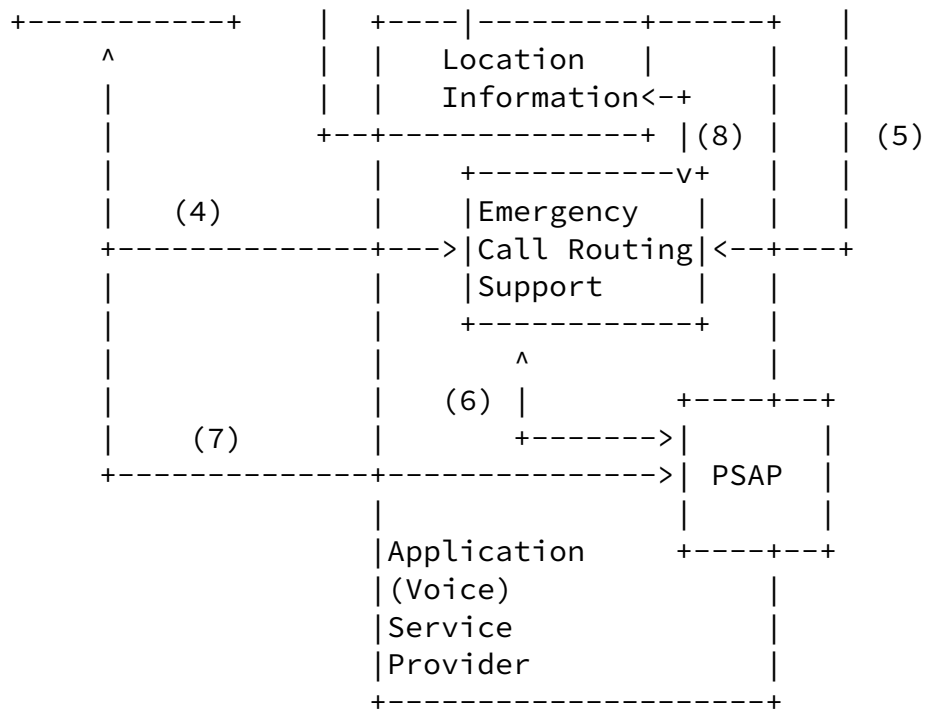


Figure 1: Framework

Figure 1 shows the interaction between the entities involved in the call. There are a number of different deployment choices, as it can be easily seen from the figure. The following deployment choices need to be highlighted:

- How is location information provided to the end host? It might either be known to the end host itself (due to manual configuration or provided via GPS) or available via a third party. Even if location information is known to the network it might be made available to the end host. Alternatively, location information is used as part of call routing and inserted by intermediaries.

- Is the Access Infrastructure Provider also the Application (Voice) Service Provider? In the Internet today these roles are typically provided by different entities. As a consequence, the Application (Voice) Service Provider is typically not able to learn the physical location of the Emergency Caller.

Please note that the overlapping squares aim to indicate that certain

functionality can be collapsed into a single entity. As an example, the Application (Voice) Service Provider might be the same entity as the Access Infrastructure Provider and they might also operate the PSAP. There is, however, no requirement that this must be the case. Additionally it is worth pointing out that end systems might be its own VoSP, e.g., for enterprises or residential users.

Below, we describe various interactions between the entities shown in Figure 1 are described:

- (1) Location information might be available to the end host itself.

- (2) Location information might, however, also be obtained from the Access Infrastructure Provider (e.g., using DHCP or application layer signaling protocols).

- (3) The Emergency Caller might need to consult a directory to determine the PSAP that is appropriate for the physical location of the emergency caller (and considering other attributes such as a certain language support by the Emergency Call Takers).

- (4) The Emergency Caller might get assistance for emergency call routing by infrastructure elements (referred as Emergency Call Routing Support entities). In case of SIP these enities are proxies.

- (5) Individual Emergency Call Routing Support entities might need to consult a directory to determine where to route the emergency call.

- (6) The Emergency Call Routing Support entities need to finally forward the call, if infrastructure based emergency call routing

  is used.

- (7) The emergency caller might interact directly with the PSAP without any Emergency Call Routing Support entities.

4. Security Threats

This section discusses various security threats related to emergency call handling.

4.1 Denial of Service Attacks

A (distributed) denial-of-service attack (DoS attack) on a PSAP, for example, might make the PSAP unreachable for emergency calls. Since a particular PSAP is responsible for a certain geopraphical area, the entire area might be affected (if no other backup PSAP is available). DoS attacks might appear in many different flavors ranging from standard SYN flooding attacks to attacks where a human operator is involved and needs to determine whether a call is in fact a true emergency call. In some cases this might lead the case where the emergency staff (police, ambulance, etc.) might need to rush to the indicated emergency scene (potentionally an arbitrary location) and will therefore not be available for other rescue assignments during that time.

As such, PSAPs can be seen as a particularly valuable target since the consequences of an unreachable PSAP has severe consequences.

Attacks against the routing infrastructure enables an adversary to prevent all nodes attached to this network to sent emergency calls. Attacks against entities that assist in the call routing (such as attacks against the directory service) might make it difficult or impossible for emergency call to reach its intended PSAP.

4.2 Call Identity Spoofing

If an adversary is able to make emergency calls without the need to disclose its identity (such as a SIP URI or NAI) then prank calls cannot be traced back. If the call is proxy-routed, the PSAP will not see the IP address of the caller in signaling. Additionally, it might be necessary for the Emergency Call Taker to initiate a voice, video or instant messaging exchange towards the Emergency Caller.

Trying to find an adversary that placed a crank call is difficult if somebody uses an open 802.11 access point, even if you can find the owner of that access point. This problem is no different than somebody placing an emergency call from a payphone.

If the adversary is never authenticated (neither to the PSAP nor to the Access Infrastructure Provider) then it is possible to trace the call back to a make a particular entity accountable.

A standard requirement for emergency systems is that emergency calls

must also be placed in absence of any authentication. An adversary will typically exploit these weaknesses and he will always find networks that do not perform network access authentication of the user prior to providing network access. As such, the emergency infrastructure cannot neither rely on network access authentication nor on authentication of the caller towards the PSAP or the Application (Voice) Service Provider.

It is necessary to point to the fact that authentication in the emergency case might require the authorization procedure to be skipped. For example, in an emergency case it is still possible to authenticate the user of an emergency call but without considering that its credits are exhausted.

4.3 Location Spoofing

An adversary might want to made-up faked location information in order to fool the emergency personnel. This is made particularly easy if the location information is provided by the Emergency Caller either via manual configuration or via GPS. Spoofing is more difficult if an entity proving Emergency Call Routing Support inserts location information into emergency call signaling. In this case the adversary needs to route the call via some intermediaries. This is possible since these devices are often, by their nature as IP devices, addressable from an arbitary physical location. The usage of VPN (or other tunneling mechanisms) and proxies further complicates the ability to infer the physical location from the IP address seen by the PSAP.

4.4 Impersonating a PSAP

An adversary might pretend to operate a PSAP. When either an end host or an intermediate device wants to determine the PSAP that is responsible for a particular geographical area by sending a query to the directory an adversary might return a faked response. Returning an incorrect response message does not require the adversary to be somewhere along the path. It is sufficient for an adversary to be located in a broadcast medium and the adversary has to reply as soon as a query is observed (if no security protection is utilized). If the response indicates a legitimate but inappropriate (i.e., a PSAP that is authoritive for a different geographical area) then the emergency call interaction will be able to continue but will suffer from delays until the emergency call can be forwarded to the correct PSAP, potentially involving human interation (by the Emergency Call Taker).

4.5 Signaling Message Modification

An adversary that is located along the signaling path might modify the content of emergency calls, such as location information or identity information. This might lead to a denial of service attack against the emergency personell, disruption of the emergency call, delayed call setup, etc.

An adversary might want to inject signaling messages to terminate or redirect the call to another location. Dropping or delaying signaling messages is also possible for an on-path adversary.

Depending on the capability of the signaling protocol the range of possible attacks might have been documented already.

4.6 Modification of the Emergency Call

An adversary along the media path might want to modify the data traffic part of the emergency call (voice, video or instant message). An attacker can change the message on-the-fly and fool the PSAP to receive meaningless or bogus messages. The response messages to Emergency Caller might also be subject to change, for example by injecting a recorded failure message.

4.7 Loss of confidentiality

An adversary might eavesdrop an emergency call and use the information to future sessions as part of replay attacks. The ability to eavesdrop also allows to learn details about the emergency situation which might be of interest for the press or other media organizations. Please note that the location of the adversary is important regarding the eavesdropped area. For example, an adversary in a WLAN is typically able to see a small amount of traffic due to the coverage area of typical WLAN network.

Reavealing the true identity of the user as part of the privacy override mechanism might conflict with the users privacy settings.

4.8 Replay Attack

An adversary might want to use eavesdropped information to mount attacks in the future. This might be necessary if information cannot be re-created by the adversary (for example, asserted location information). The ability to replay messages or individual objects the specific property of these messages and objects is important. For example, asserted location information might bind location information and a timestamp with a digital signature together that makes it difficult to reuse this object beyonds its lifetime.

[Editor's Note: It is sometimes hard to tell what are real threats and what security threats are addressed already by certain solutions outside the scope of the working group. Addressing all standard security threats is a long process if certain mechanisms are required in an case that largely or completely mitigate against these threats.]

4.9 Corrupting Configuration Information

An adversary might overide all locally configured emergency numbers. This might be particular problematic if these emergency numbers are dynamically retrieved using some mechanisms. As such, an Emergency Caller would start a call that either leads to a blackhole (as such it is a DoS attack), the Emergency Caller connects to a rogue PSAP or to an inappropriate PSAP.

5. Security Requirements

[Editor's Note: A few requirements below are already addressed by a number of requirements and solution specific documents today. In order to keep the document short it would be reasonable to focus only on the difficult security threats and requiremens for emergency calls rather than enumerating everything that could happen to an emergency call. The working group should decide how to proceed with this particular issue and what threats and requirements should be elaborated in more detail.]

Compiling security requirements to address the threats listed in the previous section might is impacted by several constraints:

Security mechanisms may lead to a certain performance overhead (e.g., several roundtrips).

A certain security infrastructure is required that might lead to deployment problems. For example, end user certificates, certificates for networks, usage of authorization certificate, etc. might need to be deployed before any of these mechanisms are useful.

Many of these aspects are related to regulatory and legal requirements that may vary from country to country. Typically,

these mechanisms cannot be mandated by an IETF specification.

Some of the requirements impose solutions that are out-of-scope of the ECRIT working group.

Given the above-listed constraints the requirements that have to be addressed by work that is done within ECRIT have to be highlighted. Other requirements have to be read as 'if you would like to address this threat, then you might want to consider this requirement' rather than 'any solution must address fulfill this requirement'.

5.1 Denial of Service Attacks

It is difficult to address all possible denial of service attacks that might lead to disruption of an emergency call since a number of IETF protocols are used in order to provide this functionality. Hence, care must be taken when protocol extensions are developed that the chance for a denial of service attack is not increased. Even without using any security mechanisms (such as authentication and key exchange protocols) some degree of security has to be provided.

It is important to understand that the ability to mount DoS attacks must also be considered as part of the architecture work when legal

and regulatory requirements are known and need to be fulfilled.

5.2 Call Identity Spoofing

A standard requirement to prevent identity spoofing is to authenticate the Emergency Caller. Authentication mechanisms that require multiple roundtrips and as such might delay the call are often not desirable or cannot be mandated.

5.3 Location Spoofing

An Emergency Caller might in many cases know its own location information because it was obtained via civic or geospatial location extensions for DHCP, via manual configuration or via GPS. Unfortunately, information provided by the end host is untrustworthy particularly when it is as important as location information. Two approaches have been discussed in the past that place lead to a few requirements:

- Location Information is asserted by the Access Infrastructure Provider. As such, the end host might use GPS but uses a protocol to allow the network to assert the location information. This approach also has its limitations if the coverage area of the wireless network is fairly large.

- Location Information is added to the emergency call via an Emergency Call Routing Support entity. Depending on the protocol used for call routing and on the properties of this protocol it might be necessary to return the asserted location information to the end host since intermediate nodes might not be allowed to insert objects into the call setup messages (at least not in all parts of the messages, such as bodies). These signaling entities, in general, do not know the physiscal location of the user. Thus, they have to rely on somebody else to actually provide the location, e.g., the Access Infrastructure Provider.

As it can be seen from these two options the main difference is based on the type of protocol that is used in the message communication. This has an impact on the semantic and on the availability of certain attributes (such as identities that are used by these protocols) and on deployment constraints. Based on the observation that the Access Infrastructure Provider is closest to the end host and is therefore the most likely entity that knows something about the physical location of the end host it seems to be reasonable to assume that some entity that asserts the location information is actually available in this particular network.

The following requirements need to be provided in order for asserted

location information to accomplish its goals:

- Location Information MUST be integrity protected to prevent modifications by third parties.

- The recipient of the asserted location information object MUST be able to determine the party that asserted the location information in order to verify the assertion. As such, authentication of the asserting party (the entity that created the assertion) MUST be provided.

- The asserted location information MUST include a timestamp to limit its validity in order to reduce replay attacks.

- The recipient of the asserted location information MUST have a way to verify that the asserting party is indeed authorized to create such an assertion. As such, authentication is insufficient if not further authorization decision can be associated to the authenticated identity.

- The recipient of the asserted location information SHOULD have a mechanism to determine the Emergency Caller based on the provided assertion.

The last bullet deserves further discussion: If some information about the Emergency Caller identity has to be included then only for the purpose of tracability and this functionality might not of general use since an adversary will always find networks that do not authenticate the user prior to providing network access. Furthermore, the goal of a number of network access authentication protocols is to prevent disclosure of the user identity to entities other than to the user's home network. Note that the term 'user idenity' does not require that this identity directly points to the 'real' identity of a user. A court might want to require this identity to be resolved and to determine the user behind this identity. Even if the access network would like to assertain the user's identity as part of the asserted location information it is, in many cases, not even possible for the Access Infrastructure Provider.

If the authenticated user identity is not available to the Access Infrastructure Provider then only a few other identities might be useful, such as the IP address or the MAC address. Other identities, such as the Host Identity, might not be available since they are only used by very few protocols. An assertion that indicates the network in combination with the IP and/or MAC address (together with a timestamp) might provide some limited degree of traceability only if the user was authenticated directly to this particular network.

Providing the IP address allows some obvious attempts to cheat to be caught. Hence, there is the question whether some identity should be added at all given the potential limitations and the potential small amounts of cut-and-paste attacks. Using end user based

authentication in addition to the asserted location information would be helpful (e.g., using end user certificates) but will impose a serious deployment problem. Given the fact that emergency calls must still be allowed even without end user authentication certainly defeats the purpose of these mechanisms. A partical attempt to address some phrank calls is to classify emergency calls based on the availability of the provided attributes. If suspicious information is being provided that may well be wrong then additional verification steps need to be taken. For example, if a report of a large fire on a Manhattan street is received then the PSAP may wait to dispatch until it gets a second person to call in. This approach obviously has some limitations as well.

5.4 Impersonating a PSAP

The Emergency Caller SHOULD be able to determine conclusively that he has reached an "authorized" or "legitimate" emergency call center. This requirement is meant to address the threat that a rogue, possibly criminal, entity pretends to accept emergency calls and disrupts the emergency infrastructure. Particularly the caching properties of a distributed directy might be expoited. Typically, the following properties are assumed:

- The interaction between the directory access client and the directory access server MUST be integrity and replay protected.

- The directory access server MUST be provide data origin authentication thereby ensuring that the provided data items are indeed from the claimed source.

- The directory server MUST provide information to ensure that it is authoritive for the provided information.

Unlike in the PSTN case IP based networks provide a better opportunity to spoof a PSAP since physical access to the cable plant is required in the PSTN case, while this may not true for the IP case.

5.5 Signaling Message Modification

To protect signaling messages against modifications either individual attributes SHOULD be protected (such as location objects) or the entire signaling message communication SHOULD experience end-to-end protection. This requires integrity and replay protection to be

applied. Authentication of the data sender and the data receiver SHOULD be provided to prevent a man-in-the-middle attack.

5.6 Replay Attack

In order to protect signaling messages (or individual attributes) to be replayed in future protocol sessions integrity and replay protection mechanisms SHOULD be provided.

5.7 Loss of confidentiality

In order to prevent leakage of information exchanged during the emergency call (both signaling and data traffic) confidentiality protection SHOULD be provided. The mechanisms to accomplish this functionality are typically different for the data traffic and for the signaling messages and various scenarios, such as hop-by-hop, end-to-middle, middle-to-middle and end-to-end security, need to be considered. Particularly the key management aspects for end-to-end security mechansisms imposes a deployment burden and hence need to be critically analysed in order to determine its applicability in the given context.

5.8 Modification of the Emergency Call

To protect a man-in-the-middle attack to modify or inject data traffic into the communication between the Emergency Caller and the PSAP integrity, replay and data origin authentication SHOULD be provided. Since the signaling messages are used to authenticate the end points and to distribute the required keying material it is necessary that either the key exchange protocol itself and the signaling messages experience appropriate security protection. The term 'appropriate' refers to the given context, the used signaling protocol and the key exchange protocol.

Please note that the interactive nature of a voice communication already provides a some degree of protection. However, with the introduction of instant messaging the freshness of the Emergency Call needs to be provided by other means.

5.9 Corrupting Configuration Information

Devices SHOULD be assured of the correctness of the local emergency numbers that are automatically configured. If we assume a fixed, global emergency service identifier that requires no configuration and only configure local "traditional" emergency numbers, users are not likely to suddenly dial some random number if a rogue configuration server introduces this as an additional emergency

number. The ability to override all locally configured emergency

number is of more concern. If the Emergency Caller does not use the infrastructure to route the call to the appropriate PSAP then the security of the directory service is of importance for security.

6. Security Considerations

This document addresses security threats and security requirements. Therefore, security is considered throughout this document.

7. References

7.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", March 1997.

7.2 Informative References

[I-D.schulzrinne-ecrit-requirements] Schulzrinne, H. and R. Marshall, "Requirements for Emergency Context Resolution with Internet Technologies", May 2005.

Authors' Addresses

Hannes Tschofenig
Siemens
Otto-Hahn-Ring 6
Munich, Bayern 81739
Germany

Email: Hannes.Tschofenig@siemens.com

Henning Schulzrinne
Columbia University
Department of Computer Science
450 Computer Science Building
New York, NY 10027
USA

Phone: +1 212 939 7042
Email: schulzrinne@cs.columbia.edu
URI: http://www.cs.columbia.edu/~hgs

Murugaraj Shanmugam
Technische Universitat Hamburg-Harburg
Department of Security in Distributed applications
Harburger Schlossstrasse 20
Hamburg-Harburg 21079
Germany

Email: murugaraj.shanmugam@tuhh.de

Intellectual Property Statement

copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.