SIP Dealing with Location Based Information

Jose Costa-Requena, Haitao Tang, and Inmaculada Espigares Del Pozo

Abstract: **This paper introduces an approach for providing Location Based Services (LBS) using Spatial Location (SLO) information within the Session Initiation Protocol (SIP). The aim of the paper is to set up the framework for providing LBS services in devices connected to wire and wireless IP networks. This method uses SIP as transport and the SLO as data format inserted in the SIP payload. It analyses the relationship among the network elements involved in the architecture, and its functionality for providing access to network services regardless of the user location. In conclusion, this proposal enables SIP to support location-related services such as messaging, location-based commerce, and any location-based computing. Furthermore, it describes the advantage of adding the user location information for network resources optimization in mobile environments and future networks.**

Index Terms: **SIP, location info, mobility, LBS, load balancing.**

I. INTRODUCTION

Location based services (LBS) are associated with the spatial or geographical locations of some entities while processing the service. The key for achieving LBS consists in how the spatial location of an interested entity is collected, notified to the interested entities, and available upon request from service provisioning elements.

The Session Initiation Protocol (SIP) [1] has been selected as signaling protocol for the new Universal Mobile Telephony Services (UMTS) networks by the standardization entity named 3GPP (3^{rd} Generation Partnership Project) [2].

SIP has the generic nature for managing media sessions as well as supporting network presence and messaging. SIP can transport any content data. Therefore, it is a protocol that can be used to carry spatial location information as the content data. In addition, its infrastructure can be utilized for obtaining, storing, and facilitating that spatial location information to location-aware applications.

There can be many mechanisms to locate or place a given SIP entity. SIP security mechanisms are then used to meet the possible security constrains involved for obtaining the location information of a given entity. In addition, the mechanism to obtain and register the location information associated with the user is already proposed in SIP for enabling emergency call [3].

Thus, the aim of this work is to make spatial location information available to enable location-based applications.

Manuscript received July 21, 2001.

The authors are active researchers at Nokia Corporation in Helsinki, Finland, e-mail: Jose.Costa-Requena@nokia.com, Haitao.Tang@nokia.com, Inmaculada.Espigares@nokia.com.

II. LOCATION BASED SERVICES

Location based services (LBS) services are ranging from only-dealing location content (retrieving information according to the user location), up to services involving network management and resources optimization in wireless communications.

In the former type of services, adding the location information during the terminal registration enhances the user to access specific network services, and augments the information content value. The geographical data can help to improve the information content because it is related to the real user location. Thus, the user gets the right information and the services according to his actual situation. They can be either crucial services ("e.g., the nearest Emergency service assistance point") or they can require the precise location to provide an accurate information (e.g., the LBS service of looking for the nearest Pharmacy or requesting a taxi providing the exact location, etc.). These services are related to the information content, and when the user requests a specific service the provider immediately checks the user location and supplies the closest site for attending the user request. In this case, the knowledge of location information is used to send the data requested based on the user location. It does not matter whether that information comes from the closest server or from an application server situated far away from the user. In these cases, the information content supplied to the customer is the one that has to be linked to his exact position.

Regarding to the latter services, the network architecture, its elements and capabilities are especially involved. The user leaves the own network and requests services to what it is known as *Visited network*. The provider that manages the *Visited Network*, based on roaming agreements with the Home Network operator, assigns a server to attend all the transactions coming from the temporal user. Nowadays, the Server that attends to the new user is assigned according to the specified terminal requirements and the services agreed between operators. The user at the *Visited network* starts a registration process. When the registration is completed, a new SIP server becomes its server to handle all the transactions desired while the user is under the coverage of that network. This server assignment is based just on the terminal capabilities, service agreements, and network load balancing.

This paper also describes how this server selection is enhanced based on user geographical location knowledge. The server assignment either in the Visited network or in the Home network is optimized based on the user location and profile. Mostly, the location information improves the assortment of the SIP server with new criteria. The directives for this decision are the combination of the geographical data, the user profile, the server load balancing calculation, and the terminal capabilities. Hence, new components are studied to build up the service pro-

1229-2370/01/$10.00 © 2001 KICS

visioning. Thus, these services are related to the improvement of information processing based on the user location.

III. SIP AND THE SPATIAL INFORMATION

The following sections describe the main characteristics of SIP and SLO applying to spatial location information management.

SIP is an application layer signaling protocol used for creating, modifying and terminating multimedia sessions among different parties. It is used as a call control protocol in IP Telephony replacing the traditional telephone services. As mentioned above, SIP has been chosen as the signaling protocol for the core network in the next generation of wireless communications. It is a text-based protocol for providing call signaling over IP networks among network entities. SIP includes the main indispensable headers for establishing a multimedia transaction. The header contents can be easily extended with new features for adapting new services. SIP, by itself, only provides call control. However, SIP accepts complementary information inserted as SIP payload for other applications. Currently, SIP is able to set up a voice call carrying the information of a more detailed multimedia session using protocols as the Session Description Protocol (SDP) [4].

An SIP session usually involves the following entities: the User Agent (UA), the Proxy Server, the Redirect Server, the Registrar Server, and the Location server. The UA is located in the user device. The Proxy, Redirect, Registrar and the Location server are network elements, interacting with the user and other network entities simultaneously.

The UA is normally placed in the user terminal. The UA initiates an SIP request and contacts the user when either an SIP request or response is received. The Proxy Server is an intermediate entity that behaves as a client and server simultaneously. It is a network element that receives the messages (requests) from the user (UA). It can analyze and modify the request before forwarding it to other servers. The Redirect Server is another network entity that receives the messages from the user and maps the address indicated by the request into zero or other addresses. Thus, the client should try again the addresses returned by the Redirect Server to contact the callee or another SIP server that can handle the message in case of special requirements. The Registrar Server accepts the user registration (REGISTER). The user information received within the REGISTER message is extracted and forwarded to be stored at the Location Server. The Location Server is an element used by a Redirect or Proxy server to obtain user information (including the geographical information of the callee, etc.). The Location Server is an SIP entity allocated within the Registrar server. Its function is collecting the information of the user that has been registered under the SIP Registrar. Therefore, the Location Server is improved with the storage of location information for implementing LBS services or any other location-aware applications.

An SIP message consists of a start line, one or more header fields, an empty line (carriage-return line-feed, CRLF) and an optional body. Basically, the start line indicates if the message is a request or a response. The main requests are INVITE, ACK, OPTIONS, BYE, CANCEL, and REGISTER; while common responses are 100 Informational, 200 Success, 300 Redirection; 400 Client Error, 500 Server Error, and 600 Global Failure.

SIP supports name mapping and redirection services enabling *personal mobility*. Personal mobility is the end user's ability to originate, receive calls, and access to subscribed services on any terminal on the network. It is also considered the network facility to identify end users as they roam among network sites. Personal Mobility is based on the use of an exclusive personal identity based on the SIP URL.

The SIP also supports terminal mobility by proxying and redirecting request to the current network location where the user is logged in. The user can register his current network location and also the service profile required for the session using SIP.

Therefore, it is remarkable when we can provide spatial location information as part of SIP signaling flows to facilitate the development of location-based services.

The Spatial Location (SLO) information is an XML [5] structure defined at the IETF for representing the user location information. The basis of SLO is providing a common and extensible container in which are placed the user identifiers, security factors, location representation, and other parameters to manage the user location according to his specifications. The objective of the SLO is to be secure and self-contained to carry the user information. Furthermore, SLO has to fulfil user needs and meet the privacy requirements proposed by the Cellular Telecommunications Industry Association (CTIA) [6].

CTIA indicates that mobile customers have to (1) be well informed of location collection and use practices *prior* to collection; (2) have the meaningful opportunity to consent to the collection and use of this information for location-based services; and (3) ensure the security and integrity of any collected location information.

Additional location data structures are under development in other organizations such as WAP [7] Location Forum, and Location Information Forum (LIF) [8]. Those data structures can also be easily placed as SIP payload and it has to be indicated in the headers that the message must be managed in a different way.

The advantage of using SIP for providing the location information is that it can be inserted into any type of payload in the body of the SIP messages. The content is indicated in the headers and upon receiving the message and checking the headers, the server recognizes what kind of data is included in the payload. Afterwards, it is up to the server to process the message accordingly.

An example of adopting SLO as SIP payload is already described to support emergency call services [3]. SIP is simply the transport means for connecting two entities over IP. SIP does not determine the purpose of the connection. The content of the SIP payload indicates what kind of session is going to be established. Normally, SIP carries an SDP packet describing an audio or video session, indicating end terminal's capabilities.

Furthermore, SIP has the capabilities for providing user mobility, and this is indispensable for providing services regardless of where the user is attached to the network. SIP is designed to perform a session establishment based on the transaction of the media characteristics indicated in the payload of the message. In the same manner, the Spatial Location (SLO) information [5], or

Table 1. Components of the SLO location data structure.

Attribute	Data Type	Priority
Datum	- WGS84	Mandatory
Coordinate	- Latitude	Mandatory
	- Longitude	Mandatory
	- Altitude above WGS84 reference ellipsoid	Optional
	- Altitude above mean sea level	Optional
Location	- Horizontal accuracy, by radius of a circle from the positioned point	Optional
Accuracy	- Altitude accuracy, by range from the positioned point	Optional
Time	- Real time of the measurement/fix	Mandatory
Speed	- Ground speed	Optional
	- Vertical speed	Optional
Direction	- Direction of movement	Optional
Course	- Direction from the current position to a defined destination	Optional
Orientation	-Horizontal orientation	Optional
	- Vertical orientation (pitch)	
Unspecified Attributes	- Attributes enabling some extensibility	Optional

the data structure defined in WAP Location Forum can also be considered as session data to be carried in the initial call set up, and placed in the SIP payload instead of the usual SDP packet. In this case, SIP is not used for establishing a complete session between two users. It just registers, stores, and exchanges the user location information. Thus, taking advantage of SIP flexibility for extending new capabilities, this paper analyses the spatial location use to enhance network resources management, and provide location-based content.

Nowadays, wireless and Internet access is rapidly converging. Mobile commerce is expected to grow at seemingly incredible rates with the increase of mobile users. Hence, adding SLO or any location data structure improves SIP to support location-based applications and even resources optimization. These new promising applications are location-based messaging, and other location based services. An SIP user enhanced with SLO can establish a session with any SIP-speaking provider of location-based services. The SIP message transfers the user SLO information to a provider. Afterwards, the provider can send back a response to the user based on his spatial location. Additionally, the knowledge of the user location in conjunction with his usage profile improves the load balancing optimization. Hence, when user request the network access and provides his location data, the provider can assign a nearby server based on the transaction load expected from that user according to his usage profile. With this approach it reduces the latency in the transactions when high bandwidth is required according to the user history. Obviously, the provider will not assign a closer server for another user who does not have a profile with high traffic requirements.

IV. LOCATION INFORMATION FORMATS

There are various formats adopted for representing the location information. The different efforts to define those data structures are handled at WAP Location Forum [7], LIF [8], and IETF [9]. The Spatial Location (SLO) is the data format defined at the IETF as shown in Table 1[1].

[1] WGS 84stands for World Geodesic System 1984. WGS 84 is an earth fixed global reference frame, including an earth model.

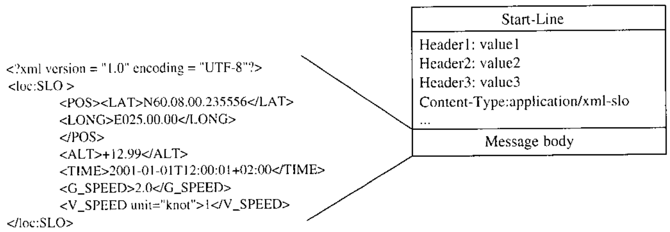


Fig. 1. SIP message containing location information.

Thus, in this paper is selected the utilization of SLO as an alternative to prove the feasibility of implementing global network architecture for location-based services using SIP. Nevertheless, the same architecture can be used with other type of location data structures. The differences reside in notifying the type of structure that the SIP payload is carrying and it is done using the message headers *(Content-Type:)*.

Fig. 1 shows how the SIP message body contains the XML file with the location information using the attributes defined in Table 1. [2]

V. SPATIAL LOCATION SYSTEM BASED ON SIP

This section describes the SIP-based architecture where SLO is included to enable LBS services. The principal objective is to establish an SIP based group of entities being able to establish/acquire and provide their spatial location data. SIP is a well-designed mechanism to achieve a common understanding among network entities to establish a media session. In this case, all of them have to agree in SLO data management that later on will be used for third party applications to provide LBS. In the above section is mentioned that a self-description procedure for security and data management is included as part of the SLO structure.

The user can be temporarily registered at the local SIP server. The registration simply indicates that the user is located at that network point for a short period of time. Thus adding the SLO or

[2] Namespaces are omitted to simplify the figure.

any similar location data structure as the SIP Content Type during the registration provides complementary information about his spatial location. Hence, the SLO is linked to the user identifier used in SIP (URL), which is stored in the Location Server during the registration.

SIP also has a server discovery mechanism to find out where an appropriate server that understands the SLO data is located. In this scheme the SLO data format is part of the protocol payload, and it is indicated in one of the protocol headers.

One important SIP feature is its capability of handling the terminal requirements. SIP provides server discovery mechanism using pre-defined configuration, multicast queries to a DHCP server or DNS query. Thus, the user terminal can obtain immediately the address of its closest SIP server when starts up. That server is the contact point to which the terminal will address all the outgoing transactions such as the initial registration (REGISTER). The server contacted has to be able to handle the user-originated transactions. In case that the local server cannot attend the user requirements regarding services or requested applications, it will either forward the registration to another server or will return to the user the address of the capable server. The user and terminal requirements are indicated using different methods. It is possible to use the *Require* header to indicate what concrete server features are necessary to manage this message. It can be indicated also in *Contact* header, where the user sets his priorities and media capabilities.

In what follows, there is an example of *Contact* header in a REGISTER message. It shows that the user is logged at *nokia.com*, she is able to speak English (en), Spanish (sp), and French (fr) and the terminal has audio, video, and SLO data capabilities (media). The duplex parameter "full" indicates that the UA can simultaneously send and receive media. However, she only wants callers to use the terminal if the call is of priority "urgent" or higher.

```
Contact: Inmaculada <Quex@nokia.com>;
language:"en,sp,fr"; media:"audio,video,application/SLO";
duplex:"full";
priority="urgent"
```

Other parameters included in the Contact header could have been *class, feature, mobility, methods, description*, and others. The *class* parameter indicates whether the UA is found in a residential or business setting. The *feature* parameter enumerates additional features of the UA. The *mobility* parameter indicates if the UA is fixed or mobile. The *methods* parameter indicates the SIP-methods that the UA understands. The *description* parameter describes, as text, the terminal. Further information is detailed in the "SIP caller preferences and capabilities" specification [10].

This feature with similar examples enabling application layer mobility is also presented in other documents [11].

In this paper, since the SLO or any location information structure is a new feature for SIP, there remains the possibility that a particular SIP registrar contacted do not support this content. Thus, once the user has the initial server to address, it can send the SIP REGISTER message containing the SLO payload in its body and the service requirements in the Contact field. The message also has specific header for discovering the Location Server. An SIP header called "Require" is included in the REGISTER message. The Require header field is used by clients to tell user agent servers about options that the client expects the server to support in order to properly process the request. Although an optional header, the Require must not be ignored if it is present. Thus, the SIP REGISTER has the "*Require: SLO-server*" header for indicating that the incoming message contains a Spatial Information and needs to be processed by a Location Server. In case that the server contacted has no SLO Capabilities the user receives back a response where the Contact header includes the address of the new SIP server to handle the message ("*Contact: slo-server.nokia.com.*").

Fig. 2 shows the process described in this section. Firstly, a basic SIP server discovery using DHCP is illustrated. Following, the SIP/SLO enabled server discovery allows the final registration. As the example shows, the message contains the *Require* header, and the *Contact* header returning the SIP server address is able to handle the SLO data.

In the next section, the advantage of using the location information for selecting the right server to attend the user based on terminal requirements and usage profile is described. Hence, initially the assignment of the SIP server that has to attend the user requests is decided based on the terminal requirements. Now, this decision can be improved with other criteria. Afterwards, the user is registered and his location information is available in the Location Server. The user registration including his geographical location enables the implementation of Location Based Services as value-added in the information content.

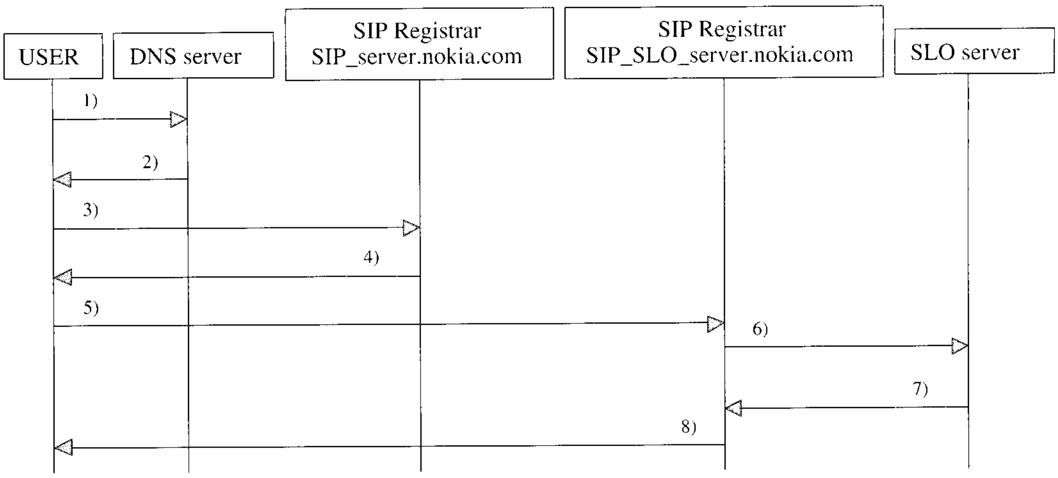
VI. SCENARIOS WITH SIP AND LOCATION INFORMATION

In this section the issues of providing the user location information to the network and the management of SLO data are considered. In addition, various advantages derived from enhancing SIP with location data information is depicted. To implement LBS services, SIP messages need to be assisted with an additional data like SLO or any structure with similar characteristics. In this case, the SLO and the architecture requirements to develop a location-based architecture are illustrated. Due to the adoption of SIP for the 3G mobile networks and its flexibility, this architecture is easily implemented using SIP entities.

A. Push and Pull Services

This subsection describes the procedure to implement Location Based Services using SIP and SLO considering basic registration based only on terminal capabilities, based on the server discovery mechanism. The terminal registers and provides its media and SLO requirements without any further resource analysis forced to the serving SIP server. In case that the SIP registrar has a SLO server accomplishing the user device requirements, it responds to the target after treating the content of the SLO structure. Once the user is trusted and the SIP payload is accepted, the message is forwarded to the SLO-enabled Location Server.

The user registers and provides the SLO data. That data is stored in the Location Server databases and can be used later



1) DNS SRV query; DHCP

2) SIP_server.nokia.com

3) REGISTER sip:SIP_server@nokia.com
To: joseph@nokia.com
From: joseph@nokia.com
Contact: Inmaculada <Quex@nokia.com>;
language:"en,sp";media:"audio,application/xml-slo"
Cseq: 1 REGISTER
Require: SLO-server
Content-Type: application/xml-slo
*****ENCRYPTED DATA*****

4) SIP/2.0 300 Multiple Choices
To: joseph@nokia.com
From: joseph@nokia.com
Cseq: 1 REGISTER
Contact: SIP_SLO_server@nokia.com

5) REGISTER sip:SIP_SLO_server@nokia.com
To: joseph@nokia.com
From: joseph@nokia.com
Cseq: 1 REGISTER
Require: SLO-server
Content-Type: application/xml-slo
*****ENCRYPTED DATA******

6) Registration request to theSLO server

7) Registration response from theSLO server

8) SIP/2.0 200OK
To: joseph@nokia.com
From: joseph@nokia.com
Cseq: 1 REGISTER

Fig. 2. SIP/SLO Server Discovery Process.

for requesting any LBS service. Once the user has registered the location information, he can perform either a pull or push service for requesting a location-based service. During the registration the user can set a certain Policy mechanism to third parties trying to access his information (with additional SLO attributes). The user can make his information available to the public or totally private. In the latter case the user has to give the pertinent authorization to the third party requesting the information. Furthermore, it is necessary to establish a secure channel for the transmission of that information between the Location server and the LBS provider. However, if the target designates the Information to be totally public, any un-trusted entity can request his location information.

Based on these assumptions, an intermediate element is required between the Location Servers where the user profile and the LBS information are stored.

Information User (IU) is a new entity needed to implement LBS in combination with the proposed SIP/SLO architecture. IU is habitually a value-added service/application that can access to the user location information. After user authorization, the IU accesses to the Location Server for obtaining the user spatial location information. This information can be required by any external elements such as a routing module, a signaling protocol, a lifesaving service, etc. The IU also performs a server role when it acts as an interface for accessing directory services or local facilities (taxi, restaurants, etc.). In this case, the IU knows the user's spatial location through the Location Server, and takes care of requesting the services needed by the user through third parties.

The approach facilitates the implementation of the pull and push services model depending on the user specifications.

The user can subscribe to an application server (IU) for obtaining some information about local sites for entertainment, leisure, or health services. In this case the user gives the rights to the service provider to access his location information stored in the Location Server. Hence, this application sends some information periodically to the user according to his geographical location at each moment.

Furthermore, the user can request similar services only when he wishes so. In this case, the user does not need to register to any LBS server, he just sends the query to the IU requesting some LBS service and sends his location information keeping his anonymity. The SIP server listening to the user gets his petition and makes the query to the IU that contacts the third party without revealing the identity of the user that made the request. Finally, the LBS server returns the information to be forwarded to the user. In this model, the user is hidden from third parties application servers and can access to identical services.

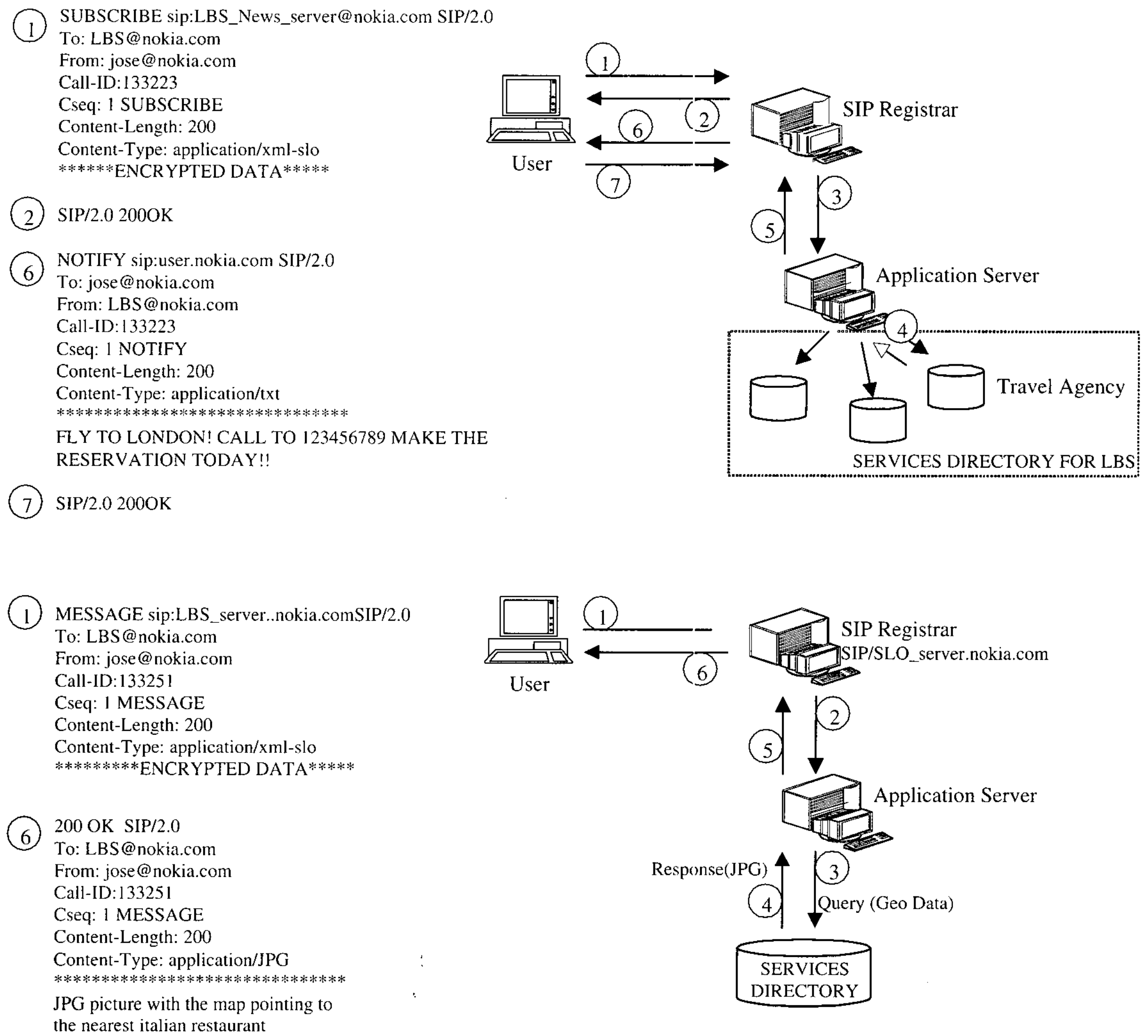


Fig. 3. SIP/SLO messages in the Pull and push LBS services scenarios.

In resume both types of services can be easily implemented using SLO as part of SIP. It is possible to use the existing methods available on SIP like SUBSCRIBE, NOTIFY, and MESSAGE to perform both kind of services.

SIP flexibility allows the extension of new features to facilitate business opportunities to third parties. Operators can implement their own LBS applications or other external service providers can offer those services. User then selects his favorite. The technique opens a free competition for services provisioning.

Fig. 3 represents the push and pull services model using the SIP methods. In the first case, the user registers to get information regarding hot offers. He is pushed to buy flight tickets. In the second case, the user requests personalized information regarding the nearest Italian restaurant, and he receives a map (JPG picture) with the indication of the most popular Italian restaurants according to his location.

B. Location Aware Signalling

This subsection refers to the advantage of knowing the user location when roaming for optimizing the assignment of the SIP server. During the network access the user will gain some benefits in terms of transactions latency if he provides his location. Furthermore, the provider can manage efficiently in terms of resource reservation (bandwidth requirements) when assigning the SIP server to the user based on his location.

This server assignment procedure is adopted in the next generation of wireless networks (UMTS or 3G networks). This section introduces briefly a description of the 3G infrastructure. An example of how the location information is useful and also the main benefits are presented.

B.1 3G Networks Architecture

The UMTS network is IP based and uses SIP as the signaling protocol. For that reason the entities involved are similar to the IP networks with different naming. SIP is used as network layer to communicate among all UMTS entities. Basically, the 3G elements have the same functionality than the SIP network elements described in the SIP overview.

The Call Processing Server (CPS) is the network element that controls the Call State, performs the right decisions, and handles the routing process to set up the call. The Call Processing Server (CPS) implements the Call State Control Function (CSCF) as part of the core network access control. The CSCF is the logi-

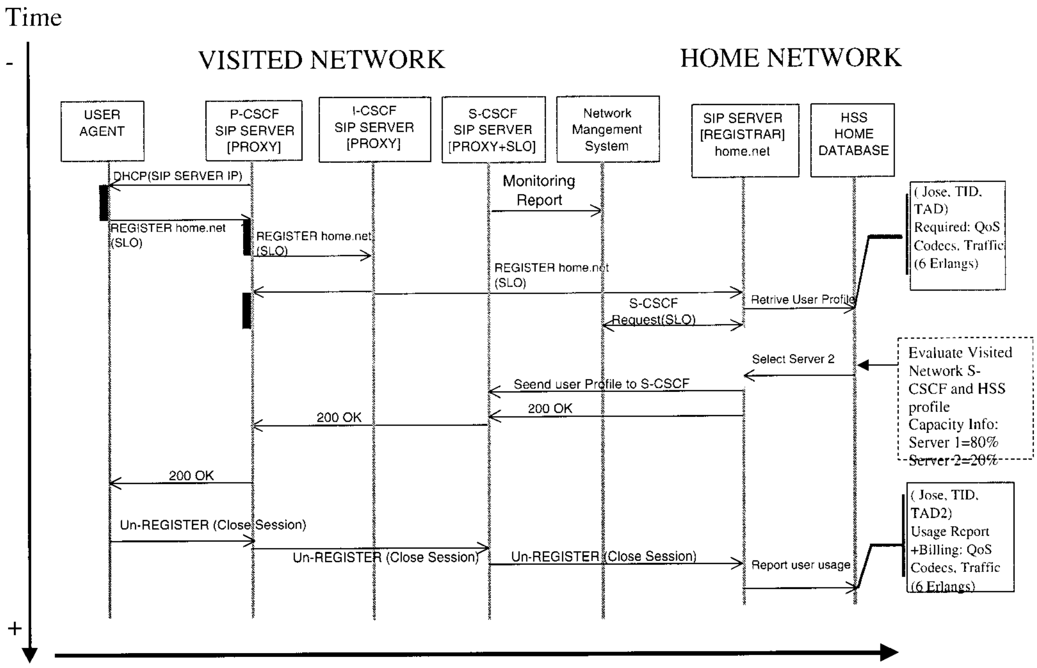


Fig. 4. 3G networks user registration and S-CSCF assignment.

cal entity that contains the SIP functionality for terminating the user-to-network signaling applied for the network access. The CSCF, as part of the CPS contains the SIP functionality of proxy server, redirect server, and registrar.

Depending on the precise role of the CSCF, it may include additional terminology. Thus, I-CSCF stands for Interrogating CSCF, which is the CSCF to which others CSCFs query to obtain the callee address for routing the call to reach the end terminal. The S-CSCF stands for the Serving CSCF. After the terminal registration, the S-CSCF attends all the messages coming from the subscriber. Finally, P-CSCF stands for the Proxy CSCF. This is just an intermediate CSCF that receives the incoming messages from the terminal, and forwards them to the S-CSCF attending the user.

The Home Subscriber Server (HSS) is like a Database that stores user profile. The HSS keeps the user information regarding user/service profile and terminal capabilities. Thus, when the user intends to register in the network the HSS returns the user credentials and the profile to indicate what are the user requirements based on his subscription. That information is used for authenticating and authorizing the user to access to the network. Moreover, the user profile stored in the HSS is also utilized for assigning the S-CSCF that has to attend the user while he is connected in the network.

After assigning the S-CSCF and while the connection is active, the user profile at the HSS is downloaded into the S-CSCF. When the user terminates or un-registers from that network, the profile is sent back to the HSS for billing and service control purposes.

When the user registers in the Home Network he is assigned to an S-CSCF located either in the Home Network or in the Visited Network. From the S-CSCF the network establishes a set of connections to all the applications that the user requires. From the traffic generated at the User Equipment point of view, the P-CSCF and the S-CSCF will represent the bottleneck in the 3G architecture. The rest of this section argues the benefits of using the location information for the right selection of the S-CSCF. Fig. 4 illustrates the message flows.

B.2 Early Congestion Prediction

The previous subsection described the procedure for registering the user in 3G networks. It was also foreseen that the assignment of the right S-CSCF would be the main issue to avoid future congestion problems during the ongoing sessions.

This section proposes an "Early Congestion Prediction" mechanism implemented using a traffic knowledge of the network and the location information. This mechanism evaluates the present congestion in the network and based on that predicts the congestion that the new user would create in a certain location area. This measure helps to assign the appropriate network element for serving the new user on that area. The actual congestion avoidance mechanisms detect congestion problems, and inform the traffic source to ask for slowing down the transmission rate. An IETF group already studies these proposals defined as "Explicit Congestion Notification" [10].

This mechanism is very relevant in low-bandwidth delay-sensitive TCP connections. In wireless networks where the network paths that traverse, such as the Air interface, are a very low rate links, the mechanism "Early Congestion Prediction" is very useful. These low rate links can saturate the available bandwidth even in the previous links where the rate is definitely much higher. The buffers in the routers along the path previous to the low rate link take the responsibility of absorbing the bursts.

The inconvenience is that due to the congestion in this link, the routers along the path will overflow their buffers and they will not absorb the burst. This congestion would be spread along the path, avoiding other connections to perform properly.

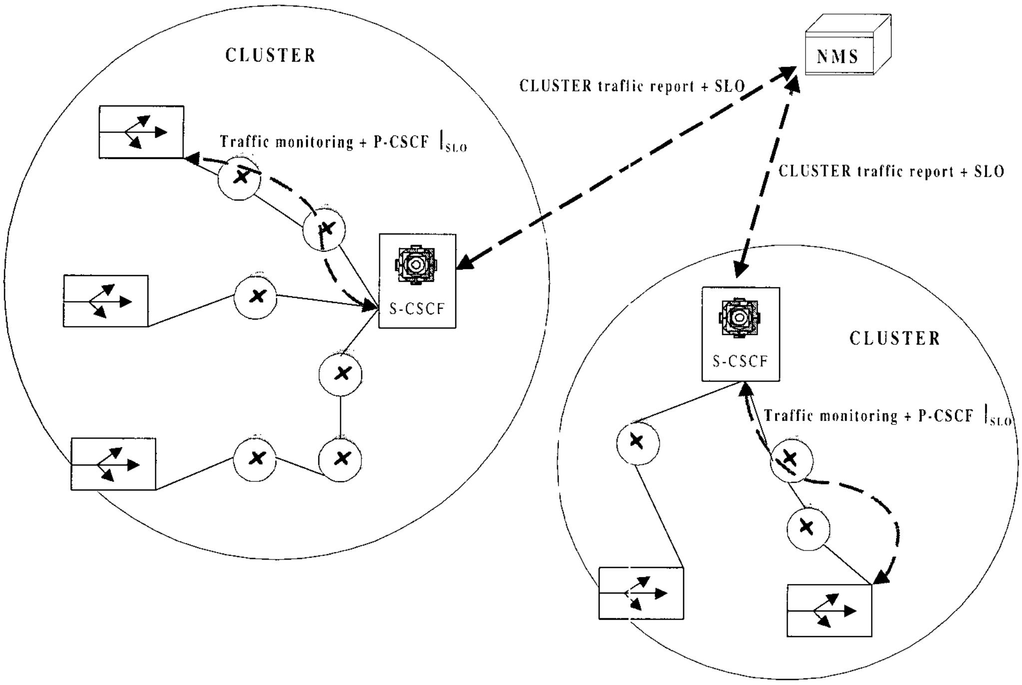


Fig. 5. Clustering structure for traffic monitoring.

B.3 Location Information Motivation

The proposal described in this paper refers to the initial phase when the terminal is registering in the network. Based on SIP, the registration can be either accepted when the server is capable of managing the user requirements or addressed to another SIP server if the local server cannot handle those requirements. This decision is favorable, and the registration is just accepted when the SIP registrar meets the terminal requirements regarding media or other requested capabilities. This assignment is done based on a policy decision at the Home network. The Home network receives the registration message from the user equipment and after authentication and authorization procedure, based on the HSS information, assigns the S-CSCF.

Thus, including the user location during the registration, the SIP server can check the user profile from previous connections. The user profile records the traffic generated, session duration per connection, and the minimum bandwidth necessary for the QoS required according to the real time transactions requested in previous connections. Hence, that information, and the user's location help the provider to assign a local server or a remote one improving the criteria. The network management and resources optimization are also enhanced.

Usually, the load distribution is done randomly assuming the same capacity for all the available servers, and an exponential negative session duration distribution. In this approach the different types of users (regulars and eventual) are not considered. It is important where regular users with long connections and lower bandwidth profile will masquerade the connections for eventual users with higher bandwidth and low latency required.

There is some research on the area to implement a scalable mechanism for traffic monitoring in IP networks. This paper proposes the usage of sampling mechanism implemented in a "Flow Agent" [12], which collects the flow information and forward the data to the central collector. This proposal defines the same Agent-based procedure but it defines a new cluster-based infrastructure. The objective is to use existing mechanisms but avoiding the explosion of monitoring messages from the multiple agents to the collector. This initiative defines a hierarchical structure where the monitoring systems are reporting traffic and flow characteristics to the local collector. Using the 3G nomenclature the collector would be the S-CSCF receiving all the monitoring information.

In this scenario, every S-CSCF would collect all the traffic samples and would evaluate all the flow for predicting in real time the most probable links with congestion. That information would be processed including the location information of all the P-CSCFs under the same cluster connected to the S-CSCF. Fig. 5. With this approach the S-CSCF has a valuable map where all the paths from the multiple P-CSCFs to the S-CSCF are included. The path information contains the actual traffic and an approximate prediction.

This traffic-path information that includes the location information would be processed at the S-CSCF and forwarded to the network management system that would have a clear knowledge of the network status. That information could be used with the user profile stored in the HSS to assign the best S-CSCF.

The clear idea is to assign an SIP server closer to the geographical location of the user when he has a profile with antecedents of high network utilization. Thus, the bandwidth resources required by the user are set closer to his actual location mapping physical-distance to network-distance avoiding transport latency in real time transactions for eventual users. It can be known if the user needs high capacity looking at his precedent transactions.

Based on higher frequency call setup attempts it creates the computational efficiency needed for processing the required sig-

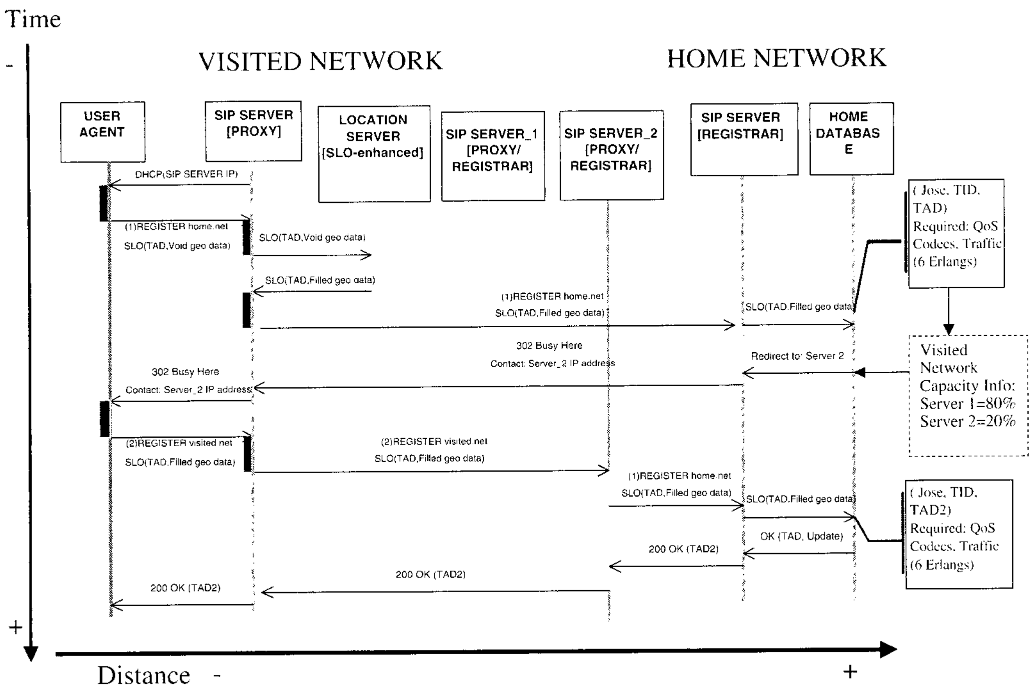


Fig. 6. User registration using SLO with SIP for resource optimization.

naling. Therefore, when the regular user has lower QoS (in terms of bandwidth/latency) requirements and the call attempts derived from his usage profile is low; the network can assign an SIP server far away from his actual situation, in terms of network distances.

The user can be assigned a server even at the Home Network, because of the lower (but long duration) number of transactions generated. In this case the load produced for managing the signaling is very low. Furthermore, if the provider assigns the nearest server it will exhaust the resources and will degrade the time response in the local server for the eventual users. Fig. 6 shows the user registration in a visited network using SLO in the SIP payload, where it is assumed a roaming agreement having the knowledge about SIP servers, its topology, and its capacity in the visited network.

Therefore, with the network resources information and the user location, the Home Network can optimize the server assignment guaranteeing the QoS and bandwidth required by the user.

In any case, the attending SIP server can pertain to the same provider or it can be part of the Visited Network when the user is roaming. In the latter situation some roaming agreements and user information exchange are needed. SLO provides some data structures to facilitate the user location and identification during roaming situations. The user is identified at the home server by a global identifier, named Target Identifier (TID). That identifier is stored at the Home Network together with the user profile. The TID was assigned to the user already when he registered at the home SLO server. Furthermore, there is another element that is used to access the user information. It is called TAD (Target record Accessing [13]), and is the Home Network that initially assigns it. Hence, the Home server stores both the user TID and TAD. There can be more than one TAD for a given TID. The TAD can be time or place dependent and it is only temporally attached to the Visited server.

Now, when the Target moves to a different location he is going to update his data. Thus, the default TAD assigned to the target by the home server serves as a default record accessing ID. The Visited network contacts the Home network using the default TAD. After the authentication, the visited server allocates a temporary TAD for the target, and informs its default location server of the target's current temporary TAD. Its default location server can then bind the two TADs (the default and the temporary). The TAD points unequivocally to the actual location of the target and is used for finding out the actual point of attachment of the terminal.

C. *Privacy Control*

Some location-based services (like other services over network) may cause privacy risks and concerns to the users/service providers. Privacy policy or indicator could be carried in the SIP payload as well whenever needed, to inform or indicate the other SIP end entity on the privacy requirements and risks associated with the LBS service. In addition, a Privacy Control Information Base (PCIB) should be set in a LBS-related SIP end entity (SIP user or network server).

However, privacy control can only be achieved by a relevant application program interface (API), a LBS-extended SIP API in this case. This API should take care of the responsibility to alert the user, get the user consent, or block certain LBS service requests according to the user's privacy setting in the PCIB, and possibly the privacy information from the other end entity involved in the required LBS service. While the API design is not under the scope of this work, we strongly recommend that a privacy control mechanism should be built into the API.

VII. CONCLUSIONS

Along this document it has been analyzed the inclusion of the SLO data as part of the SIP payload during the registration. The user registered using the SLO data makes his location available for spatial location based services. The user location information is stored in the SIP Registrar database, and it is globally available to facilitate the provisioning of location based information.

Those databases can be accessed using SIP signaling for requesting the user information previously registered. SIP already provides authorization and authentication when accessing to the SLO data to prevent certain security attacks. In short, the SIP is enhanced by adding spatial location capabilities, via extending the protocol with some specific headers for requesting the user location information.

In addition, this enhancement has made it possible to develop an efficient way to manage and assign network resources based on the user profile and his location information, such as an efficient way to implement load balancing when assigning the serving SIP server to the user in a roaming environment.

REFERENCES

[1] M. Handley *et al.*, "SIP: Session initiation protocol," IETF, RFC 2543-bis, Internet Draft, draft-ietf-sip-rfc2543bis-04.txt, 2001.

[2] Third Generation Partnership Project. Available at http://www.3gpp.org.

[3] J. Costa-Requena and H. Tang, "Application of spatial location information to SIP," Cluster Computing Journal. Available at http://www.baltzer.nl/cluster/cluster.asp, 2001.

[4] M. Handley, V. Jacobson, and C. Perkins, "SDP: Session description protocol," IETF, Internet Draft, draft-ietf-mmusic-sdp-new-00.txt, 2000.

[5] M. Korkea-aho *et al.*, "A common patial location dataset," IETF, Internet Draft, draft-korkea-aho-spatial-dataset-00.txt, 2000.

[6] M. Altschuk and R. Coleman. "Petition of the cellular telecommunications industry association for a rulemaking to establish fair location information practices," Rcd 6170.FCC.

[7] WAP Forum. Available at http://www.wapforum.org.

[8] Location Inter-operability Forum. Available at http://www.locationforum.org.

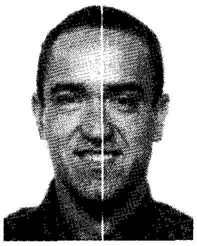
[9] The Internet Engineering Task Force. Available at http://www.ietf.org.

[10] H. Schulzrinne and J. Rosenberg, "SIP caller preferences and callee capabilities," Internet Draft, draft-ietf-sip-callerprefs-04.txt, work in progress.

[11] H. Schulzrinne and E. Wedlund, "Application-layer mobility using SIP," *Mobile Computing and Commun. Review* (MC2R), vol. 4, no. 3, July 2000.

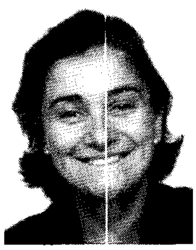
[12] S. Dawkins *et al.*, "End-to-end performance implications of slow links," RFC 3150.

[13] P. Phaal, S. Panchen, and N. McKee, "InMon Corporation's sFlow: A method for monitoring traffic in switched and routed networks," RFC 3176.



Jose Costa-Requena graduated with M.Sc. in Telecommunications Engineering in Polytechnic University of Valencia (Spain) in April 1999 having done his Master's project at Helsinki University of Technology (HUT), Finland. He joined HUT to work on SCN/IP telephony Interoperability architecture in 1998. Since 1999 he has been a Ph.D. student at the Networking Laboratory leading a group of M.Sc. students on the routing testbed project. In 1999 he joined Nokia and started working on SIP matters. Since 2000 he has been actively involved in standardization and security issues.

Haitao Tang is a senior R&D staff member in Nokia Research Center (Helsinki). He received his M.S. degree in communication and electronic systems in 1986, and his Doctor of Technology degree in computer science and engineering in 1998. He has continuously worked on data networking and applications since 1992.



Inmaculada Espigares Del Pozo graduated with M.Sc. in Telecommunications Engineering in Polytechnic University of Valencia (Spain) in December 1999, having done her Master's project at Helsinki University of Technology (HUT), Finland. In 1998 Inmaculada joined HUT to work on SIP telephony and services development. Since 1998 she has been working on SIP mainly in services related matters. In May 2001 she joined Nokia Mobile Phones in Helsinki and she is involved in the design of next generation services.