IMPROVING 'PRESENCE' SITUATION IN THE SIP BASED IP TELEPHONY NETWORKS

R. A. Akerkar

Technomathematics Research Foundation, Kolhapur, India
rakerkar@tmrfindia.org

ABSTRACT

The convergence of Voice and Data networks has changed the rudiments of telecommunication. With the spread of Cellular networks, convergence of voice, data, and cellular network is envisaged allowing users to seamlessly access services across networks while on the move. In this paper we discuss problems related to "Presence" in a converged network. A key issue in presence is to trace a user correctly and consistently. In a SIP (Session Initiation Protocol) based IP telephony network the presence information is kept as "Registration" bindings. We propose to use an alert mechanism, similar to paging in mobile network to notify the user of expired SIP registrations, so that the device/user can update the registration bindings. This improves the overall functionality of a SIP based IP telephony network and allows users to receive calls that would have otherwise been missed. In case the call does not mature, we inform the called party about the missed call so that the conversation can mature asynchronously. This solution would increase the reach of VoIP (Voice over Internet Protocol) to people and regions where internet penetration is low.

KEYWORDS

Network Protocols, Wireless Network, IP Networks, SIP, SICAR, SIPAR

1. INTRODUCTION

Last decade has seen the emergence of data networks as the prime information carrier for the masses. The explosive growth of internet has revolutionized the field of telecom as well. With the convergence of data and voice networks becoming a reality, people are increasingly realizing the benefits of a converged network. While circuit switched networks provide better voice quality, packet switch networks are more efficient. IP telephony has emerged as a popular application of converged networks.

Besides providing cost benefits to both users and service providers, it also allows the service providers to port the services available in IP networks to telecom networks.

SIP [1] is increasingly being used as a multimedia session establishment and control protocol between end users of an IP telephony network. The advantages of SIP are its simplicity and the flexibility to support other standard protocols like Real-time Transfer Protocol (RTP) [2], Stream Control Transmission Protocol (SCTP) [3]. Along with the Internet the cellular networks have also witnessed high growth. In many countries the number of cellular phones has crossed that of fixed wire line phones. With more intelligence being embedded in the mobile phones, they are set to become miniature computing device for the masses in future. The convergence of the cellular and IP networks has been propelled by the development of technologies like Multimedia Messaging Service (MMS)[4], Wireless Access Protocol (WAP)[15,17], General Packet Radio Service(GPRS) [5], SIP CGI [19], CPL (Call Processing Language) [20], VoiceXML [21], and Unified Mobile Telecommunication Systems (UMTS) [6] which allows the users of a cellular network to access the Internet using their cell phones and other portable devices. Presence services constitute information about the willingness, availability and capability of a user to communicate with other users. Presence information plays an important role in successful call establishment as it allows the network to correctly identify the capabilities of called party and thereby allowing the caller to make the call only when the network is sure of the called party being willing and available to receive it.

In SIP based IP telephony network, users update their presence information by using SIP Register messages. The Register message allows different contact addresses to be associated with a user's SIP URI. These bindings are used for determining the exact location of the user and allow the SIP proxy servers to forward the calls to the address mentioned in these bindings. These bindings are valid for a certain period of time and need to be periodically refreshed. On their expiry the called party will not be able to receive calls. Currently SIP provides no mechanism to discover the location of users with expired registrations. Calls made to such users cannot be completed by the SIP proxy server and the caller gets back an error message. These calls are defined as missed calls. In case of a IP to PSTN call, the called party may be roaming in a foreign country or busy in a meeting and unwilling to accept the call; in such case also the call will not mature resulting a missed call.

This paper intends to improve the presence situation in the SIP based IP telephony networks by proposing a mechanism similar to paging to discover the user's current location and alert him of his expired SIP registrations. In the solution being proposed by us we use paging to locate a user's mobile device and then use "Push" mechanism to notify him about his expired registrations. Two models of solutions, viz., Server Initiated Client Activated Registration (SICAR) and Server Initiated Principal Activated Registration (SIPAR) are proposed. These solutions are very much flexible in nature and can be initiated either for all the calls missed by the user or only for some prioritized calls, which are defined by the user.

The proposed model of solution aims to bring together the cellular and VoIP(Voice over IP) technologies. The presence information in cellular networks is more consistent and accurate than in VoIP networks and thus by integrating both we wish to improve the presence in VoIP networks. Currently VoIP is limited to regions with internet connectivity but cellular networks are more widely spread. In developing countries like India where a large percentage of population resides in areas with negligible internet access but with good cellular network coverage, extending the VoIP networks into the cellular network would make the VoIP calls accessible to a greater number of people. So a solution consisting of using cell phones to receive updates about the VoIP calls and setting up of internet kiosks across the rural areas [7] would allow the rural population to converse using inexpensive voice over IP technology. To best of our knowledge, this is the first time push mechanism of the Cellular System is used to proactively resolve the presence challenges in the IP Telephony system.

The rest of the paper is organized as follows: in section 2 we discuss about presence and challenges in it. Section 3 provides a background on SIP and the details of the SIP Registration process. In section 4 we look in to the details of Paging and various Push technologies. Section 5 explains the proposed solution for expired SIP registrations. In section 6 our conclusions are stated. In the final section references are provided.

2. TECHNOLOGY

Presence technologies initially developed for Instant Messaging services, are now being widely used for other services like remote control of Network Appliances [8], remote monitoring of patient health, inventory control etc. The presence technology consists of three elements in it – *presence entity, presence server and subscriber*. The presence entity is the user/device whose information is sought and subscribers are the users who wish to have this information. The presence server stores, updates, and provides the presence information. The presence server can be logical entity located along with a registrar or a proxy server in a SIP based system.

The presence technology faces quite a few challenges like the subscriber willing to get information about a presence entity must first discover the presence server associated with that entity. The presence server must also be capable of receiving all the updates related to a presence entity. This updated information must be synchronized that is it must be updated at all the associated databases and must be sent to all the subscribers who have subscribed for the information.

Using an underlying protocol like SIP for presence information can solve many of the challenges like the SIP server discovery [9] can be effectively used for discovering presence entities and presence servers. SIP has an extension for event notification [10] consisting of “Subscribe” and “Notify” messages, which provides asynchronous event notification to the users. This can be effectively used for distributing presence information.

Currently various efforts are being made at different levels to integrate SIP and various mobile technologies. Like Open Mobile Alliance (OMA) [18], an industry wide association of different telecom companies, has also developed SIP Push OTA specification [11] that proposes to use SIP as a transport bearer for WAP Push [12]. These efforts are directed towards improving the Instant messaging services.

3. SESSION INITIATION PROTOCOL

The Session Initiation Protocol (SIP), an application layer protocol, is used for establishing maintaining and terminating multimedia sessions between end users in an IP based real time application. Users in a SIP session establish connections using User Agents, which forward the requests to the Proxy or Redirect Server. The proxy/redirect server uses the location services provided by a Registrar to determine the location of the users and forwards request to them. The Location service provided by the Registrar is a two step process:

- **User Registration**

The user registers himself at the Registrar of a particular SIP domain. He sends a REGISTER message to the Registrar containing his SIP URI and a set of contact addresses. These contact addresses are bound to his URI [13].

The Registrar might want to authorize the user before processing his register request. The Registrar has an expiration time set with each of the contact address. This expiration time period can be set by the user by including an Expiration header field in the Register message or in the absence of the Expiration header the Registrar may set a default value. On successful registration the Registrar returns an acknowledgement containing all the registered contact address back to the user. Figure 1 shows the registration process in detail.

- **Location Service**

The Location service is used by the Proxy servers to discover the current address binding of a user so that calls meant for that user can be established. The Proxy server queries the called party ’s Registrar for a mapping of the called party ’s SIP URI with the contact address. The Registrar may reply back with more than one address binding. If the called party has moved to another domain then the new URI of the called party is returned else if no matching records are found for the called party then an error response is send back. The SIP registrar server doesn’t further attempt to determine the current location of the user. A binding for a particular SIP URI is removed after its expiration time period has elapsed.

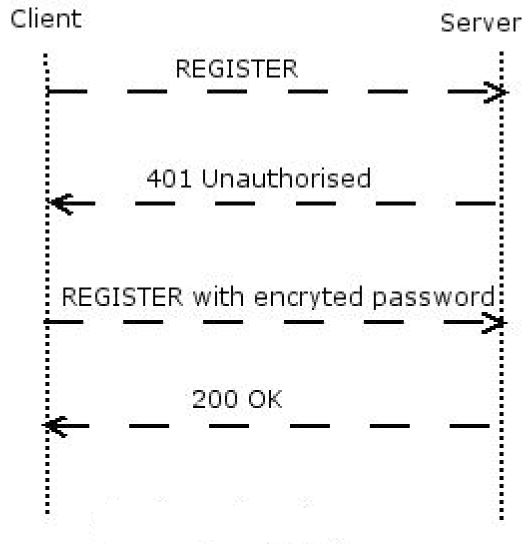

Figure 1. Registration in SIP

4. PAGING AND IMPROVING PRESENCE

In cellular networks paging is used to determine the exact location of a mobile station. A mobile station can be either in *active* state or in a *standby* state. In the active state the mobile station's location is constantly monitored by the base station. In standby state the mobile tracking is done in a much coarser way, with the mobile station being located in a *paging area*. A paging area is a group of cells in which the mobile is expected to be located. When the mobile station needs to be located then a request is sent to all the base stations in the paging area.

In the solution model proposed by us for improving presence in VoIP networks we plan to use a technique similar to mobile paging. When the user's SIP registrations are expired, then the registrar server is unaware of the user's current location. But if the registrar server is aware of the user's mobile number(s), then the registrar can contact each of these mobile stations and try to renew the registration bindings. Thus the SIP registrar uses a model similar to *paging*, in which the registrar has the knowledge of the expected bindings of the mobile user and then proactively contacts each of these to determine the binding for the user. This proactive nature of the SIP registrar resembles the *push* architecture.

Wireless Access Protocol or WAP developed as a joint effort of various different cellular companies, is a layered protocol stack consisting of different layers for application, session, transaction, security and transport services. The WAP model consists of a user making a request using a WAP browser from his mobile phone. This request is first routed through a WAP gateway, where the request is processed and the gateway then retrieves the content from a content provider. The content then is formatted and sent back to the user. The WAP Push model consists of a Push Initiator (PI) which initializes the push mechanism and contacts a Push Proxy Gateway (PPG) by the Push Access Protocol (PAP) [14]. The PPG then uses the Push Over-The Air protocol for forwarding the data to the Mobile station. Figure 2 shows all the entities of WAP push architecture.

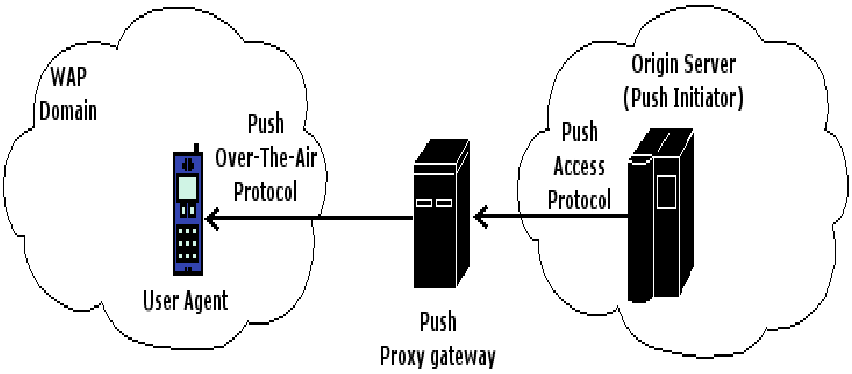

Figure 2. WAP PUSH

5. SOLVING SIP REGISTRATION EXPIRATION USING PUSH

5.1. Missed calls in SIP

The concept of missed calls is not well formalized in SIP based IP telephony networks. However, we have taken the concept from the cellular network. In a SIP based IP telephony network the miss call can be defined in the following cases –

- While receiving a call, the called party doesn't answer the call or does not lift up his user agent. The user agent could be an IP based soft phone or a PSTN based phone.
- The user's all SIP registrations have expired and therefore the Registrar cannot locate the user's exact location. So though the user may be willing to receive calls but due to registration expiration the system cannot locate him.

Solving the first kind of missed calls requires the system to store a list of calls not received by the user. This information can be sent to the user either in an email or as an SMS message in an asynchronous fashion.

5.2. Basic Philosophy

In a normal IP telephony call the SIP Proxy server on receiving a call for a registered user of its domain, enquires the called party's current address bindings from the Registrar. It then forwards the call on the received address bindings. If no bindings are available and if the server recognizes the requested URI as a valid user of the domain then it generates a 480 Temporarily Unavailable response message (see figure 3) and sends it back to the proxy. After this the Proxy checks the user profile of the user (called party) to determine if the called party needs to be notified of the missed call. The called party can set priority levels to filter out calls depending upon his willingness to accept some all or none calls.

The system provides the called party with the option to set different priorities to different incoming calls. For incoming calls with higher priority than the current value set by the called party the Push mechanism is initiated. The proxy server in this case acts as the content provider and it then calls the *Push Proxy Server (PPS)*.

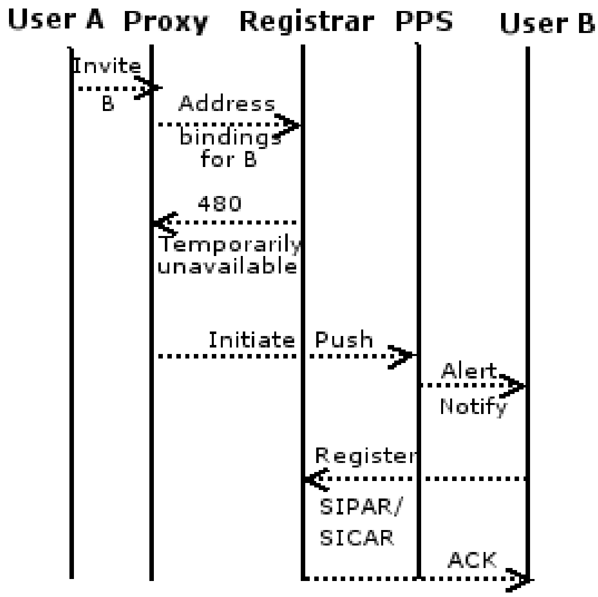

Figure 3. Basic SICAR/SIPAR message flow

The PPS acts as the PI and initiates the push mechanism. Depending upon the registration mechanism used two different models of registrations are proposed –

- *Server Initiated Client Activated Registration (SICAR)*
- *Server Initiated Principal Activated Registration (SIPAR)*

The underlying assumption behind these registrations is that the user has a mobile phone and the SIP servers have the knowledge about the mobile phone. This information can be collected and stored in a user profile database when the user joins the SIP domain. The user profile also stores information about the user subscribing to the call return and call screening services. Call return allows users to automatically call the most recent caller. Call Screening allows the notification facility to be triggered only for a particular set of URIs. If the user is not located due to timeout of SICAR and SIPAR, then a SMS is send to the user alerting him of the missed call. The alert messages to the users can be sent in their local languages as this would make the system more accessible to the non-English speaking population.

5.3. SICAR

SICAR requires the user's mobile device to have a client program running on it. The registration expiration information is send by the SIP Registrar server to this client program. The client program on receiving the notification automatically registers the user's current mobile device at the registrar. Thus SICAR provides the facility of automatically registering a user's mobile device without his intervention. However, the user has the choice to cancel the request as "Do Not Disturb".

In a WAP environment the PPS sends a PAP Push Submission message to the PPG. This submission message consists of the Mobile station International ISDN (MSISDN) number of the user as the device address. On receiving this push submission message the PPG accepts the message only if the message confirms the Document Type Definition (DTD) specifications of the PPG and if PPG is ready to deliver the message to the OTA client. The PPG then pushes the message on to the client device. The message is addressed to the client program residing on the user's mobile device. The control entity of the message is set to service load so that the client program on receiving the message automatically starts the registration procedure. The client program then sends its response back to the PPG containing a registration request with the user's SIP Authentication password and MSISDN number of the user's mobile device.

The PPG decodes the information and forwards it to the PPS. The PPS acts as a proxy for the user and sends a SIP Registration message on his behalf to the Registrar of the domain. The registrar may first require the user to authenticate him; in that case the PPS has to resend the registration message with the user's encrypted password. The registrar then registers the user and binds the contact address with his SIP URI. It then sends an acknowledgement back to the PPS indicating a successful registration, containing all the current bindings for the user. The PPS then sends the acknowledgement back to the user. If the user has also subscribed for the call return facility then the user should be prompted about making the last call again. This call then would be made as a normal IP telephony call and would be routed through the signaling and media gateway. Figure 4 shows message flow in SICAR.

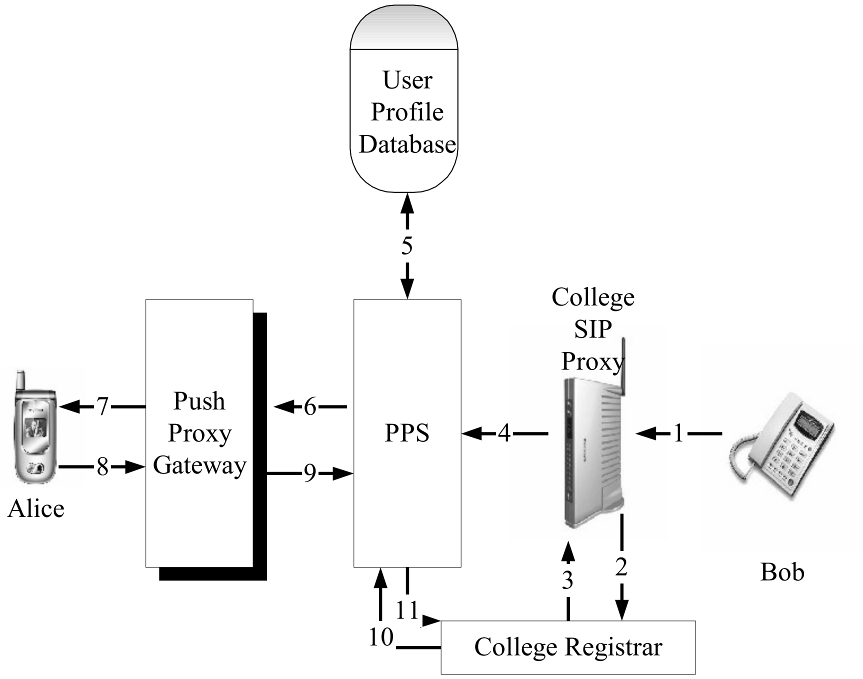

Figure 4. SICAR

General Packet Radio Services allows mobile devices to access internet simultaneously with other mobile applications at high speeds. In GPRS networks a user's mobile device can be in active, standby or idle states. In the active and standby states the Serving GPRS Support Node (SGSN) is aware of the mobile location and for these states the SICAR model of solution can be used. In idle state there is no logical GPRS context for the Mobile Station (MS) and the SGSN is not aware of the MS's location, so in idle states the SIPAR model of solution can be used.

For users in GPRS networks the solution for expired registrations can be provided by two means

- GPRS can act as a bearer for WAP. In this model of solution the same WAP Push solution proposed would be used, but the delivery mechanism for the WAP information would be GPRS.

- In the standby state only the mobile device's routing area is known. This routing area can comprise of a one or more cells. In this state the SGSN first pages the MS to determine its actual location and only then the data transfer can start. After this the PPS sends the alert message to the Gateway GPRS Support Node (GGSN). GGSN routes the message to the SGSN which further caters to the MS. The alert message is addressed to the client program residing in the MS.

5.4. SIPAR

In SIPAR the user's mobile device doesn't contain a client program running on it. The registrar expiration is send as an alert message to the user. The user can then register any device on which he wishes to receive the calls. Thus SIPAR requires human intervention to refresh the registration bindings. For users using WAP enabled phones the PPS pushes the alert message to the PPG. The PPG then pushes the message on to the client device. The control entity of the message is set to service indication so that only an alert is send to the user and depending on user preference the page can be loaded. A Wireless Markup Language (WML) page must be used to display the contents of the Registration page. The displayed page must query the user for his password, address of the device to be registered and the time period till which this current binding value would be valid. The user after filling in the required entries submits the page. The content of the page is then send to the WAP Gateway where it is decoded and then further send to the PPS. The PPS then contacts the Registrar and registers the new device. If the user on receiving the alert notification wishes to register later then he can disallow the loading of the registration page.

For users not having internet enabled phones the alert notification is send as a Short Messaging Service (SMS).

The PPS in this case acts as the Short Messaging Entity (SME) and submits the message at the Short Messaging Service Center (SMSC). The SMSC forwards the message to GSMSC, which further after consulting the Home Location Register (HLR) routes the message to the user's mobile device.

5.5. Security Issues

There are certain privacy and security issues to be handled in this kind of model.

Privacy - A user might not wish to be disturbed during a particular activity like while he is in a meeting. In such cases the user can set a flag in the user profile database, which indicates that he should not be disturbed.

Message Integrity - A user while registering through the WAP Push message will be required to enter his password for authorization at the Registrar. The bearer sends this password over the air (OTA) to the PPG. In these situations some unauthorized person can sniff the password or may try to alter the message contents. This kind of security threat can be avoided by using Wireless Transport Layer Security (WTLS) [16] security mechanism available in WAP.

Imposter Registrations – If the information in the user profile database is not updated like the user has a new MSISDN number, then the WAP Push message will be send to some other user leading to imposter registrations. To avoid such registrations it is mandatory that the SIP Registrar server first performs user authorization before making changes.

5.6. User Profile Database

The user profile database is maintained at the proxy server of the SIP domain. It contains the following information

- The SIP URI of the user.
- MSISDN number of user's mobile device.
- Flag set to determine if the user subscribes to the alert notification mechanism.
- Flag set to determine if SICAR or SIPAR is to be used.
- Information to determine the kind of bearer technology used for sending the alert mechanism to the user. This can be GPRS, WAP or SMS.
- Language in which alert notification are to be received.
- Flag set to true if call return is enabled.
- Flag set to true if call screening is enabled.
- List of SIP URIs for which the alert messages should be sent.
- Default expiration time for the contact binding.
- The email id of the user where the missed call information will be sent.

6. Conclusions

The solutions projected in this paper can be enhanced further by integrating Unified Messaging within it. For users whose location isn't determined by the SICAR or SIPAR, the alerts and notifications can be sent as an email or as voice mail. We have shown that a push mechanism of the cellular networks can be effectively used for solving the problem of missed calls in SIP network by refreshing expired SIP registrations The ubiquitous nature of mobile devices and their widespread usage provides better presence information about users than IP networks. Thus we have here demonstrated that better presence information available in cellular networks can be effectively ported to IP networks, thereby improving the presence information in IP based telephony networks. Using WAP Push/SMS has a further great advantage of using the existing cellular infrastructure thereby allowing the service providers to provide additional services without requiring further investments. This kind of solution brings together the cellular, voice and data networks together and achieves convergence in true sense. We expect in future more such applications will come up which would further harness the services of cellular networks.

References

[1] J Rosenberg, H Schulzrine, G Camarillo, A Johnston, J Peterson, R Sparks, M. Handley, E Schooler, (2002) "*Session Initiation Protocol*" IETF RFC 3261.

[2] H Schulzrine, S Casner, R Frederick, V Jacobson, (2003) "*RTP : A Transport Protocol for Real Time Applications*" IETF RFC 3550.

[3] R Stewart, K Morneault, H Schwarzbauer, T Taylor, I Rytina, M Kalla, L Zhang, V Paxson, (2000) "*Stream Control Transmission Protocol*" IETF RFC 2960.

[4] "*Universal Mobile Telecommunication Systems (UMTS); Multimedia Messaging Service (MMS)*"; 3GPP TS 22.140 version 6.7.0 Release 6, Available at www.etsi.org

[5] "*General Packet Service Registration (GPRS) Service description Stage 2*" ETSI EN 301 344 v7.4.1 (2000-09) available at www.etsi.org

[6] "*Universal Mobile Telecommunication Systems (UMTS); General UMTS Architecture*" 3GPP TS 23.101 version 6.0.0 Release 6. Available at www.etsi.org

[7] Ashok Jhunjhunwala, Anuradha Ramachandran, Alankar Bandyopadhyay, (2004) "*n-Louge: The Story of a Rural Service Provider in India*". The Journal of Community Informatics, Vol. 1 , Issue 1, 30-38.

[8] Arjun Roychowdhury and Stan Moyer, (2001) "*Instant Messaging and Presence for SIP Enabled Network Appliances*". Proceedings of Internet Telephony Workshop.

[9] J Rosenberg, H Schulzrine, (2002) "*Session Initiation Protocol (SIP): Locating Sip Servers*" IETF RFC 3263.

[10] A B Roach, (2002) "*Session Initiation Protocol (SIP): Specific Event Notification*" IETF RFC 3265.

[11] Open Mobile Alliance "*SIP Based Push Requirements*" OMA-RD-SIP_Push-V1_0-20050130-C

[12] Gunther Pospischil, Johannes Stadler, Igor Miladinovic, (2001) "*A Location-based Push Architecture using SIP*" Wireless Personal Multimedia Communication.

[13] A Johnston, S Donovan, R Sparks, C Cunningham, K Summers, (2003) "*Session Initiation Protocol Basic Call Flows*" IETF RFC 3665..

[14] Fergus Wills "*WAP Push Technology Overview*" Openwave Technical Document (www.openwave.com)

[15] "*Wireless Application Protocol Architecture Specification*" (2001), WAP-210-WAP Arch -20010712.

[16] "*Wireless Transport Layer End-to-end Security*" (2001), WAP-187-TLE2E-20010628-a.

[17] Wireless Application Protocol www.wapforum.org.

[18] Open Mobile Alliance www.openmobilealliance.org.

[19] J Lennox, H Schulzrine, J Rosenburg,, (2001) "*Common Gateway Interface for SIP*" IETF RFC 3250.

[20] J Lennox, H Schulzrine, (2000) "Call Processing Language Framework and Requirements", IETF RFC 2824.

[21] Voice Extensible Markup Language (VOICEXML) http://www.w3.org/TR/voicexml20/

Author

R. A. Akerkar is a professor of computer science and chairman of Technomathematics Research Foundation, India. He is also a guest professor at NTNU, Norway. His current research focuses on computer network systems, hybrid computing, and semantic web.