

Mobility support for wireless streaming multimedia in MarconiNet

Ashutosh Dutta, Wai Chen

Telcordia Technologies, 445 South Street, Morristown, NJ 07960

Henning Schulzrinne

Computer Science Department, Columbia University, New York, NY 10027

Onur Altintas

Toyota InfoTechnology Center, 4009 Miranda Avenue, Palo Alto, CA 94304

Abstract—

Real-time streaming content (audio and video) is mostly RTP/UDP based application which has stringent delay and loss requirements. Mobility affects the delay and loss factors for the multimedia stream delivery to a great extent, because of associated continuous handoff. This paper describes novel methods for supporting continuity of multimedia stream for the mobile clients in MarconiNet environment. In MarconiNet environment the end hosts could be (auto)mobile and these move from one cell to another, where each cell may belong to the same subnet or different subnet and different domain while receiving continuous multimedia stream content from several sources proxied via different local servers. Multimedia stream delivery to an end-host in MarconiNet's mobile environment depends upon the delay which is a composite of several factors such as cell/subnet detection, registration/configuration with new address, join/leave latency for the multicast streams. While movement detection, registration and configuration delays are similar to unicast stream delivery, join/leave latencies play an important role for delivery of multicast stream. This paper touches upon the latency factors associated, focuses on how media delivery due to mobility is affected by the join/leave latency and devises ways to improve that.

I. INTRODUCTION

Providing flexible streaming services for the mobile users in a wireless environment is gaining increasing momentum. These flexible services include multimedia content distribution in a scalable way, local and global program management, local advertisement insertion, and location based services. MarconiNet[1], [2] provides such a platform by adopting local server based control, augmented by an appropriate security and payment model between the global stations and local affiliates. Its architecture is based on standard IETF protocols such as SIP, SAP, SDP for signaling, RTSP for stream control, and RTP/RTCP for media delivery and feedback control. Mobility support when added to MarconiNet's streaming architecture provides a useful feature for mobile Internet application like M-commerce.

MarconiNet architecture is based on a hierarchical based multicast approach with global multicast and local multicast. In each cell (subnet) there is a local server

(Radio Antenna Server) which converts globally scoped incoming multicast stream to a locally scoped multicast stream. This server can act like a router but it can also be a proxy by performing application layer routing. There can be multiple sources, each multimedia stream delivery is associated with a unique globally scoped multicast address.

The architectural details and description of each functional component has been described in [1], [2]. However we present a brief overview of the architecture for clarity in section II.

In this paper we talk about the details of the techniques that would help provide these services in a highly mobile environment. This paper proposes several methods for taking care of continuous multicast stream delivery when the client is moving from one cell (subnet) to another cell (subnet) while reducing the latency.

This paper is organized as follows. Section II provides a brief overview of MarconiNet, Section III cites some of the related work, Section IV defines the latency factors involved in the delivery of this streaming multimedia. Section V details the possible alternatives to take care of seamless mobility for the multimedia stream. Section VI briefly introduces the testbed. Section VII concludes the paper with some open issues.

II. MARCONINET OVERVIEW

Figure 1 provides a brief overview of MarconiNet for streaming application, and Figure 2 provides a logical architecture. MarconiNet provides a business model for multimedia streaming distribution over the Internet in a scalable way. As illustrated in the Figure 2 there are 3 basic elements in MarconiNet architecture, namely Global Station (Primary Station), Radio Antenna Server (RAS) which are the local servers and Internet Multimedia Clients (IMC). Primary station acts like a multimedia streaming source. The Antenna server keeps an association between the globally scoped multicast address to the locally scoped multicast address.

By introducing the local servers one can have full con-

The application layer approach (specifically SIP) aims to keep mobility support independent of the underlying wireless technology and may fit better into the area which MarconiNet addresses, i.e. the provision of multimedia services.

SIP is basically an application layer signaling protocol which allows two or more participants to setup a multimedia session. Several wireless technical forums, such as 3GPP, 3GPP2, and MWIF have agreed upon SIP as the basis of the session management of the mobile Internet.

Since it seems that SIP will eventually be part of the mobile Internet protocol architecture, it may make sense to leverage some of its inherently present mobility support functions. SIP can help provide personal mobility, terminal mobility, session mobility and service mobility. For the purposes of MarconiNet, in this paper, we will elaborate on the terminal mobility support, however a comprehensive discussion on other SIP based mobility types can be found in [5]

Basically, supporting terminal mobility is to provide the ability to move with the same end-device while a session (call) is active. With SIP, being mobile does not require to modify (or add) capabilities to existing terminal's operating system. If the mobile host moves during an active session, first it receives a new address from a DHCP server (or a variant of it), and then sends a new session invitation to the correspondent host. With this new invitation, it tells its new IP address (contact) so as to forward the packets properly. Actually, this invitation is nothing more than updating the current ongoing session description. Then the mobile host also updates its registration with the home SIP server so that new sessions that would come can correctly be redirected to its new address.

In case of multimedia stream applications, basic advantage of using SIP for terminal mobility support is that the mobile host (using Mobile IP) does not have to go through first registering with its home agent to tunnel the binding updates to the corresponding host. This way, SIP helps cutting the hand-off delays.

However, whether SIP-based mobility is more appropriate for the purposes of MarconiNet needs to be investigated further. Note that the SIP mobility support approach does not necessarily exclude the Mobile IP approach, rather it may work to complement it. This paper addresses some of the ways to address the fast delivery methods in MarconiNet context.

In general, during a node movement, signaling and transport delays contribute to the latency of multimedia stream delivery associated with any two-party or multi-party communication session. Generally mobility can be categorized as pre-session and mid-session. Pre-session mobility generally does not contribute to the delay associated with the on-going session, but may add delay to any new upcoming session. Signaling such as registering

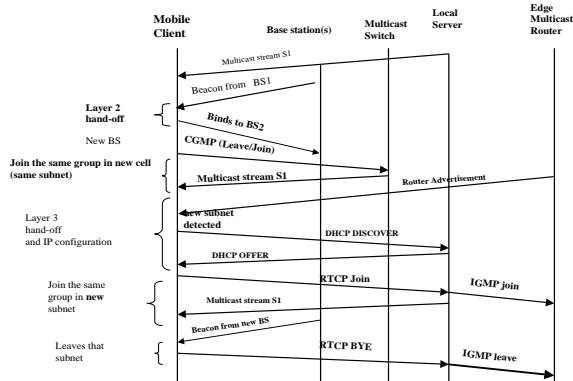


Fig. 3. Handoff flow

with a new server, notifying the communicating party with the mobile node's new contact address, inviting another user to a streaming session, would generally constitute signaling delay, while transport delay dominates the delay component associated with the mid-session mobility.

Latency associated with receiving continuous unicast/multicast stream from the same source while the client moves to the next cell would consist of several components such as detection of a new cell/subnet/domain, address acquisition, triggering for multimedia stream to be delivered in the new subnet and actual delivery of multimedia stream. While some of these factors are common to both unicast and multicast stream delivery (e.g., cell/subnet detection), we concentrate on multicast delivery scheme in this paper.

During a node movement latency is composite of actual hand-off time and join interval. Handoff delay would include the time to detect that it is in a new cell/subnet/domain, time for obtaining an IP address from a DHCP[28] or DRCP[12] server if it is moving between the subnets, and some triggering mechanism that would help initiate the multimedia flow in the new location. This would decide the join/leave latency of the multimedia flow.

Figure 3 shows a basic diagram with standard sequence of events for micro and macro mobility, and in next Section these steps are discussed in details. In this diagram some of the interim message flows (e.g., DHCP ACK etc.) are omitted for brevity. For domain mobility there are other factors such as AAA (Accounting, Authentication and Authorization) that would play a big role in deciding the delay factor for media delivery, because of establishing security association between the local servers in each domain.

A. Cell/Subnet Movement Detection

As the mobile moves away from its own cell, and begins to move to a new cell, it would likely to receive the

stream from a new local server. However in some cases in order to provide proper load balancing, there may be two local servers in the same cell served by the same base station. In that case load balancing needs to be provisioned to make sure that the client is achieving proper QoS requirement by switching to the proper server. Discovering a new cell/subnet/domain can be realized in different layers. During mobile's handoff, first movement detection takes place in layer 2, where the client would decide to switch over to a new base station based on the signal strength. In case of CDMA, soft handoff is initiated, so that each client can listen to both the base stations, and receive the stream, and then it decides as to which one would be accepted from the mixed signal. As soon as it switches over to a new base station and layer two handoff is over, it would need to figure out if the client is in a new subnet or domain altogether. Using layer 3 triggering mechanism such as router advertisement method similar to ICMP advertisement in Mobile IP it can be determined if the client is in a different subnet. An application layer detection method can as well be used, if the client is involved in a real-time communication session. However it may be faster to achieve the handoff notification using layer 2 mechanism.

B. Address acquisition by mobile

In the cases where Mobile IP's Foreign Agent approach is not instituted, as the client moves from one cell to another, where the new cell is another subnet it would either obtain a new address from the DHCP/DRCP server or can use standard Mobile IP approach to obtain a new care-of-address from the foreign agent while keeping its own IP address unchanged. There are other approaches such as cellular IP [8], HAWAII [9] for intra-domain mobility where IP address does not change once it enters a new domain. But some of these use source based routing which adds more complexity and overhead on the routers within a domain. In case of subnet movement typical time for acquiring a DHCP address would be about 15 seconds [21], although there are various other alternatives to take care of this problem such as DRCP[12] which is a light-weight version of DHCP suitable for wireless roaming environment. Using DRCP the address acquisition takes place within a few hundred milliseconds.

Although subnet detection and address acquisition are important factors for the delay, focus of this paper is about about the multicast session, join and leave latency mostly. These delay factors would be very important for channel surfing for the mobile users.

C. Triggering Multimedia stream

In traditional layer 3 method, triggering delay would consist of IGMP query report after the node moves to the new cell so as to be part of the same multicast tree. But if there is currently at least another active partici-

pant in the subnet, the mobile host can continue to receive the traffic without waiting to hear the membership query from the router. If the client is moving within a subnet but between the cells, it does not have to spend time in obtaining a new IP address, but rather handoff is taken care of at layer 2. In this case the triggering of multimedia stream can be taken care of by a variation of Generalized Switch Management Protocol [16] such as CGMP (Cisco Group Management Protocol)[17] or a variation of IGMP snooping [17] can be instituted at layer two. This method is a typical variation of layer 2 multicast. It would allow the multicast traffic to be filtered out at the multicast switch so as to save valuable wireless bandwidth when there is no active participant in the adjacent cell which is part of the same switch.

A typical query interval for the IGMP is by default 125 seconds, although this value is configurable in the multicast routers. In order to avoid the flooding in the LAN with IGMP messages this value cannot be made very small. Reference [3] shows that by using IGMP, a host will wait for 65 seconds on average in order to continue to receive the multicast traffic after the hand-over. This is because IGMP was not designed for roaming clients in a wireless environment where handoff latency and packet loss are important concerns. Similarly a typical leave latency once the host has moved to a new subnet is about 2 mins, i.e. traffic would still flow to the previous cell even after the client has moved out, thus wasting bandwidth in the previous cell.

In order to maintain minimum loss and latency during the client's movement it would be desirable to minimize the hand-off time and to provide almost instantaneous flow of multicast stream by adopting some new triggering mechanism. Similarly it may be required to avoid the waste of bandwidth because of flow of traffic due to leave latency.

In MarconiNet environment, the mobile clients use local common multicast address for indexing and use a locally scoped multicast address to tune to a particular stream from a source. Although global multicast address remains the same, there is a likelihood that the local multicast address on each of the server's local interface would be different. The mapping association between global and local address is stored locally in Radio Antenna Server's database. In order to make sure that each moving client keeps on pointing to the same multicast address or belong to the same group membership, some kind of mechanism would be necessary to pre-provision. It can be quite possible that the mapping between the global and local address association is determined by a policy decided at the local station/Antenna Server. In order to make it more effective, there can be a rule based association between the global address and local address, say using an offset of N on each octet (at least between the servers on the same domain). Otherwise if the map-

ping is different at each local server, some kind of proactive mechanism is needed to inform about the local address for the stream delivery, before the actual stream gets delivered.

V. STREAM DELIVERY METHODS

A fast delivery mechanism of multicast streaming is intended here to avoid the loss of real-time data and improve the latency. These could be achieved in several different ways before or after the client is registered/configured in the new network/subnet. This mechanism can be implemented as part of pre-registration, during-registration, post registration to the new cell (subnet), or can be obtained via proxy registration. Following are the possible approaches for taking care of triggering multicast streaming in a MarconiNet environment, when a client moves from one subnet to another.

In MarconiNet architecture, each local server (station) has two interfaces, one is an interface with a global address and the other may possible be a local address. However this does not necessarily have to be opaque to the outside world. In Figure 4, Ia, Ib, Ic are the globally known subnets connected to one of the interfaces of the servers where as ia, ib, ic are the local subnets connected to the secondary interfaces of the servers and could be local to that area. According to the MarconiNet model, server gets the multicast stream through its global interface and redirects it out through the local interface for the local clients in each cell. In this particular picture, S1, S2, S5 are local stations (servers) connected to the upstream routers. Each server (with the exception of S2 and S3) is connected to a different subnet and has a separate interface. S2 and S3 are connected to the same subnet via a multicast switch. Each station serves one cellular region underneath, which could be part of a private subnet dedicated for the local users. The base stations are not shown here, which could be IP based or just layer 2 aware base stations, and in some cases a local server can behave as the base-station on one of its local interfaces. As described in reference [18] mobility management becomes easier when the base stations are IP based, in fact in some cases these local servers can also act as the IP based smart base stations with dual interfaces. The other alternative is to connect the second interface of the server to a non-IP based base station (only layer two), which would perform handoff based on layer 2 functionality. In this figure, S2 and S3 are connected to a multicast switch which then become part of the same subnet, who can manage the traffic using GSMP (General Switch Management Protocol or CGMP for Cisco based routers). For a scenario involving S2 and S3, multicast traffic is filtered at layer two, thus saving network bandwidth in each of these cells. Figure 4 shows a block diagram implementing this scenario.

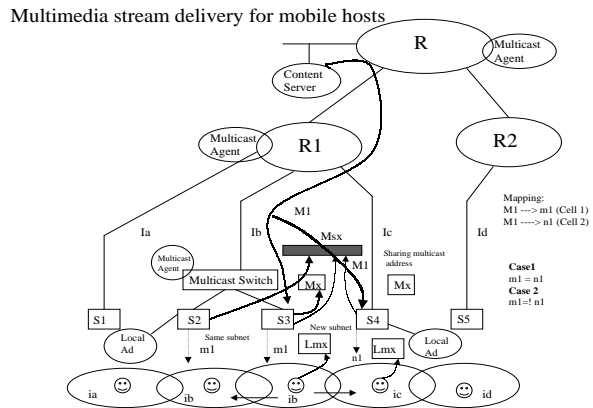


Fig. 4. Fast-handoff scenario

A. Layer two scenario

In cases where the destination cell belongs to the same subnet and both the cells are served by the same local server, then multicast stream flows in both cells in the absence of any multicast switch. Although this would save the triggering time, it would however contribute to the waste of bandwidth, if there is no active participants in the adjacent cell. However it would be a trade-off between the triggering time and bandwidth saved. Typically as soon as the change of base station is detected, CGMP Join/Leave is triggered thus starting the flow of a multicast stream or keeping it from flowing to the next cell.

B. Post-Registration

Post-registration approach is the most common approach but it takes the maximum time for the same multicast stream to get directed to the new cell. In a typical scenario, the client moves to a new cell, obtains the new IP address depending upon if it is moving to a new subnet, and then sends the join query via IGMP method. In those cases IGMP would be modified to provide aggregate group report, as suggested in [7]. But in this case there is still a latency associated with it which is related to IP address acquisition and, which could be avoided by other approaches as detailed below.

Another approach would be to use an application layer triggering mechanism such as RTCP as in Reference [1] in order to facilitate the join and leave of the session. (Experimental results of an analysis of advantage of using either of the methods are being investigated currently). In this case the end-client who is listening to a local multicast stream would send the triggering signal by sending RTCP JOIN/BYE using the SDES field of RTCP. In this scenario, triggering at the lower hierarchy would be accomplished by RTCP, but the server would trigger the multicast flow from the upstream router using IGMP method.

C. Pre-registration Method

Each local station (server) can have multiple neighboring (servers) stations. For each of these stations sharing an overlapping area with another station there would be a multicast announcement (address) associated that would be pre-provisioned, where each station can point to find out the program subscribed to (group address) by the impending mobile host. Just before a mobile node leaves (decides to leave) the cell which could be based on the threshold value of the layer 2 signal, or some other soft handoff methodology as described in CDMA, it would send an RTCP message to the local announcement address, the server in turn would announce that to the sharing multicast addresses where the neighboring stations would be listening to in the global space. The neighboring stations (servers) look up to the multicast address and check it with their own databases to see if this multicast stream is already subscribed to. If this is the case, since other clients have been listening to the same stream, then it does not do anything, if it is not playing then it sends an IGMP message to the upstream router and passes the stream to the local cells using a locally scoped multicast address, even before the client has moved to the new cell. Thus virtual soft hand-off is emulated for the associated stream. As soon as the client moves in to the next cell it can still get the same stream without any interruption. The only tradeoff with this scheme is that bandwidth would be wasted a bit because the multicast stream would be playing beforehand. Similarly the client would send an RTCP BYE to the server as it moves away from the previous server. This approach is better than the approach described in [4] where the client has to send a unicast packet to the neighboring station, that means the client has to know the addresses of the neighboring servers beforehand.

D. Pre-registration with Multicast Agent

This section defines the use of a Multicast Agent to take care of the multicast streams.

With this method there would be a multicast agent within each router which would forward the global stream to the respective global multicast addresses (e.g., for areas where these clients are trying to move) in each subnet for a specific period of time, so that each neighboring server would receive the stream irrespective of whether the mobile node is moving into that cell or not. As soon as the mobile node moves into the new cell it sends an RTCP signal to alert that the mobile node has moved in, so the respective timer does not need to get triggered. However this approach would create waste of bandwidth in each of these subnets for a period of time.

E. During Registration

In another approach, group association information can be passed on as a part of registration method. When

the node moves in, and tries to acquire the IP address in the local subnet it can send the request for that particular stream in its DHCP Discover message about the address it has been listening to. But in that case the server would have to be a registration server. At the time of obtaining the IP address from the DHCP server, the client can send the locally scoped multicast address to the server, and depending on whether the server is already part of the multicast tree, it would ignore this request or re-join the tree. It remains to be seen if the client can use RTCP before the client obtains a DHCP address in the new domain.

F. Proxy Registration

Another approach would be to deploy proxy agents in each subnet. These proxy agents would join the upstream multicast tree on behalf of the servers, even before the clients move into the cell. The neighboring proxy servers would then be listening to a common multicast address to figure out the impending host's subscribed multicast address. In this case the multicast proxy would be sending the IGMP query messages beforehand on behalf of the local servers.

VI. MOBILITY REALIZATION IN AN EMULATED TESTBED

Figure 5 shows a IEEE 802.11 based testbed, where mobility part of MarconiNet is being experimented currently. This testbed provides both indoor and outdoor mobility capability, and emulates a wireless Internet. In this testbed all three kinds of mobility scenario namely micro, macro and domain are being tested out. Application layer triggering methodology is being implemented to provide a faster join and leave latency. Each of the servers has a media database for local advertisement. In addition to be able to explore faster stream delivery methods for different kinds of mobility scenario, other aspects of MarconiNet such as QoS management based on real-time feedback of streaming content, and different location based services such as localized advertisement insertion, inviting another client to a multicast stream are being realized. AAA (Accounting, Authorization and Accounting) servers are consulted when each adjacent cell belongs to a new domain altogether.

VII. CONCLUSION AND OPEN ISSUES

In this paper we discussed the latency factors associated and presented several scenarios for obtaining fast delivery of multimedia stream in MarconiNet environment. The focus has been on faster join and leave latency for the clients (using both network layer and application layer approaches), and timely delivery of multimedia content for the moving clients. One of the open issues is how to pre-provision the mapping of the global multicast address to the local ones in all the neighboring

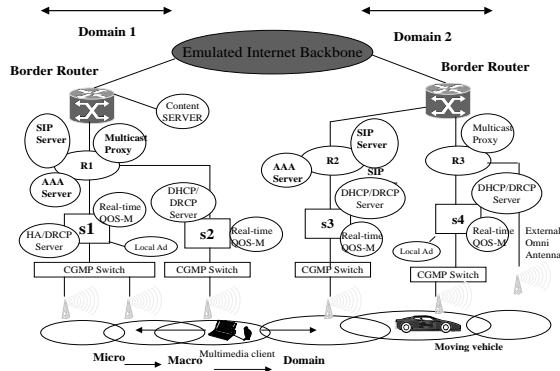


Fig. 5. MarconiNet mobility realization

servers. If the mapping is not identical in the neighboring stations then we need to find a different way of conveying the existing mapping in the previous cell so that the neighboring server would apply this mapping dynamically for the impending host. It also remains to be seen if SIP based approach can be used here to take care of delivery of localmulticast stream.

Each of these approaches is being implemented in MarconiNet test-bed so that performance of each category can be analyzed in the final version of the paper. An analysis of join/leave latency for network layer vs. application layer is also being studied.

VIII. ACKNOWLEDGEMENT

The authors would like to thank Subir Das, Archan Misra, and Anthony McAuley for several discussions on fast handoff issues, and wish to thank Dr. Mitsuo Nozawa of Toyota InfoTechnology Center whose comments helped improve the format of the paper.

REFERENCES

- [1] A. Dutta, H. Schulzrinne and Yechiam Yemini, "MarconiNet - An Architecture for Internet Radio and TV Networks" in *Proc. of NOSSDAV 1999*
- [2] A. Dutta, H. Schulzrinne, "A Streamig Architecture for Next Generation Internet" in *To appear in the proceedings of ICC 2001*
- [3] M. Flament et. al, "An approach to 4th Generation Wireless Infrastructures- Scenarios and Key Research Issues". VTC 99, May 1999
- [4] Jiang Wu, "An IP mobility support for 4GW wireless Infrastructure" 1999 Personal Computing and Communication Workshop (PCC Workshop 99), November 1999
- [5] H. Schulzrinne and E. Wedlund, "Application-layer mobility using SIP", *ACM Mobile Computing and Communications Review*, vol.4, no.3, July 2000.
- [6] Giuseppe Anastasi et al, "Group Multicast in Distributed Mobile Systems with unreliable wireless network" in the Proceedings of the 18th IEEE Symposium on Reliable Distributed Systems (SRDS'99), October 1999.
- [7] Satwant Kaur et al, "Multicast Support for mobile IP using a modified IGMP" in the Proceedings of the IEEE Wireless Communications and Networking Conference, 1999 (IEEE WCNC'99), September 1999.
- [8] Andrew T. Cambell, Javier Gomez, Andras Valko, " An overview of Cellular IP" IEEE Wireless Communications and Networking Conference, September 1999.
- [9] R. Ramjee, T. La Porta et al, "HAWAII: A Domain-based Approach for Supporting Mobility in Wide-area Wireless networks, in the Proceedings of 7th IEEE International Conference on Network Protocols, 1999(ICNP'99), October 1999
- [10] A.J. McAuley et al, "Mobile Multicast Proxy" in the Proceedings of the IEEE MILCOM Conference 1999, October 1999
- [11] G. Xylomenos and G.C Polyzos, "IP Multicast for Mobile Hosts", *IEEE Communications Magazine*, January 1997
- [12] A.J. Mcauley, S. Das et al, "Dynamic Registration and Configuration Protocol" IETF Draft, work in progress.
- [13] Upkar Varshney and Samir Chatterji, "Architectural Issues to support multicasting over wireless and Mobile networks" 1998 WCNC
- [14] Elin Wedlund, H. Schulzrinne, " Mobility Suport using SIP" in *IEEE/ACM Multimedia conference WOWMOM 1999*
- [15] A.Dutta, F. Vakil, J.C Chen, S. Baba, N. Nakajima, H. Schulzrinne, "Application Layer Mobility Management Scheme" to appear in 3G Wireless 2001 conference.
- [16] Avri Doria, Kenneth Sundell, "draft-ietf-gsmv-08.txt", General Switch Management Protocol V3, IETF work in Progress
- [17] B. Williamson,"Developing IP Multicast Networks: The Definitive Guide to Designing and Deploying CISCO IP Multicast Networks, Volume I, January 2000, Cisco Press".
- [18] A. Misra, S. Das, A. Dutta, A. McAuley, "IDMP-based Fast Handoffs and Paging in IP-based cellular networks" to appear in 3G Wireless 2001.
- [19] W. Fenner, "Internet group management protocol, version 2," Request for Comments (Proposed Standard) 2236, Internet Engineering Task Force, Nov. 1997.
- [20] Kevin Almeroth, Mostafa Ammar "Multicast Group Behaviour in the Internet's Multicast Backbone (MBone)", *IEEE Communication Magazine*, June 1997.
- [21] J-O. Vatn and G.Q. Maguire Jr., "The effect of using co-located care-of addresses on macro handover delay, in the Proceedings of the 14th Nordic Tele-traffic Seminar (NTS 14), August 1998
- [22] Perkins et al., "IP Mobility Support", Request for Comments (Proposed Standard) 2002, Internet Engineering Task Force, October 1996
- [23] M. Handley and V. Jacobson, "SDP: session description protocol," Request for Comments (Proposed Standard) 2327, Internet Engineering Task Force, Apr. 1998.
- [24] M. Handley, "SAP: Session announcement protocol," Internet Draft, Internet Engineering Task Force, Nov. 1996. Work in progress.
- [25] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: a transport protocol for real-time applications," Request for Comments (Proposed Standard) 1889, Internet Engineering Task Force, Jan. 1996.
- [26] T. Friedman and D. Towsley, "Multicast session membership size estimation," in *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, (New York), Mar. 1999.
- [27] H. Schulzrinne, A. Rao, and R. Lanphier, "Real time streaming protocol (RTSP)," Request for Comments (Proposed Standard) 2326, Internet Engineering Task Force, Apr. 1998.
- [28] R. Droms "Dynamic Host Configuration Protocol" Request for Comments (Proposed Standard) 2131, Internet Engineering Task Force, March 1997
- [29] D.A. Ferguson, "Measurement of mundane TV behaviors: Remote control device flipping frequency", *Journal of Broadcasting Electronic Media*, 38(1), 35-47