

A Strategy for Best Access Point Selection

V. Ghini, S. Ferretti, F. Panzieri
Department of Computer Science
University of Bologna
Bologna, Italy
{ghini, sferrett, panzieri}@cs.unibo.it

Abstract—This paper describes a scheme for providing mobile multimedia applications, distributed in a Wireless Local Area Network (WLAN), with seamless communications. Specifically, when multiple WLAN access points are available, the scheme enables a mobile node to select dynamically the access point to connect with, based on a combination of measured parameters such as signal strength, transmission rate, channel load and interference. The proposed scheme has been implemented as part of a cross-layer architecture which offers the possibility of concurrently exploiting multiple networks for the same communication flow generated by some distributed multimedia application. The results which have obtained from an experimental assessment of our scheme demonstrate its adequacy and effectiveness.

Keywords- Access point selection, resource-intensive multimedia applications, wireless communications, multi homing

I. INTRODUCTION

Today, there is an increasing demand for accessing and running distributed applications in mobile nodes. Wireless communications must be efficiently supported and exploited to offer seamless services to mobile clients. In practice, the services accessed by the mobile clients must be supported by the underlying communication infrastructure so that the interactions between the services and their mobile clients are not interrupted, and the service results delivered timely.

In this context, we have developed a cross-layer architecture that allows an application running in a multi-homed mobile node to exploit concurrently the multiple network interfaces available on that node, so as to guarantee the continuity (and maximize the performance) of the communications between that mobile node and its remote peers. Such an architecture can be viably exploited to support several kinds of distributed, wireless (multimedia) applications, such as Voice over IP, multimedia streaming, wireless online gaming, discovery of points of interest [1-4]. In essence, our approach exploits a proxy-based solution, i.e. a client proxy is installed on the mobile device while a server-proxy is deployed on the Internet. The two proxies interact and are in charge of reconstructing the communication flow based on the data transmitted from/to different network interfaces at the mobile client. Such an expedient makes transparent to the (client/server) applications the fact that the client node concurrently exploits different IP addresses for a single application data flow.

Having a scheme that allows to select dynamically the best network among different alternatives opens the issue of determining which one is the best to choose at each instant of the communication. This means that a selection must be made among the different available network technologies (WLAN, UMTS, etc.), and that the best Access Point (AP) must be identified among those available for a single technology. In this paper, we focus on IEEE 802.11 WLAN technology and on the specific issue of selecting the best Access Point (AP) for accessing this technology.

The use of WLANs might be a problem for certain kinds of mobile users running certain applications, since these technologies have a limited coverage, multiple obstacles can reduce the transmission range and a large number of concurrent users can consume the available bandwidth [5, 6]. The system we developed has the following main characteristics: it 1) tolerates temporary disconnections with the wireless AP so as to guarantee highly available communications, 2) ensures enough bandwidth for multimedia data transmission, 3) limits the power consumption of the mobile station transmission so as to reduce the electromagnetic radiations absorbed by the human operator. The software module which is in charge of selecting the best AP to connect with is named Access Point Monitor (APM). It operates in the mobile node and is capable of dynamically selecting the best AP quickly, using a combination of metrics. In particular, it analyzes the strength of the received signal together with other metrics related to the traffic load and channel interference. Results coming from an experimental evaluation of an implementation of the APM we have developed confirm the viability of our proposal.

The rest of this paper is structured as follows. In Section II, we outline the system architecture that embodies the software module for the selection of the AP mentioned above. Section III describes the method we employ for monitoring the APs and selecting the most appropriate AP. Section IV reports on the results obtained from an experimental evaluation we have carried out of our proposal to assess its viability. Finally, Section V provides some concluding remarks.

II. SYSTEM ARCHITECTURE

We have developed a cross-layer architecture which, implemented in a multi-homed node, offers *always-connected* services by exploiting all the networks available to the user, and by dynamically adapting their use on the basis of their performance and costs. Thus, our architecture can provide its

users with reliable communications even in presence of vertical and horizontal handoffs, and can optimize those communications, augmenting the probability of meeting application QoS requirements such as responsiveness, availability, and continuity of the service. (In this paper, we specifically focus on the sub-task of identifying which is the best access point to use, when a Wi-Fi technology is exploited.)

Our cross-layer architecture is able to manage different types of communication flows, based on either the UDP or the TCP transport protocols. Basically, it offers an adaptive use of different network technologies, transparently to the applications, by exploiting a pair of proxies for each mobile device. A client proxy needs to be executed on the mobile device, while a server proxy executes over the Internet in a fixed host with a public IP address (see Figure 1). In particular, such a cross-layer software architecture (that works at the datalink, network, transport and session layers) is instantiated in the client proxy which is in charge of managing the wireless networks and dynamically deciding which network to use. The server proxy is aware that each client can use multiple nets, and that messages from the same client (i.e., belonging to the same communication flow) may be received coming from different IP addresses. Hence, once identified all the messages that belong to the same client, it rebuilds the communication flow from that client and forwards that client's messages to their destination. Thus, the server proxy appears as the client to the destination host. Moreover, the pair of proxies implements packet load balancing and retransmission on the wireless channels to meet application QoS requirements.

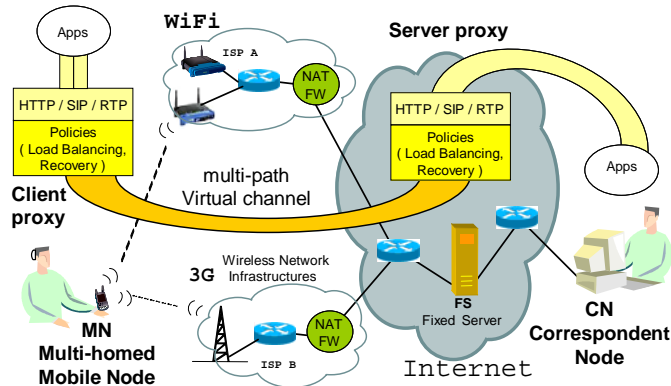


Figure 1. Cross-Layer Mobility Management Architecture

Our mobility management mechanism overcomes a number of limitations of other existing approaches. In particular, the Mobile IPv6-based architectures, such as Fast Handover MIPv6 [7] and Proxy MIPv6 [8], need that network infrastructures be modified to add IPv6 capabilities and, moreover, do not allow the simultaneous use of the multiple network interfaces of the mobile node. Others solutions at the ISO/OSI transport layers, such as Datagram Congestion Control protocol [9] and Mobile Stream Control Transmission protocol [10] require the end user applications be modified. In contrast, our solution does not require modifications of the access networks or the applications that may use “proxies”, such as Web-based and SIP/RTP-based multimedia applications.

III. ACCESS POINT MONITOR SELECTION

A. Background

As of today, the majority of WLANs are based on the IEEE 802.11 standard. When a communication with an infrastructured WLAN must occur, the mobile station has to select (and associate with) an AP. Usually, the mobile station tries to associate with the AP with the highest signal strength, due to the rate adaptation mechanism (named multi rate protocol) of the 802.11's physical layer. This mechanism allows the mobile station to tune its transmission rate in response to the quality of the wireless link to the AP. If the quality of the link is poor, then a robust modulation is required, together with coding schemes that reduce its actual transmission rate. Conversely, the higher the signal strength the higher the transmission rate. However, it has been observed that the classic approach of selecting the AP with the highest signal strength may not be appropriate when it is crucial to select the AP that provides a large amount of bandwidth. Moreover, such traditional policy may lead to poor user-perceived performance and highly unbalanced load distribution [11]. Conversely, schemes should be considered that take into account different parameters, including the following.

Maximum allowed transmission rate: it depends on the hardware of both an AP and a given mobile station. It represents the upper bound to the transmission rate used in the communications between the mobile station and the AP.

Actual average transmission rate: this parameter represents the average transmission rate used in the communications between the AP and the mobile station. It mainly depends on the quality of the link: the worse the link quality, the lower the actual transmission rate (lower transmission rate means decreasing both the loss rate and the need for retransmissions). This parameter strongly impacts the channel occupation introduced below.

Channel occupation: this parameter represents the time a channel is used for a communication between an AP and a given mobile station. It depends on both the traffic produced/received by the nodes and the specific traffic transmission rate. In fact, the lower the transmission rate the longer the channel occupation by a mobile station.

Channel interference: it represents the interference level that affects a given AP; it can be caused by different external sources, that include other APs. The interference depends on the signal strength of the APs and their mobile stations.

To conclude, it is worth noticing that, due to the contention mechanisms for the transmission in the IEEE 802.11 environment, an increment of the traffic that can affect a given AP would increase the waiting time before a transmission, thus decreasing the responsiveness of the AP.

B. The Proposed Solution

We have designed a software tool named Access Point Monitor (APM), to be run in the mobile node. It allows to select the best AP based on the knowledge of its surrounding environment and on the network background traffic it perceives. The APM operates as follows. Firstly, it monitors all

the communication channels; in particular, it “sniffs” and captures all the frames related to those channels and builds an empirical evaluation of the actual traffic it perceives on each channel. Then, the APM builds a sorted list of best APs using the traffic evaluation and the corresponding achievable hypothetical bandwidth.

In the model, the following assumptions are made: the 802.11 physical layer uses the DS-SS (Direct Sequence Spread Spectrum) with a time slot $T_{slot} = 20\mu s$; the 802.11 MAC layer uses the Distributed Coordination Function (DCF) with a minimum contention window $CW_{min} = 15$. Based on these configurations, we obtain a value for the Distributed InterFrame Space which is $DIFS = 68\mu s$. Finally, we assume that the random backoff time (hereinafter denoted as $T_{backoff}$) is, on average, $T_{backoff} = T_{slot} * CW_{min} / 2 \approx 150\mu s$.

A scheme for measuring the traffic at a given AP may work as follows. As the mobile node needs some AP, the APM starts monitoring each channel for a fixed monitoring interval T_{mon} . For each channel, the APM computes the time spent by the channel to transmit packets over the T_{mon} as follows:

$$T_{pkt}^{Ch} = \sum_{\forall pkt \in Ch} \left(\frac{size_{preamble}}{rate_{preamble}} + \frac{size_{pkt}}{rate_{pkt}} + DIFS + T_{backoff} \right) \quad (1)$$

Equation (1) computes the traffic generated on a channel Ch , by measuring the time to transmit each packet the APM receives on the channel Ch . This time is obtained by summing the time to send a packet preamble, the time to send the data packet itself (note that we have different rates for sending packets and their preambles), the DIFS and the above estimated $T_{backoff}$, which represent the time between the transmission of a frame and the next one.

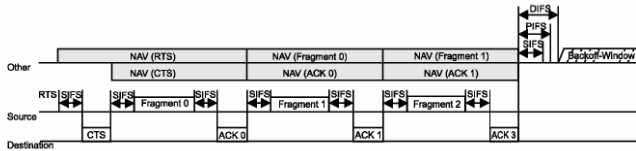


Figure 2. Transmission of a sequence of fragments

However, it is worth observing that this equation is valid only when a network packet is included within a single data-link fragment. Indeed, the equation does not take into consideration several factors. First of all, the equation does not include the time the channel in which the channel is occupied due to the transmission of the acknowledgment for the packet that follows immediately the reception of the packet. Moreover, the equation does not include the transmission of the RTS (Request to Send) and CTS (Clear To Send) frames, used to agree on a frame transmission in the WLAN. Typically, this RTS/CTS packet exchange is used when a packet is large and it is sent as a sequence of consecutive fragments. Moreover, between the transmission of a RTS and the subsequent CTS, there is a time interval $SIFS = 10\mu s$ (Single Inter Frame Space). Moreover, equation (1) does not take into account the interferences on the channel, which may lead to packet losses, which cause retransmissions. Then, when the packet to be transmitted is composed of multiple fragments (namely, a

sequence), after having obtained the clearance to send the packet, the sending node transmits the entire sequence of fragments. In this case, for each transmitted fragment, the sending node waits for an acknowledgment from the receiving node (see Figure 2). After each fragment transmission, a SIFS time interval elapses. All these different transmission possibilities may be easily monitored at the mobile node by simply looking at the information contained within the frames. As an example, when a fragment is retransmitted, the MAC layer sets the bit “retry” in the MAC header. When multiple fragments are consequently transmitted for a single packet, instead, the *More Data* field is set equal to 1.

In view of the above observations, it is possible to measure the time needed to send the sequence of fragments that compose a packet as

$$T_{seq}^{Ch} = \sum_{\forall Seq \in Ch} \left(\sum_{\forall frag \in Seq} \left(\frac{size_{preamble}}{rate_{preamble}} + SIFS + \frac{size_{frag}}{rate_{frag}} + SIFS + ACK \right) + DIFS + T_{backoff} \right) \quad (2)$$

Analogously, the time needed to send a single short packet that does not require the preliminary RTS/CTS packets exchange is

$$T_{pkt}^{Ch} = \frac{size_{preamble}}{rate_{preamble}} + SIFS + \frac{size_{pkt}}{rate_{pkt}} + SIFS + ACK + DIFS + T_{backoff} \quad (2')$$

When a fragment of a sequence is lost the sequence transmission is stopped immediately and the retransmission of that sequence is restarted from that lost fragment at the cost of an additional preliminary RTS/CTS packets exchange. Thus, it is possible to measure the time needed to resend packets that need to be retransmitted

$$T_{fraglost}^{Ch} = \sum_{\forall fraglost \in Ch} \left(\frac{size_{preamble}}{rate_{preamble}} + SIFS + \frac{size_{fraglost}}{rate_{fraglost}} + DIFS + T_{backoff} \right) \quad (3)$$

or, for a single-fragment packet

$$T_{fraglost}^{Ch} = \sum_{\forall fraglost \in Ch} \left(\frac{size_{preamble}}{rate_{preamble}} + SIFS + \frac{size_{fraglost}}{rate_{fraglost}} + SIFS + ACK + DIFS + T_{backoff} \right) \quad (3')$$

Hence, percentage of the total transmission time of the channel Ch can be calculated as

$$T_{tr}^{Ch} = \frac{T_{seq}^{Ch} + T_{fraglost}^{Ch}}{T_{mon}} \quad (4)$$

The percentage of available time (i.e., T_{free}) for the transmission on the channel Ch over the monitoring interval is equal to:

$$T_{free} = 1 - T_{tr}^{Ch} \quad (5)$$

Once obtained T_{free} , this value can be employed to estimate the hypothetical bandwidth an AP can provide on the channel Ch . In this case, there are different possibilities that may arise, depending on both the size of the packets the mobile node will send, and the variations of the 802.11 protocol implementations that are installed on the AP and the mobile devices. For instance, if the mobile device sends frames which are composed of a single frame and there is no use of RTS/CTS frames, then we obtain

$$BW_{hyp} = T_{free} * [\max size_{pkt} / (\frac{size_{preamble}^{Ch}}{rate_{preamble}^{Ch}} + \frac{\max size_{pkt}}{rate_{hyp}^{Ch}} + DIFS + T_{backoff})] \quad (6)$$

Equation (6) takes into account the rate to send the packet preamble, the maximum size (i.e., 1518B) of the packets allowed by the 802.11 standard, and the hypothetical rate for sending the packets. The hypothetical rate an AP can reach depends on the signal strength. We have evaluated the hypothetical rate of an AP based on the RSSI. Table 1 shows this evaluation. The “dBm” row of Table 1 represents the minimum signal strength for which it is guaranteed the transmission rate shown in the corresponding row “Mbps”. For example, with -77dBm the hypothetical rate cannot be greater than 36 Mbps. With strengths greater than or equal to -71dBm the hypothetical rate is always 54 Mbps.

TABLE I. HYPOTHETICAL CHANNEL RATE

dBm	-92	-91	-90	-88	-87	-86	-85	-83	-81	-78	-72	-71
Mbps	1	2	5.5	6	9	11	12	18	24	36	48	54

Vice versa, if the mobile node sends only frames which are included in a single fragment, but nevertheless the device is configured to utilize RTS/CTS frames before its data transmissions, an estimation of the available bandwidth is as follows

$$BW_{hyp} = T_{free} * \left[\max size_{pkt} / \left(\frac{RTS + 2 * SIFS + CTS + \frac{size_{preamble}^{Ch}}{rate_{preamble}^{Ch}}}{\frac{\max size_{pkt}}{rate_{hyp}^{Ch}} + ACK + DIFS + T_{backoff}} \right) \right] \quad (6')$$

Note that in this case, several RTS/CTS are sent and several waiting times $T_{backoff}$ occur.

In general, when a frame is sent, which is composed of more fragments, a single pair of RTS/CTS is employed for each frame. In this case, we may obtain an upper bound of the available bandwidth of the AP by neglecting the time related to the transmission of the (small sized) RTS/CTS frames, and considering only the time related to the transmission of fragments. This (impossible) case corresponds to the situation when a single frame, composed of a long sequence of fragments, is transmitted during the whole time interval T_{free} .

$$BW_{hyp} = T_{free}^{Ch} * [\max size_{frag} / (\frac{size_{preamble}^{Ch}}{rate_{preamble}^{Ch}} + \frac{\max size_{frag}}{rate_{hyp}^{Ch}} + 2 * SIFS + ACK)] \quad (6'')$$

To conclude, according to our approach the mobile node

selects the AP with the highest hypothetical bandwidth computed through (6), (6') or (6''), depending on the specific implementation of the IEEE 802.11 protocol one is considering.

C. Implementation

The architecture introduced above has been implemented in a Linux-based system. The monitoring and configuration of both datalink and network layers of the network adapters of a mobile host are managed by the software component (hereinafter termed Monitor) that coordinates with an implementation of the APM, the wpa_supplicant software tool, and the DHCP client component. To identify disconnections with a given AP, the Monitor communicates with the Linux kernel through a datagram-oriented Netlink socket. This way, when a network adapter changes its status, the kernel informs the Monitor, which in turn activates the adapter reconfiguration. As to the general architecture, the Monitor also communicates with modules placed at higher (session) layers that manage the different connections through different network interfaces.

As to the APM, it works at the datalink layer and it is implemented using the Linux Wireless Extensions API. The APM guides the adapter in scanning the wireless channels, creating a sorted list of the possible candidate APs for that adapter. The list is passed to the Monitor component that excludes from it the APs already associated with a different wireless adapter of the same MS, and the APs for which there are no information required for the authentication process. The wpa_supplicant is then responsible for carrying out the association and authentication procedures with the first available AP in the list. Finally, the DHCP client configures the adapter's network layer.

IV. EXPERIMENTAL EVALUATION

We have assessed the system we have implemented to understand how the APM behaves in real wireless scenarios. In this section, we report on some of the experiences made with such real tests. Basically, we were interested on the APM ability to select the best AP among a set of candidate ones, when these APs (i) are affected by different traffic configurations, (ii) exhibit different signal strengths, and (ii) use radio channels that may be overlapped. We used two different APs, involved in different types of network traffic.

As a first test, we used the two APs working on non-overlapping channels and at different maximum datarate, in particular 54 and 11 Mbps. Two different scenarios were assessed; in particular, we evaluated the results obtained without any traffic on the two APs, and then with some traffic (10.2 Mbps) on the faster AP A. It is clear that, under the same traffic conditions (no traffic), the higher the AP's bandwidth the higher the quality measured by the APM, as reported in Table II (last column). Conversely, when we add the traffic on A, the APM suggests to employ the more unloaded (even if slower) AP B; in fact, using equation (6'), we obtain a higher value for the estimated bandwidth.

To assess how the distance of an AP influences the outcomes of the APM, we varied the distance of the two APs

and the types of traffic. In these tests, both APs were set to transmit data at the maximum datarate of 54 Mbps on non-overlapping channels. Table III reports results, obtained using equation (6''), when A was placed at an approximate distance of 8 meters with respect to the mobile host, while B was placed at 40 meters. As before, traffic in A (the nearer AP) was varied. It may be observed that while in the first case A is selected as the AP to connect with, B is chosen when traffic through A is present.

TABLE II. QUALITY INDEX OF THE TWO APs DEPENDING ON MAXIMUM ALLOWED DATARATE AND TRAFFIC

AP	Maximum Datarate (Mbps)	Traffic (Mbps)	Quality Index (Mbps)(6') No Traffic / with Traffic
A	54	- / 10.2	17.79 / 3.30
B	11	-	6.75 / 6.75

TABLE III. QUALITY INDEX WITH APs AT DIFFERENT DISTANCES

AP	Approx. Distance	Traffic (Mbps)	Quality Index (Mbps)(6'') No Traffic / with Traffic
A	Near (8 meters)	- / 10.2	27.04 / 13.66
B	Far (40 meters)	-	16.95 / 16.95

Previous tests were made using the two APs working on non-overlapping channels for the transmission, i.e., 1, 6, in order to avoid interferences. An interesting issue is also to understand how the quality index measured by the APM is influenced when the APs work in overlapping channels. Tables IV and V show results when the two APs work on the same channels, with different traffic types. Not only, in this case we report results obtained using both equations (6') and (6'') to estimate the available bandwidth. It is interesting to observe that the measured indexes are the same, even when the traffics destined to the two APs are very different. This is due to the fact that the APs have to deal with the aggregate traffic, since the channel is the same.

TABLE IV. QUALITY INDEX OF APs WORKING ON THE SAME CHANNEL WITH NO TRAFFIC

AP	Channel	Traffic (Mbps)	Quality Index (Mbps) (6'') / (6')
A	1	0	27.10 / 17.51
B	1	0	27.10 / 17.51

TABLE V. QUALITY INDEX OF APs WORKING ON THE SAME CHANNEL, WITH DIFFERENT TRAFFIC TYPES

AP	Channel	Traffic (Mbps)	Quality Index (Mbps) (6'') / (6')
A	4	8.4	6.65 / 4.29
B	4	-	6.65 / 4.29

Finally, the influence of the concurrent traffic can be observed when the two APs do not operate on the same channel, but on overlapping ones, as shown in Tables VI and VII. In this case, the quality index is not the same, but it is clear an high traffic through A influences the quality measured at B (see the result on Table VI, with respect to Table VII).

TABLE VI. QUALITY INDEX WITH APs WORKING ON OVERLAPPING CHANNELS

AP	Channel	Traffic (Mbps)	Quality Index (Mbps) (6'') / (6')
A	1	0	27.39 / 17.69
B	2	0	27.39 / 17.69

TABLE VII. QUALITY INDEX WITH APs ON THE SAME CHANNEL, DIFFERENT TRAFFIC TYPES

AP	Channel	Traffic (Mbps)	Quality Index (Mbps) (6'') / (6')
A	4	9.2	10.27 / 6.63
B	5	0	19.31 / 12.47

V. CONCLUSIONS

In this paper, we have presented a scheme, termed Access Point Monitor (APM) that enables dynamic selection of the best AP reachable by a mobile node. To this end, the selection policy uses both the signal strength and the traffic and packet losses on the wireless communication channels. Such a scheme has been implemented as part of a distributed architecture that allows mobile nodes to use seamlessly utilizing multiple wireless networks for the same (application-layer) communication flow. An experimental evaluation of this architecture, described in this paper, has shown the adequacy of our APM scheme.

REFERENCES

- [1] S. Ferretti, V. Ghini, "Discovering Points of Interests Through A Web 2.0, Location-based Architecture", in *International Journal On Advances in Internet Technology*, IARIA, issn: 1942-2652, Vol. 2, N.2-3, 2009, 238-252.
- [2] S. Ferretti, V. Ghini, "A Web 2.0, Location-based Architecture For A Seamless Discovery Of Points of Interests", in *Proceedings of 5th Advanced International Conference on Telecommunications (AICT 2009)*, Venice (Italy), IEEE Communications Society, May 2009.
- [3] V. Ghini, S. Ferretti & F. Panzneri, "Mobile Games Through the Nets: a Cross-Layer Architecture for Seamless Playing", in *Proceedings of the International Workshop on Distributed Simulation & Online gaming (DISIO 2010) - ICST Conference on Simulation Tools and Techniques (SIMUTools 2010)*, Torremolinos (Spain), ICST, March 2010.
- [4] V. Ghini, G. Lodi, F. Panzneri, "Mobile E-Witness", *Multimedia Tools and Applications*, 37(3): 293-318 (2008).
- [5] G. Kasbekar, J. Kuri, and P. Nuggehalli, "Online Association Policies in IEEE 802.11 WLANs," in *Proc. of 4th Int'l Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, 1-10, 2006.
- [6] T. Korakis, O. Ercetin, S. V. Krishnamurthy, L. Tassiulas, and S. K. Tripathi, "Link Quality based Association Mechanism in IEEE 802.11h compliant Wireless LANs", in *Proc. of 1st Workshop on Resource Allocation in Wireless Networks (RAWNET)*, 2005.
- [7] R. Koodli, "Fast Handover for Mobile IPv6", IETF RFC 4068, July 2005.
- [8] S. Gundavalli et al., "Proxy Mobile IPv6", IETF Internet Draft, draft-ietf-netlmm-proxymip6-01.txt, June 2007.
- [9] E. Kooler et al., "Datagram Congestion Control Protocol (DCCP)", IETF RFC 4340, March 2006.
- [10] M. Riegel, M. Tuexen, "Mobile SCTP," IETF Internet draft, draft-riegel-tuexen-mobile-sctp-07.txt, Oct. 2006.
- [11] A. Balachandran, P. Bahl, and G. Voelker. "Hot-spot congestion relief and service guarantees in public-area wireless networks", *SIGCOMM Computer Communication Review*, 32(1), 2002.