

Service Differentiation via Adaptive Gateway Discovery in Ad Hoc Networks Connected to Wired Networks

Mari Carmen Domingo

Telematics Engineering Department, Catalonia University of Technology (UPC),
Av. del Canal Olímpic s/n, 08860 Castelldefels (Barcelona), Spain
+34 93 413 70 51
cdomingo@entel.upc.edu

Abstract. A gateway discovery mechanism is necessary to allow wireless nodes in an ad hoc network to route their packets towards a fixed network. Real-time applications have quality of service parameters and require a gateway discovery mechanism that helps them to maintain their requirements. Therefore we propose in this work a new adaptive gateway discovery scheme that adjusts the frequency of the gateway advertisements dynamically. This protocol is able to differentiate services between applications and it cooperates with real-time flows to maintain the desired quality of service. Simulation results investigate the performance of the proposed adaptive scheme and show its effectiveness in comparison with the hybrid mechanism in the simulated scenarios.

1 Introduction

Ad hoc networks [1] have been designed as wireless mobile devices that are able to communicate without having to resort to a pre-existing network infrastructure and without the intervention of a system administrator.

Originally, the investigation was centered in developing isolated and independent ad hoc networks to cooperate in military operations or natural catastrophes like hurricanes. However, this kind of networks is restricted to certain particular environments. Therefore, more recently, the attention has been focused in studying the interaction between ad hoc networks and other types of networks like cellular networks, infrastructure-based WLANs (Wireless Local Area Networks) [2] or wired networks.

The communication between wireless ad hoc networks and infrastructure-based networks is essential to extend Internet beyond its traditional scope, to remote inaccessible areas, making Web services available anytime, anywhere. In addition, ad hoc networks can be seen as an easy way to reduce the congestion in hotspot areas, allowing users to communicate themselves directly without the presence of an access point like in infrastructure-based IEEE 802.11 WLANs; besides, ad hoc networks are able to access Internet via a gateway as well.

The mobile nodes in a wireless ad hoc network must be able to detect available gateways and select one of them if they want to have Internet access.

Real-time applications have special quality of service requirements that must be satisfied to function properly.

We argue that new gateway discovery mechanisms should be designed thinking over the requirements of real-time flows because the selected gateway discovery mechanism will affect the overall performance of the ad hoc network.

Our objective will be to design a new gateway discovery protocol that helps real-time flows to maintain their quality of service parameters. Some different approaches have been developed in literature that propose different gateway discovery schemes, but none of them is related to service differentiation and quality of service improvement for real-time flows. This is the main contribution of this paper.

The paper is organized as follows: Section 2 describes related work about Internet gateway discovery methods. Section 3 remarks the importance of quality of service provision in wireless ad hoc networks. The proposed adaptive gateway discovery scheme is presented in Section 4. Section 5 shows our simulation results and finally Section 6 concludes this paper.

2 Related Work

In order to communicate the ad hoc and the Internet network packets must be transmitted to a gateway as it is illustrated in Fig. 1. This device implements the protocol stack of the ad hoc as well as the fixed network, routing the packets from one network to the other. The protocol stack used by mobile nodes, gateways and Internet nodes is shown in Fig. 2.

The Internet Draft “Global Connectivity for IPv6 Mobile Ad Hoc Networks” [3] describes how to provide Internet connectivity to mobile ad hoc networks modifying an existing routing protocol (like AODV [4] in the example) so that it is able to discover gateways.

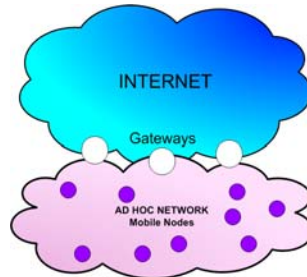


Fig. 1. Interworking scenario.

Three main approaches have been developed to detect gateways:

- Proactive gateway discovery [5] [6]: The gateways periodically broadcast advertisement messages that contain information about the global prefix length and the IPv6 address from the gateway. These messages are flooded throughout the entire network. The mobile nodes use this information to autoconfigure a new routable IPv6 address and select the address of one of

the gateways as default route. The mobile nodes select the best Internet-gateway by its distance in hops or by other parameters.

- Reactive gateway discovery [7]: A mobile node that wants to send packets towards Internet broadcasts a message to the group of gateways within the ad hoc network. The gateways receive this message and reply to it accordingly. The mobile node selects the gateway which offers the best route towards Internet in terms of number of hops or other parameters.
- Hybrid gateway discovery [8] [9]: This method combines the reactive and proactive approaches; it defines a transmission range where the gateways periodically send advertisement messages and they are propagated around a limited zone (a certain number of hops away from the gateway). A mobile node receiving these messages can obtain information about the global prefix length and the IPv6 address from the gateways carried in this message to discover the global prefix. Afterwards, this mobile node autoconfigures a new routable IPv6 address and selects the address of one of the gateways as default route. The mobile nodes select the gateway that is either closer in terms of number of hops or that is more appropriate because of other parameters. If a mobile node wants Internet connectivity and it is outside the gateways transmission range and the propagation zone of the gateways advertisements, it broadcasts a message to the group of gateways in the ad hoc network. If another mobile node receives this message, it rebroadcasts it until it arrives to a gateway that responds sending back a reply. The mobile node selects the reply of the gateway which offers the best route towards Internet in terms of number of hops or due to other parameters.

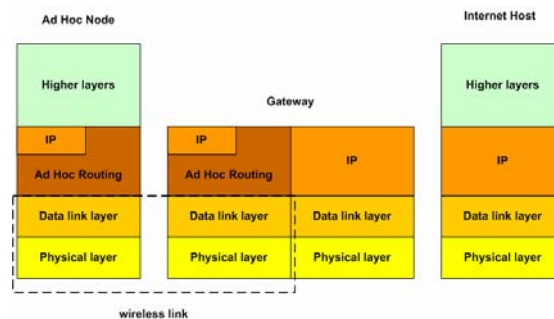


Fig. 2. Protocols architecture.

From here on the different approaches that have been proposed in the literature are modifications of the already mentioned gateway discovery strategies.

However, these existing approaches are methods to discover gateways that treat all the traffic in the same way and do not consider differences between real-time and best-effort applications. In the next sections we will remark the importance of providing quality of service to real-time applications in wireless ad hoc networks and of introducing gateway discovery mechanisms that differentiate service levels between best-effort and real-time traffic.

3 Quality of Service Provision in Wireless Ad Hoc Networks

Quality of Service (QoS) can be defined as the ability of the network to offer a required service demanded by a particular application, establishing some type of control over its end-to-end delay, jitter, traffic loss or bandwidth.

It is a very challenging topic to provide QoS in wireless ad hoc networks [10] due to the intrinsic properties of this kind of networks: variable capacity of the links, topologies that change dynamically, etc; furthermore, in wireless networks the packet loss rate and the jitter of the applications are higher in comparison with wired networks due to the existence of fading, interference between neighbouring nodes, etc.

Our objective is to provide QoS to real-time applications in wireless ad hoc networks, differentiating services between real-time and best-effort traffic.

We are interested in studying the performance of multimedia applications in wireless ad hoc networks connected to wired networks. We have selected a specific type of real-time application that implies burstiness and that contains end-to-end delay information: VBR Voice-over-IP (VoIP) [11]. The ITU-T recommends in its standard G.114 that the end-to-end delay should be kept below 150 ms to maintain an acceptable conversation quality [12]. Delays from 150 to 400 ms are acceptable provided that administrators are aware of the impact of quality, and latency larger than 400 ms is unacceptable.

There exists a relation between the QoS provisioning and the gateway discovery method. The hybrid and specially the proactive approaches have a better performance with respect to end-to-end delay, because GWADV messages are sent periodically and not only when it is needed, as in the reactive approach. Thus, real-time applications are able to find a route towards Internet for their traffic sooner. But, on the other hand, if a real-time application has delay problems due to congestion and more GWADV messages are sent, the congestion will be increased and the performance of the delay sensitive applications will be seriously damaged. In the next section we introduce an adaptive gateway discovery approach that has been mainly designed to reduce congestion problems and it helps real-time applications to maintain their QoS parameters even in the presence of excessive traffic.

4 Proposed Adaptive Gateway Discovery Mechanism

The proactive and hybrid approaches have a better performance with respect to end-to-end delay. We have selected as reference model a hybrid gateway discovery mechanism instead of a proactive one because the proactive schemes propagate the GWADVs through the entire network and therefore they introduce more overhead, a circumstance that could seriously damage real-time traffic in the presence of congestion. On the contrary, the GWADVs of the hybrid approach are propagated only a limited number of hops away from the gateway (advertisement zone).

A parameter of the hybrid approach that is directly related to the functioning of this method is the gateway advertisement interval. We have studied the consequences of sending GWADV messages. If mobile nodes receive a GWADV, not all profit in

the same way when they receive this gateway message: some nodes benefit and others are harmed with more packets that introduce more congestion and that they don't in fact really need. Consequently, the gateway advertisement interval should be carefully chosen.

We consider a network where best-effort and real-time traffic sources send traffic from the ad-hoc towards the fixed network. We want to provide quality of service, differentiating services between real-time and best-effort applications.

The destination nodes of the real-time traffic in the fixed network periodically monitor the end-to-end delays of these flows. To achieve it, a 'timestamp' or generation time of the packet is introduced in the header of the real-time application protocol (the RTP protocol (Real-time Transport)) and the average end-to-end delay is calculated at the destination node as a time difference. If the end-to-end delay of one or more real-time sources becomes greater than 140 ms, QoS_LOST messages will be sent to the real-time traffic sources that have latency problems to warn them about the situation (see Fig. 3).

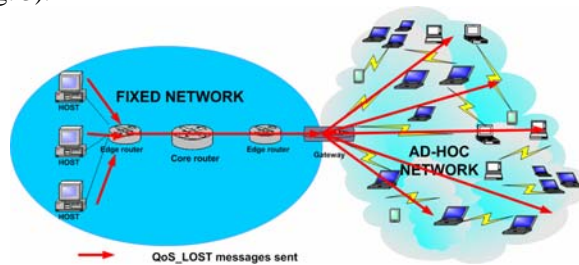


Fig. 3. QoS_LOST messages sent from the fixed towards the ad hoc network.

When a node in the ad hoc network receives a QoS_LOST message, it will react executing a QoS mechanism to improve the QoS of its real-time flow; for example the authors in [13] propose a new protocol, named DS-SWAN (Differentiated Services-Stateless Wireless Ad Hoc Networks), to support end-to-end QoS in ad hoc networks connected to one fixed DiffServ domain. DS-SWAN warns nodes in the ad hoc network when congestion is excessive in the ad hoc network for the correct functioning of real-time applications. These nodes react by slowing down best-effort traffic. Simulation results indicate that DS-SWAN significantly improves end-to-end delays of real-time flows without starvation of background traffic.

However, it is not the scope of the paper to suggest which QoS mechanism should be employed to improve the QoS of real-time flows.

We have made the following assumptions aiming to design our adaptive gateway discovery approach:

- Congestion appears in the ad hoc and not in the fixed network. A Diff-Serv domain in the fixed network may prevent that congestion is introduced in the core routes.
- The destination nodes periodically measure the end-to-end delays of the real-time flows and if these delays are larger than a threshold (140 ms, because the ITU-T recommends to keep these delays under 150 ms and the system needs some reaction time) then QoS_LOST messages will be

sent to the source. We are interested in the arrival of QoS_LOST messages to the corresponding gateway that is crossed when these messages travel towards the real-time sources.

- If a route for real-time traffic from the ad hoc towards the fixed network is broken, not only the source node but also the gateway should be warned about the situation with a Route Error (RERR) message. This means that the intermediate node that detects the link failure should send a RERR message not only to the real-time source node but also to the gateway.

We propose a new mechanism where the gateway periodically checks if it has received a RERR coming from a real-time application. In this case, the gateway sends a GWADV message unconditionally because thus this source will be able to find a new route towards the destination sooner with more probability. This does not necessarily mean that this source will sure find a new route by this procedure: It can be possible that the GWADV does not arrive to the source because it is more than TTL hops (advertisement zone) away from the gateway or because there is no route anymore available to reach this source. However, the gateway should not avoid sending this GWADV message because it should try to help a source with routing difficulties.

On the contrary, if the gateway has not received any RERR informing that a real-time source has routing problems, it should do the following:

The gateway should periodically check if it has received QoS_LOST messages during the last T seconds from real-time flows having problems to keep their end-to-end delays below 150 ms.

The gateway should calculate:

$$\alpha(t) = \frac{P}{F}, \quad (1)$$

where P = number of real-time sources having end-to-end latency problems and F = total number of real-time sources using that gateway.

We consider a threshold γ , where $0 \leq \gamma \leq 1$. It is fulfilled:

If $\alpha(t) \geq \gamma$, no GWADV messages should be sent by the gateway to the ad hoc network, because if real-time flows have QoS problems due to excessive congestion, it is not recommended to introduce more traffic overload in the network with these messages.

On the contrary, if $\alpha(t) < \gamma$, GWADV messages should normally be sent towards the ad hoc network.

This functioning method serves the purpose that real-time sources do not increase their end-to-end latency problems if congestion is excessive.

The value of γ should be carefully chosen by operators according to their own needs. In the extreme case $\gamma = 1$, the gateway discovery approach is equivalent to the hybrid scheme.

On the other hand, it is important to think about the role of background best-effort traffic sources. They don't have an active role in the decisions taken by the gateway to help real-time sources to discover the best route for sending their packets towards

Internet. Sometimes they will take advantage and sometimes they will be at disadvantage due to the intrinsic functioning of this gateway discovery approach; what is a fact is that this mechanism has not been designed thinking on them although they will try to profit from the situation if it is possible.

The functioning of this adaptive scheme is illustrated in Fig. 4.

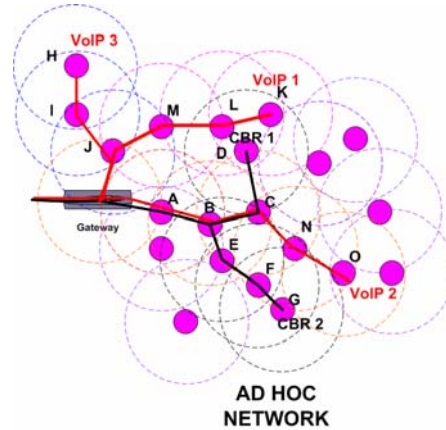


Fig. 4. Example network.

It shows an example of an ad hoc network where three VoIP real-time and two CBR best-effort flows have been established to send packets towards Internet through the gateway. If we consider that the VoIP flows VoIP1 and VoIP3 have problems to keep their end-to-end delays under 150 ms, QoS_LOST messages will be sent to these VoIP sources in the ad hoc network through the gateway to warn them about the situation. The gateway takes advantage of this information and it periodically calculates the percentage of VoIP sources that route their packets towards Internet through it and that have end-to-end delay problems. In our example this percentage is $\alpha(t) = 2/3$. If the threshold for latency problems is set to be $\gamma = 0.5$, then it follows that $\alpha(t) \geq \gamma$, which means that no GWADV messages should be sent by the gateway to the ad hoc network, because the number of VoIP sources having delay problems due to excessive congestion is larger than the threshold and this means that the network should not be overloaded with more traffic if it is not strictly necessary. If afterwards one of the VoIP sources solves its QoS problems, the gateway will calculate a new percentage $\alpha(t) = 1/3$. Now GWADV messages should be sent towards the ad hoc network because $\alpha(t) < \gamma$. The advertisement messages will be propagated around a limited zone (a certain number of hops away from the gateway); in this case we consider an advertisement zone of TTL = 4 hops. This means that the gateway advertisement messages will be received by the sources VoIP1 (route gateway-J-M-L-K), VoIP3 (route gateway-J-I-H) and CBR1 (route gateway-A-B-C-D). The other sources will not receive the GWADV messages because they are more than

4 hops away from the gateway and they would have to do a route discovery in the case that the route towards the gateway breaks.

5 Simulations

We have run simulations with the NS-2 tool [14] to investigate the performance of our proposed approach. The system framework is shown in Fig. 5. A scenario where an ad-hoc network is connected via two gateways to a fixed IP network has been selected. The chosen scenario consists of 20 mobile nodes, 2 gateways, 3 fixed routers and 3 corresponding hosts.

The mobile nodes are uniformly distributed in a rectangular region of 1000 m by 500 m. The gateways are placed with x, y coordinates (150,250) and (850,250). Each mobile node selects a random destination within the area and moves toward it at a velocity uniformly distributed between 0 and 3 m/s. Upon reaching the destination, the node pauses for a pause time, selects another destination and repeats the process. Five different pause times have been used: 0, 20, 50, 125 and 200 seconds. The dynamic routing algorithm is AODV [4] and the wireless links are IEEE 802.11b.

Background traffic is generated by 6 of the mobile hosts, while VBR VoIP traffic is generated by 15 of the mobile hosts. The destinations of each of the background and VoIP flows are chosen randomly among the three hosts in the wired network.

We assume that best-effort CBR background traffic and real-time VBR VoIP traffic are transmitted. We have proposed CBR as background traffic instead of TCP. The reason is that TCP performs poorly in an ad-hoc network because packets that are lost due to link failure and route changes trigger TCP's congestion avoidance mechanisms [15]. On the contrary, many authors [16] use CBR as background traffic successfully.

The VBR mode is used for VoIP traffic. We employ a silence suppression technique in voice codecs so that no packets are generated in the silence period. For the voice calls, we use the ITU G. 726 or "adaptive differential pulse code modulation (ADPCM)" codec. The VoIP traffic is modelled as a source with exponentially distributed on and off periods with 1.004 s and 3.587 s average each and two frames (20 ms audio sample each frame) are carried in each packet (80 + 80 bytes payload). Frames are generated during the on period every 20 ms with size 80 bytes and at a constant bit rate of 32 Kbps without any compression. VoIP is established over real-time transport protocol (RTP), which uses UDP/IP between RTP and link layer protocols. Packets have a constant size and are generated at a constant inter-arrival time during the on period. The VoIP connections are activated at a starting time chosen from a uniform distribution in [10 s, 15 s].

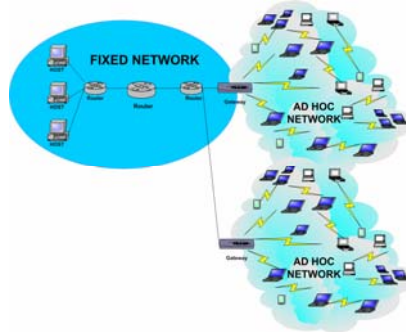


Fig. 5. Simulation framework.

Background traffic is Constant Bit Rate (CBR) with a rate of 48 Kbit/s and a packet size of 120 bytes. To avoid synchronization, the CBR flows have different starting times that have been randomly chosen.

We have run simulations to assess the following three performance measures:

- Average end-to-end delay for real-time (VoIP) traffic: Defined as the time it takes for data packets to arrive from the source node to the destination node.
- Packet delivery ratio: Defined as the number of real-time (VoIP) packets successfully delivered over the number of real-time (VoIP) packets generated by the sources.
- Routing overhead: Defined as the amount of control packets (for gateway discovery and routing) divided by the sum of the control packets plus the data packets.

We have evaluated and compared the performance of the hybrid approach “Hybrid scheme” with our proposed adaptive scheme discussed in Section 4: “Proposed adaptive scheme”. In both approaches a TTL=2 hops is used as advertisement zone.

Fig. 6 shows the average end-to-end delay for VoIP traffic. We can observe that in both schemes the end-to-end delays for VoIP traffic are increased with smaller pause times, because when the pause time is very low the routes of the existing flows break down frequently and the routing protocol continuously does new route discoveries processes that increase the latency. On the contrary, when the pause time is higher, the average link duration is increased as well as the duration of the routes.

We notice that the average end-to-end delays for VoIP sources are lower when our proposed adaptive scheme is used, because less GWADV messages are sent in congestion conditions. The gateway periodically checks if it has received QoS_messages associated with VoIP sources having end-to-end delay problems. If the percentage of VoIP traffic sources having latency problems exceeds a predefined threshold (in this case this threshold is set to $\gamma = 0.15$), no GWADV messages are sent by the gateway. Therefore, no more traffic overload is introduced in a congested network and as a consequence the latency of VoIP flows is diminished. When the pause time is 0 or 200 seconds (extreme cases), the routes break down frequently due to continuous mobility or because of the isolation of certain nodes that don't find

neighbours to act as intermediate nodes for their packets and as a result new route discoveries have to be done frequently and congestion is increased more; hence real-time sources have very often delay problems. If a scheme like the proposed adaptive scheme is used, the reduction of congestion is very effective in comparison with the hybrid scheme. A similar trend is observed regarding the jitter for VoIP traffic.

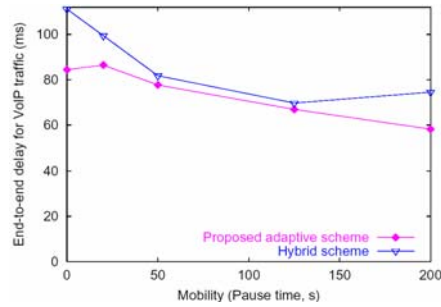


Fig. 6. Average end-to-end delay for VoIP traffic.

The packet delivery ratio for VoIP packets is illustrated in Fig. 7.

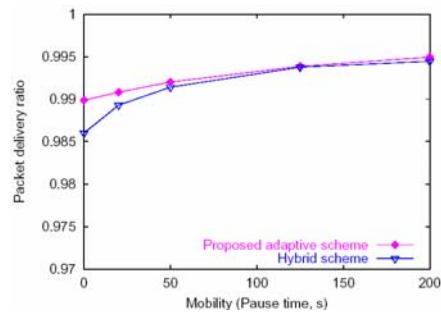


Fig. 7. Packet delivery ratio for VoIP packets.

As we can appreciate from the figure, the higher the mobility of the nodes, the best performs our approach in comparison with the other. There is a trade-off between the signaling overhead and the proactivity of the protocol in both approaches; however, when mobility is high and due to the frequently routes breakage, in the hybrid protocol signaling overhead is higher in comparison with the proposed adaptive approach. Consequently, more congestion is introduced and as a result, the packet delivery ratio is decreased. The differences in packet delivery ratio are lesser as mobility is decreased.

The overhead of control packets is depicted in Fig. 8. We can see that the proposed approach has a lower overhead than the hybrid approach. The differences in overhead are lower in the extreme cases (pause time 0, 20 and 200) because in the proposed approach less GWADV messages are sent by the gateways but, on the other hand, more sources start a route discovery mechanism as in a reactive approach when they need to find a gateway and moreover if a RERR message arrives to a gateway it sends

GWADV messages unconditionally so that the overhead is increased. However, the routing overhead is a metric that shows how much network resources does the protocol need to do its work and we can appreciate that the number of network resources is much lower using the proposed mechanism.

Besides, we have done more simulations that show that with the proposed scheme there is no starvation of best-effort traffic.

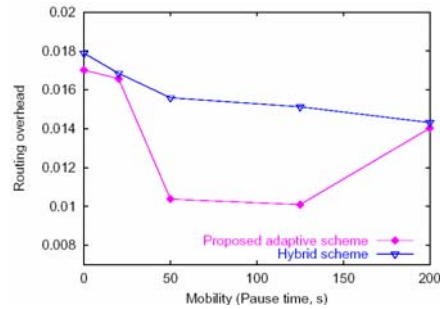


Fig. 8. Overhead of control packets.

6 Conclusions and Future Work

We have proposed a new adaptive gateway discovery protocol that differentiates services between applications through gateway selection. The scheme introduced outperforms the hybrid scheme in terms of average end-to-end delays and jitter for real-time flows, packet delivery ratio and routing overhead. What is more, using this scheme there is no starvation of best-effort traffic.

As future work we are planning to do more simulations that evaluate the effectiveness of our gateway discovery scheme when scalability is introduced (with respect to the number of real-time sources, the number of gateways, the traffic load, etc). Besides, we think it could be interesting to analyze the performance of the proposed protocol in a real network.

Acknowledgment

This work was partially supported by the "Ministerio de Ciencia y Tecnología" of Spain under the project TIC2003-08129-C02, which is partially funded by FEDER.

References

- [1] S. Basagni, M. Conti, S. Giordano, and I. Stojmenovic, "Mobile Ad Hoc Networking", IEEE Press & Wiley Inter-Science, 2004.

- [2] D. Cavalcanti, C. Cordeiro, D. Agrawal, B. Xie and A. Kumar, "Issues in Integrating Cellular Networks, WLANs, and MANETs: A Futuristic Heterogeneous Wireless Network", IEEE Wireless Communications Magazine, Vol. 12, No. 3, pp. 30-4, April 2005.
- [3] R. Wakikawa, J. T. Malinen, C. E. Perkins, A. Nilsson, and A. J. Tuominen, "Global connectivity for IPv6 mobile ad-hoc networks", Internet Engineering Task Force, Internet Draft (Work in Progress), July 2002.
- [4] C. E. Perkins, E. M. Belding-Royer, and I. Chakeres, "Ad Hoc On Demand Distance Vector (AODV) Routing", IETF Internet draft, draft-perkins-manet-aodvbis-00.txt, Oct 2003.
- [5] Y. Sun, E. M. Belding-Royer and C. E. Perkins, "Internet Connectivity for Ad Hoc Mobile Networks", International Journal of Wireless Information Networks, vol. 9, issue 2, 2002, pp. 75-88.
- [6] C. Jelger, T. Noel and A. Frey, "Gateway and address autoconfiguration for IPv6 adhoc networks", Internet-Draft, draft-jelger-manet-gateway-autoconf-v6-02.txt, April 2004.
- [7] J. Broch, D. Maltz and D. Johnson, "Supporting Hierarchy and Heterogeneous Interfaces in Multi-hop Wireless Ad Hoc Networks", in Proceedings of the IEEE International Symposium on Parallel Architectures, Algorithms and Networks, June 23-25, Perth, Australia, pp. 370-375.
- [8] P. Ratanchandani and R. Kravets, "A Hybrid Approach to Internet Connectivity for Mobile Ad Hoc Networks", in Proc of the IEEE WCNC 2003, Vol. 3, pp. 1522-1527.
- [9] J. Lee, D. Kim, J. J. Garcia-Luna-Aceves, Y. Choi, J. Choi and S. Nam, "Hybrid Gateway Advertisement Scheme for Connecting Mobile Ad Hoc Networks to the Internet", in Proc. of the 57th IEEE VTC 2003, Vol. 1, Jeju, Korea, April 2003, pp. 191-195.
- [10] K. Wu and J. Harms, "QoS Support in Mobile Ad-hoc Networks," Crossing Boundaries- the GSA Journal of University of Alberta, Vol. 1, No. 1, Nov. 2001, pp.92- 106.
- [11] D. Chen, S. Garg, M. Kappes and K.S. Trivedi, "Supporting VBR Traffic in IEEE 802.11 WLAN in PCF Mode," in Proc. OPNETWORK'02, Washington D.C., Aug. 2002.
- [12] ITU-T Recommendation G.114, "One way transmission time", May 2000.
- [13] M C. Domingo and D. Remondo, "An Interaction Model and Routing Scheme for QoS Support in Ad Hoc Networks Connected to Fixed Networks", vol. 3266 of Lecture Notes in Computer Science, Berlin, 2004, Springer Verlag, pp. 74-83, ISBN 3-540-23238-9.
- [14] NS-2: Network Simulator, <http://www.isi.edu/nsnam/ns>.
- [15] A. Jain, A. Pruthi, R.C. Thakur, and M.P.S. Bhatia, "TCP analysis over wireless mobile ad hoc networks", IEEE Personal Wireless Communications, New Delhi, India, Dec. 2002.
- [16] P.B. Velloso, M. G. Rubinstein and M. B. Duarte, "Analyzing Voice Transmission Capacity on Ad Hoc Networks", International Conference on Communications Technology - ICCT 2003, Beijing, China, April 2003.