TD(06)029

COST 290 6th MCM, Berne, Switzerland, May 09-10th 2007

QAMNet: Providing Quality of Service to Ad-hoc Multicast Enabled Networks

Harald Tebbe,
University of Ulm, Department of Media Technology, Oberer Eselsberg, 89069 Ulm, Germany
harald.tebbe@informatik.uni-ulm.de

Andreas Kassler
Karlstad University
Universitetsgatan, 65188 Karlstad, Sweden
kassler@ieee.org

Abstract

Mobile Ad-hoc networking has been considered as one of the most important and essential technologies that support future Pervasive Computing Scenarios and 4G networks. In a mobile Ad-hoc Network (MANET), support for multicast communication is essential in order to minimize overhead for group communication. For group conferencing services, controlling the end-to-end delay is important to provide high quality communication. In this paper, we present QAMNet, an approach to improve the Quality of Service (QoS) for multicast communication in MANETs. We extend existing approaches of mesh based multicasting by introducing traffic prioritization, distributed resource probing and admission control mechanisms, adaptive rate control of non-real-time traffic based on Medium Access Control (MAC) layer feedback so as to maintain low delay and required throughput for real-time multicast flows. Simulation results show that our approach is as scalable as mesh based deliver structures and does not require significantly more states than normal mesh based multicasting protocols. As we reuse signaling packets of mesh based multicast packets, we do not introduce additional signaling overhead.

Keywords

Ad-hoc Networking, Pervasive Computing, Quality of Service, Multicast Communication

Working Group 2

QAMNet: Providing Quality of Service to Ad-hoc Multicast Enabled Networks

Harald Tebbe and Andreas J. Kassler, *Member, IEEE*

Abstract— **Mobile Ad-hoc networking has been considered as one of the most important and essential technologies that support future Pervasive Computing Scenarios and 4G networks. In a mobile Ad-hoc Network (MANET), support for multicast communication is essential in order to minimize overhead for group communication. For group conferencing services, controlling the end-to-end delay is important to provide high quality communication. In this paper, we present QAMNet, an approach to improve the Quality of Service (QoS) for multicast communication in MANETs. We extend existing approaches of mesh based multicasting by introducing traffic prioritization, distributed resource probing and admission control mechanisms, adaptive rate control of non-real-time traffic based on Medium Access Control (MAC) layer feedback so as to maintain low delay and required throughput for real-time multicast flows. Simulation results show that our approach is as scalable as mesh based deliver structures and does not require significantly more states than normal mesh based multicasting protocols. As we reuse signaling packets of mesh based multicast packets, we do not introduce additional signaling overhead.**

Index Terms—**Mobile Communications, Wireless Networks, Ad-hoc Networking, Pervasive Computing, Quality of Service, Multicast Communication**

I. Introduction

MOBILE Ad-hoc networking has been considered as one of the most important and essential technologies that support future Pervasive Computing scenarios [1] and MANETs will be an integral part of 4G networks like the one proposed by the EU IST-DAIDALOS project (see e.g. Daidalos project homepage at http://www.ist-daidalos.org). A MANET is a collection of mobile nodes (MN) that communicate using wireless links without support from any pre-existing infrastructure network. Packets are delivered from a source to a destination using packet forwarding capabilities of intermediate nodes. Therefore, MNs act as both end systems and routers. In such an autonomous system, MNs are free to move randomly which causes several problems to state-of-the art routing protocols for both unicast and multicast delivery.

The increasing popularity of multimedia and group communication leads to the need to support advanced services like QoS or multicast communication. In multicast communication a source is sending only one packet with a group address as a destination. The network will be in charge of replicating that packet only when necessary to make it reach all the destinations, i.e. all the nodes that have joined the group associated with that specific group address. This leads to minimal bandwidth consumption and high scalability, which is essential in MANETs. Quality of Service is mainly related to resource allocation and management, where it is necessary to decide how to allocate resources such that QoS requirements of all flows can be satisfied.

However in MANETs, several problems arise. Node mobility and channel variation lead to dynamic per node bandwidth availability. The properties of the standard 802.11 MAC and the shared medium make it very difficult to provide QoS, even for wireless Local Area Networks (LANs). Dynamic changes in topology, caused by node mobility and energy constraints lead to frequent re-routing. It is hard to estimate available resources in a multi-hop wireless environment due to node mobility and channel contention. Also, it is hard to perform resource reservation, as the shared medium reservations require global coordination. Violations still can occur as bandwidth fluctuates.

This paper is structured as follows. In chapter 2 we provide an overview on related work in the area of Quality of Service management for Mobile Ad-hoc networks and multicasting. In Chapter 3, we introduce our approach QAMNet for providing QoS for multicast in MANET. With QAMNet, no additional messages are needed compared to mesh based multicast routing protocols. Instead, intermediate nodes measure resource availability and update additional fields within those messages that are used to create and maintain the mesh. Therefore, our approach can be seen as a resource aware mechanism to build QoS enabled delivery meshes without the need for

Harald Tebbe is with University of Ulm, Department of Media Technology, Oberer Eselsberg, 89069 Ulm, Germany (e-mail: harald.tebbe@informatik.uni-ulm.de).

Andreas J. Kassler is with Karlstad University, Department of Information Technology, Universitetgatan 2, 65188 Karlstad, Sweden (e-mail: kassler@ieee.org).

This paper originally appears in: *Proc. 1st International Symposium on Wireless Pervasive Computing 2006 (ISWPC)*, 16 - 18 January 2006, Phuket, Thailand.

additional control messages. We distinguish between best-effort and real-time packets which receive preferential treatment. To manage resource availability, QAMNet features distributed resource probing and admission control interweaved with multicast mesh creation and maintenance. Adaptive rate control of non-real-time traffic is based on Medium Access Control (MAC) layer feedback and helps to maintain low delay and required throughput for real-time multicast flows. To cope with mobility and transient bandwidth fluctuations, QAMNet resorts to regulation mechanism. Chapter four evaluates our approach using simulations and finally chapter five presents our conclusion.

II. Related Work

A. QoS support in MANET

Currently, there exist several QoS architecture proposals for MANETs. These architectures are mainly based on previous work carried out in terms of QoS support in infrastructure networks. The main ideas are to introduce as few states as necessary because topology changes due to mobility make it hard to maintain reservation states. The goal of those architectures is to provide some sort of service differentiation and soft QoS rather than hard guarantees as this is almost impossible given the dynamics of a MANET. In such an environment, it is a bad idea to pin resource reservations to a specific route because such reservations must be re-established whenever route changes occur. INSIGNIA [2] is a QoS model that comprises an in-band signaling protocol with support for adaptive reservation-based services in ad-hoc networks. In-band information concerning the required resources for the flow is used by intermediate nodes to establish, maintain and restore per-flow soft-state reservations. Admission control is performed hop by hop and the destination host informs the source node of the result of the reservation using a QoS report mechanism. Reservations are maintained as long as packets associated with a particular flow are periodically received at intermediate nodes to refresh timers. INSIGNIA copes with re-routing by triggering the admission control and resource reservation through the new path while reservation state along the old path times out. SWAN [3] is a QoS model that tries to keep the network underutilized. Local rate control is used for best-effort traffic based on MAC delay measurements to set aside resources for real-time data. Distributed admission control is performed for real-time flows by the source, based on the result of an end-to-end request/response probe that senses the available bandwidth through the path from the source to the destination. SWAN resorts to dynamic regulation of real-time sessions when congestion/overload conditions occur (e.g due to node mobility) by sending regulate messages to the source, which might lead to a new probing request being sent to the destination. FQMM [4] was developed for assuring a certain level of service differentiation in ad-hoc networks. It is a hybrid Integrated Services (IntServ)/Differentiated Services (DiffServ) model and supports hybrid per-flow/per-class provisioning. Source nodes are considered as ingress nodes, which use classification, marking, policing and shaping. Intermediate nodes only apply shaping. Traffic conditioning is carried out according to a traffic profile. A relative adaptive traffic profile is defined as the relative percentage of the effective link capacity, in order to keep the differentiation between classes predictable and consistent under the dynamics of the network. It is used to keep consistent differentiation between sessions (flows/aggregates).

B. Multicasting and QoS in MANET

There are several Ad-hoc multicast routing protocols that do not provide support for QoS. While MAODVs [5] delivery structure is tree based, ODMRP [6] uses a delivery mesh, which tries to reduce problems of tree-based approaches in the presence of mobility because tree-structures are fragile and need to be frequently readjusted when connectivity changes. Using delivery meshes that span all multicast group members, multiple links do exist which provides redundancy to route breaks caused by mobility of nodes. This minimizes packet loss and avoids frequent tree re-organization but leads to higher overhead due to redundant transmissions.

Only very few works support QoS for multicast like [7] but those approaches introduce network state and additional signaling. The main challenge for providing QoS in Multicast MANETs is the heterogeneous nature of the branches of a multicast tree or mesh as perceived quality varies among users [8] and depends on the resource availability and mobility of a receiver. A tight integration/interaction of resource management and distributed admission control mechanisms developed for unicast QoS architectures with the multicast protocol seems to be beneficial in order to avoid overhead and minimize network state. Additional signaling packets for reservation protocol should be avoided as this contributes to network congestion, especially in high mobility scenarios.

III. QAMNet – Providing Quality of Service for Multicast in MANETs

A. Overview on QAMNET

In this paper, we present QAMNet, an approach to improve QoS for multicast communication in MANETs using a standard 802.11 MAC without QoS support. We extend existing approaches of mesh based multicasting [6] and unicast QoS provisioning [3] by introducing service differentiation (real-time and best-effort traffic class), distributed resource probing and admission control mechanisms, adaptive rate control of non-real-time traffic based on MAC layer feedback so as to maintain low delay and required throughput for real-time multicast flows. Our approach is very scalable as we will show in the evaluation section and does not require more states or signaling than normal mesh based multicasting protocols. This is a big advantage for small devices with memory and power constraints compared to related work, e.g. [7].

B. Resource aware mesh creation in QAMNET

When a QAMNet node in a MANET has real-time traffic to transmit to a multicast group, it starts with flooding the entire network with the first data packet piggybacking the control/signaling information. In contrast of ODMRP, where the control/signaling information is composed of a Join-Query message, we additionally piggy back it with a probing request, which contains a bottleneck bandwidth (BB) and a required bandwidth (RB) field. We refer to the first data packet as the Join-Probe packet. The source broadcasts a Join-Probe packet and upon reception of the first, non duplicate, Join-Probe packet, intermediate nodes set pointers towards their upstream nodes and rebroadcast it, after modifying the probing request information. To do so, each intermediate node additionally updates the bottleneck bandwidth field, if the local bandwidth availability at the given node is less than the current value. Bandwidth availability at the local node is calculated similar to SWAN X[3]X.

Once a Join-Probe packet reaches a multicast receiver, BB indicates the bottleneck bandwidth found along the path. The receiver waits a small time period to collect all Join-Probe packets received from other branches of the multicast mesh. The receiver evaluates if the BB with the largest value is greater than RB and if so creates a Join-Reply, piggy backing a Probe-Response which contains the largest BB and the RB field. It also sets a real-time forwarder flag (RTF_FLAG) for the given multicast group. Note, that at this point the node already has set a state in order to be a multicast forwarder for the group so we just add one more flag. The Join-Reply is relayed by the intermediate nodes all the way to the source following the pointers set when the Join-Probes was propagated. By this way the forwarding mesh is constructed in a similar way than ODMRP X[6]X. In addition, every intermediate forwarder node collects the Probe-Response messages and updates BB field with the maximum of all received Probe-Response messages when forwarding to the source. It also sets it RTF_FLAG for the given multicast group if the forwarded BB value is larger than RB.

Note, that the Join-Probe does not introduce additional control packet overhead compared to ODMRP as it is piggy backed on ODMRPs Join-Query message, which is flooded throughout the network. Only a few bits together with the RB and BB fields are required. In a similar way, the Probe-Response is piggy backed on ODMRPs Join-Reply and also does not contribute to additional overhead. In that sense, Join-Probe and Probe-Response are disseminated periodically throughout the network gathering information on the resource availability of individual nodes at the same periodicity as the messages that are responsible for the mesh creation and maintenance (Join-Query and Join-Reply) without creating additional signalling packets. Currently, we flood these messages periodically (e.g. at an interval of 3 sec.) but that could change depending on the dynamicity of the network in terms of topology changes, e.g. using mechanisms similar to Motion Adaptive Refresh X[9]X.

Once the Join-Reply reaches the source, it multicasts (real-time) packets with the help of the (real-time) forwarder nodes through the forwarding mesh. For all packets with real-time constraints, the source sets the Type of Service (ToS) bit in the IP-header and sends it via MAC-layer broadcast. Before an intermediate forwarding node re-broadcasts the packets, the classifier of that node checks, if the RTF_FLAG for the given group is set. If it is set, the packet bypasses the nodes shaping mechanism, remains unregulated and is directly passed to the MAC layer for re-broadcasting. If the RTF_FLAG is not set, the node will set the ToS bit in the header to zero and put the packet into the shaper, if the MF-flag for the given group has been set through the reception of the proper Join-Reply X[6]X.

C. Regulation of Best-Effort traffic in QAMNET

Each QAMNet node in the MANET regulates independently best effort and rejected real-time traffic (BB smaller than RB, RTF_FLAG not set) through the usage of a traffic regulator. In contrast to SWAN [3], which uses the MAC layer feedback based on RTS-CTS-DATA-ACK sequence to regulate the shaping rate of the regulator, we use the MAC layer back-off delay of 802.11b as feedback.

If the back-off increases, we decrease the rate at which BE and rejected real-time traffic enters the MAC-layer at each intermediate node, otherwise we increase it according to an Additive Increase Multiplicative Decrease (AIMD) control algorithm. We introduce AIMD here because we want to achieve fairness and efficiency in allocating resources in a distributed way. We use this mechanism to control the amount of BE traffic a node injects into the MANET based on the load of other nodes that compete for resources locally. This also helps to set aside resources for the real-time packets. Although the backoff mechanism employed by 802.11 has not been shown to be effective and fair in a MANET environment, it gives important information on the congestion state in the environment of a node. As the rate-regulation of the traffic regulator is based on that scheme, this could lead to somewhat sub-optimal performance. We are currently evaluating alternative approaches to improve the fairness of our mechanisms.

D. Coping with mobility and transient bandwidth fluctuation in QAMNet

In order to cope with false admission and bandwidth availability fluctuation that might cause flows to be admitted while there are in fact no resources left, we include dynamic regulation of real-time traffic. Each node monitors utilization of its real-time traffic class when estimating available resources not only during the Probe-Request phase but also periodically. When it detects violation, e.g. more resources are actually consumed than have been set aside for the real-time class, it selects randomly one of its real-time flows and sets congestion experienced (CE) bit in all packets belonging to the given flow. Also, RTF_FLAG for that flow is set to zero which results in the next real-time packet of the given flow to enter the shaper. QAMNet uses soft states to maintain the RTF_FLAG and sources periodically transmit Join-Probe packets. A node with an already activated RTF_Flag copies infinity into the BB-field because the session has already been admitted. Signaling messages are treated with high priority and bypass the local shaping mechanism, too.

This mechanisms also helps to cope with the mobility of nodes. When nodes are mobile and routes break, the mesh structure leads to a higher probability that alternative paths are available. Also, as nodes periodically flood the Probe-Request (as it is also done with ODMRP), the mesh structure will be updated periodically with the mobility of nodes. That results in new nodes participating in the mesh, which also participate in the processing of the Probe-Request packets. That helps in establishing the reservation states according to the RTF_FLAG for the given real-time packets which will then bypass the shaper at new nodes within the modified mesh thus getting highest priority. However, there might be some time where reservation states at those new nodes are not yet established as we use a 3 sec. interval to maintain the mesh structure. During this time, nodes might experience bandwidth fluctuations due to mobility of nodes. QAMNet deals with this situation by using the regulation mechanisms described above.

IV. EVALUATION OF OUR APPROACH

We have implemented QAMNet as part of a modified ODMRP protocol [6]. We have simulated both the original version and the QAMNet extensions based using NS-2 [10]. We compare then the QAMNet extensions both with ODMRP, and its Steiner Tree (ODMRP-ST) variant [11] with reduced forwarding overhead.

A. Basic Simulation Setup

Throughout all our simulations the channel capacity was fixed to 2 Mbit/s using standard 802.11b DCF MAC layer and the communication range was set to 250 m. Each MN waits 0.025 s before sending a Join-Reply to collect other Join-Queries and each MN waits 0.015s before propagating an updated Join-Query. Join-Queries with piggy backed probing requests are sent every three seconds, irrespective of the mobility of nodes. We set the threshold for the MAC backoff-delay to 0.02 s, which influences the AIMD rate control mechanism for the BE-traffic.

We randomly distributed 50 nodes over an area of 1500m x 300m. The nodes move using a random waypoint mobility model with varying pause times at an average speed of 10m/s. At the begin of the simulation, nodes are static for pause time seconds and move then to a random destination inside the simulation area at a speed uniformly distributed between 0 and 20 m/s (mean speed = 10m/s). Once they reach the destination this behavior is repeated until the end of the simulation. We used seven different pause times: 0, 50, 100, 150, 200, 250, and 300 seconds. For each pause time, five different scenarios where simulated. The results were obtained as the mean values over these 5 runs. We used one CBR real time and two non-real time sources each with 15 receivers. Each of the sources generates 330 byte packets at a rate of 45 packets per seconds (118,8kbps), a typical video conferencing data rate.

B. Simulation Results

As can be seen from Figure 1, the average delay of real-time packets (QAMNet RT) can be controlled by QAMNet efficiently and is bounded between 15 and 50 msec only marginally increasing with higher mobility. This is the result of controlling the load of the real-time packets at the MAC layer through the usage of distributed admission control and the regulation of best-effort traffic of the shaper. The drawback of our approach is the additional delay for the best-effort packets (QAMNet-BE) as those packets are regulated by the shaper. As there are much more BE packets than RT packets in our simulation, the average delay over all packets is marginally higher when using QAMNet compared to ODMRP (denoted as ODMRP-BE) as the high traffic leads to high delay due to MAC contention and long queues for BE traffic.

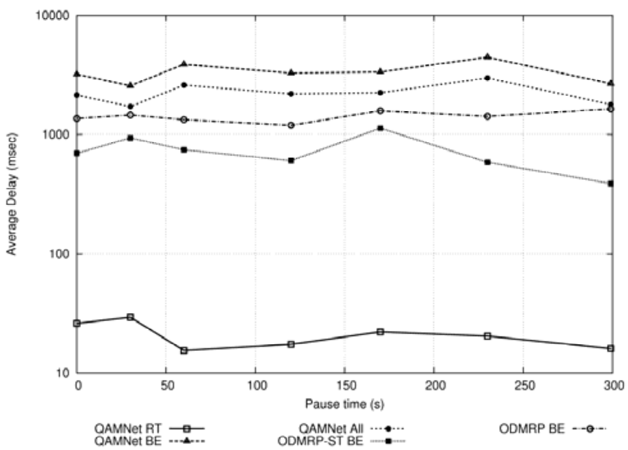

Figure 1. Average Delay in msec as a function of pause time.

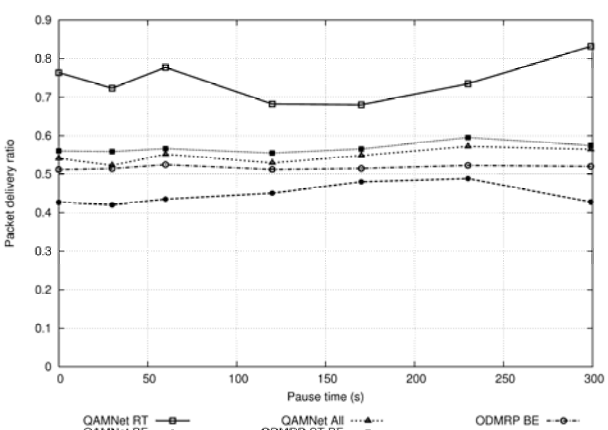

Figure 2. Packet Delivery Ratio as a function of pause time.

The delivery ratio of real-time packets (RT) is significantly higher than for BE packets and stays between 68 and 84%. In contrast, the delivery ratio for the BE packets is lower than when using ODMRP or the Steiner tree variant. This is because BE packets are dropped once the shaper queue is full and cannot accept more BE packets. But as can be seen from Figure 2, the average delivery ratio over all packets is higher when using QAMNet then using standard ODMRP. This is also due to the preferential treatment of QAMNet control messages which bypass the shaper and are thus treated as real-time packets.

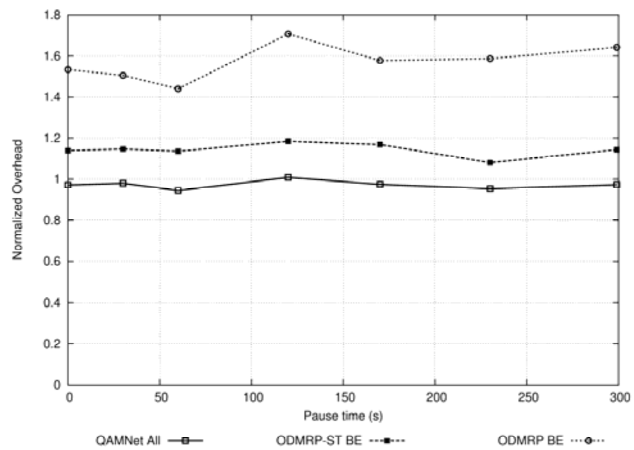

Figure 3. Normalized Overhead as a function of pause time.

This effect is also visible in Figure 3, where we plot the normalized overhead as a function of the pause time. Using QAMNet limits the overhead to stay below one for all scenarios whereas ODMRP shows higher normalized overhead (between 1.4 and 1.7). The overhead of the Steiner Tree variant (which minimizes the number of forwarding nodes in the multicast mesh) is lower than plain ODMRP but still higher than using our approach. The preferential treatment of control packets in QAMNet leads to higher routing stability and thus to lower re-transmission of control packets which reduces the overhead. In addition, QAMNet does not introduce additional signalling packets compared to ODMRP. This shows the scalability of our approach wrt. mobility, where QAMNet can control the overhead effectively and maintain low delay for real-time traffic.

V. Summary and Conclusion

We have developed a mechanism to provide Quality of Service for Multicasting in Mobile Ad-hoc networks. Our approach QAMNet extends existing approaches of mesh based packet forwarding in MANET and adds distributed resource probing and admission control, which is interweaved with the mesh creation process. In that sense, our approach can be categorized as resource aware mesh creation. We also use local traffic regulation of non real-time flows in order to control the delay of real-time packets. We cope with mobility of nodes and transient bandwidth fluctuation by resorting to a regulation process. Our approach is scalable as it does not require significantly more states than standard mesh based multicasting protocols. We do not introduce additional control packets compared to mesh based multicast routing protocols as we re-use signaling packets designed originally for mesh creation and maintenance and piggy back resource requirements and probing information. Our evaluation showed that when using QAMNet the delay and packet loss rate of multicast real-time packets in mobile Ad-hoc networks can be significantly reduced and limited to at most 50 msec over the whole range of mobility of the nodes. In addition, the overall overhead is reduced by 60% at the expense of increased delay and packet loss for best effort traffic. This effect is more visible at higher load within the MANET.

It seems natural to extend our work and combine it with the standard SWAN approach so that we can support QoS for both unicast and multicast communication. That would require a generic interface between the MAC and routing layer to provide feedback on MAC layer utilisation to the routing protocol and to adapt the shaping rate. SWAN works with any MANET unicast routing protocol as the probing messages are separated from the routing protocol. However, in our approach, we integrated the probing with the multicast routing in order to reduce the overhead introduced by probing messages for the multicast scenario. In our future work we will extend the simulation by including other mobility and traffic models. We will also make the probing requests adaptive to change with mobility and traffic patterns. In addition, delay jitter is also an important parameter to assess Quality of Service management mechanisms for e.g. audio streams. We will consider that parameter in future simulation runs. Finally, we plan to implement QAMNet on small PDAs and test it in a real setup. We will also extend our approach to cover hybrid MANETs, i.e. mobile Ad-hoc networks that are connected with the public internet using one or more gateways. Finally, we plan to extend MAODV in a similar way so that we can also support QoS for tree-based multicast routing protocols in a MANET.

References

[1] Jun-Zhao Sun, *Mobile Ad-hoc Networking: An Essential Technology for Pervasive Computing*. In: Proceedings of International Conferences on Infotech & Infonet, Beijing, China, C:316 – 321

[2] Seoung-Bum Lee, Gahng-Seop Ahn, Xiaowei Zhang, and Andrew T. Capbell, *INSIGNIA: An IP-Based Quality of Service Framework for Mobile Ad-hoc Networks*. In: Journal Parallel and Distributed Computing, vol. 60 nº4, Apr. 2000, pp. 374-406.

[3] Gahng-Seop Ahn, et al, *Supporting Service Differentiation for Real-Time and Best-Effort Traffic in Stateless Wireless Ad-hoc Networks (SWAN)*. In: IEEE Transactions on Mobile Computing, vol. 1, no. 3, 2002.

[4] H.Xiao, K.Chua, W.Seah and A.Lo, *A Flexible Quality of Service Model for Mobile Ad-hoc Networks*. In: Proceedings of Vehicular Technology Conference (VTC), Tokyo, Japan, May 2000, pp. 445-449.

[5] E.M. Royer and C.E. Perkins, *Multicast Ad-hoc On-Demand Distance Vector (MAODV) Routing*. Work in Progress, Internet-Draft, draft-ietf-manet-maodv-00.txt, July 2000.

[6] Yunjung Yi, Sung-Ju Lee, William Su, and Mario Gerla, *On-Demand Multicast Routing Protocol (ODMRP) for Ad-hoc Networks*. Work in Progress, Internet-Draft, draft-yi-manet-odmrp-00.txt, March 2003

[7] K. Bür, C. Ersoy, *Multicast Routing for Ad-hoc Networks with a Multiclass Scheme for Quality of Service*. LNCS, ISCIS 2004, Antalya, October 2004.

[8] B. Wang and J. C. Hou, *Multicast Routing and Its QoS Extension: Problems, Algorithms, and Protocol*. In: IEEE Network, Vol 14, 01/2001.

[9] Soon Y. Oh, Joon-Sang Park, Mario Gerla: *E-ODMRP: Enhanced ODMRP with Motion Adaptive Refresh*, in Proceedings of ISWCS (International Symposium on Wireless Communication Systems) 2005, September 5-7, 2005. Siena, Italy.

[10] K. Fall, K. Varadhan, *ns, Notes and Documentation.* The VINT Project, UC Berkeley, LBL, USC/ISI, and Xerox PARC, November 2003

[11] P. Ruiz and A. Gomez-Skarmeta, *Mobility-aware mesh construction algorithm for low data overhead in multicast Ad-hoc routing.* In: Journal of Communications and Networks (JNC), vol. 6, no. 5, Dec. 2004.