Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Real-Time Mesh Networks for Industrial Applications

André Herms*, Edgar Nett*
Stefan Schemmer**

**University of Magdeburg, Department of Distributed Systems, Magdeburg, Germany*
(e-mail: {aherms, nett}@ivs.cs.uni-magdeburg.de) .
***rt-solutions.de GmbH, Cologne, Germany*
(e-mail:schemmer@rt-solutions.de)

Abstract: Wireless LANs (IEEE 802.11) are increasingly used in industrial applications. They reduce cabling costs, increase flexibility and enable mobile applications for maintenance or logistics tasks. Mesh networks provide a self-configuring and -healing wireless backbone for large scale deployments (e.g. in process automation). This paper presents a routing algorithm which provides QoS in wireless mesh networks, thus leveraging their use in industrial applications. It allows reserving bandwidth for real-time flows based on measurements of the physically available bandwidth. Thus it fully utilizes the bandwidth while still preventing congestion. Simulation results demonstrate the reliability of the algorithm and its advantage over previous works.

1. INTRODUCTION

Wireless networks based on the IEEE standard 802.11 (WLAN, WiFi) are increasingly used in industrial applications. They provide reduced cabling costs, e.g. for the connection of remote parts of a production plant in the process industry, increased flexibility, and the integration of mobile devices, e.g. for mobile control panels, hand scanners, and automated guided vehicles (AGVs).

Wireless infrastructure usually consists of Access Points (APs) that provide wireless access to the mobile clients and are connected over a wired backbone of copper or fibre optics. However, this solution has limitations, especially for large scale installations that are typical for process automation. Deployment costs for wired infrastructure and its inflexibility to topology changes are the main drawbacks. These can be overcome by using wireless links between the APs as well. A currently available solution to set up such a wireless backbone is the so-called Wireless Distribution Service (WDS). WDS uses static routes which have to be manually defined on every AP. Mesh networks which are already part of some commercial products provide more flexibility, better usability, and a higher availability. In a mesh network, links between APs are automatically discovered and routes are dynamically calculated, based on the current topology. This way, a backbone network is created that in contrast to the static WDS automatically reacts to connectivity changes and compensates link failures whenever possible. IEEE 802.11s, a standard which currently is in development, will provide a common specification that enables interoperability between different products. The continuous monitoring and adaptation makes the wireless backbone self-configuring, self-healing, and self-optimizing.

WLAN mesh networks can be used in industrial applications like process automation for wireless propagation of sensory data. The sensors themselves are connected to the wireless backbone by an energy efficient technology like Wireless HART. Dedicated gateway nodes forward the sensory data into the WLAN backbone, which provides significantly more bandwidth than the wireless links of the access network. The WLAN backbone can additionally be used for other real-time applications like connecting the cameras of a video surveillance system or cordless phones using Voice over IP. With a sufficient density of APs, the mesh network provides several redundant routes that result in a high availability of the communication service.

The applications mentioned above are real-time dependant. In contrast to this, existing mesh network solutions only provide a best-effort communication without any guarantees regarding bandwidth or end-to-end latency. Especially in highly loaded networks, this can lead to significant problems. When the network becomes saturated—which means that more data is generated than the network is able to deliver—the resulting congestion leads to long latencies and high packet loss rates. Using congestion control techniques that adapt the sender's rate like TCP does (Nagel, 1984) is not an option, because the data rate is fixed and defined by the application. As the control packets are subject to the same problems, the routing process will be affected as well. As a consequence data packets may not be delivered at all.

In this paper, we present a routing architecture for wireless mesh networks, which prevents congestion using bandwidth reservation. To achieve a high utilisation of the available bandwidth, the architecture includes a calibration process. It measures the mutual interference between the APs and thereby determines how much bandwidth is actually available in the network. Furthermore, this paper presents an approach how the results of the initial calibration can be continuously updated via passive measurements and network simulation.

978-1-1234-7890-2/08/$20.00 © 2008 IFAC

10.3182/20080706-5-KR-1001.4009

The paper is structured as follows: Section 2 presents related work. Section 3 describes the general routing architecture (3.1), the challenges in determining the available bandwidth (3.2), the calibration process, which meets these challenges (3.3), and the model behind the calibration process (3.4). In Section 4, results from a simulation-based evaluation are presented, while Section 5 proposes future work and concludes the paper.

2. RELATED WORK

Several wireless communication technologies exist which provide reservation of resources. One popular example is the Bluetooth standard with its SCO links (Bluetooth, 2007) that are mainly intended for audio communication. However, multi-hop communication with Bluetooth (Scatternets) is still work in progress and the technology only provides low data rates.

A better integration of QoS for multi-hop communication is achieved by special sensor network protocols like ZigBee, Wireless HART or ISA SP100. They are designed for systems with autonomous power supply. Wireless communication is limited to the minimal requirement of delivering sensory data. ZigBee was developed to work on cheap hardware with low power consumption. Wireless HART is based on the same IEEE 802.15.4 standard as Zigbee (IEEE, 2005), which only provides low data rates. ISA SP100 is currently under development. It is a common standard that will allow the use of different real-time protocols (like HART). All these technologies provide the possibility of guaranteeing QoS, but allow only limited data rates that are not sufficient for all applications.

The integration of resource reservations in mesh routing protocols based on the WLAN standard was considered by various authors (Mahrenholz, 2007; Xue, 2003; Chen 2005; Kuo 2005). In contrast to our approach, the reservation is decentralised. This avoids the problem of having a single point of failure, which is especially important as they are developed for Mobile Ad-hoc Networks (MANETs), which lack any fixed infrastructure. The common underlying idea is that reservation requests are propagated along a route. The required resources are reserved locally, on each station. The routing itself is provided by slightly modified standard mesh routing protocols like AODV or DSR. The decentralised coordination requires a high degree of consistency between the distributed resource views of all members. In (Herms, 2007b) we have shown that it is not sufficiently solved by these protocols and suggested an extension that ensures consistency. However, those protocols require a large number of control messages. In this work we avoid the problem by using a centralised coordinator that has a consistent view of the available resources. Static redundancy could be used to make this single point of failure sufficiently reliable. This assumption is plausible as the mesh networks are connected to a fixed wired infrastructure in the application context considered.

3. MEETING REAL-TIME REQUIREMENTS IN MESH NETWORKS

3.1 Routing Architecture

Usually, the routing layer in a wireless mesh network performs just one single task: Finding the next neighbour the packet is forwarded to, according to the topology information, so that it will reach its destination. To meet the real-time requirements, the routing layer has to provide additional functionality. It must find a route from the source to the destination that meets the bandwidth requirements of the flow. Furthermore, all established routes and corresponding reservations must be stored so they can be considered when a new route is calculated.

In our architecture, a special entity, the QoS manager, which runs on a dedicated station in the mesh network, provides the QoS functionalities (Herms, 2007). The other, ordinary stations implement only a small set of functionalities, which allows using smaller and cheaper hardware. These basic functionalities comprise a multi-hop communication service, which is implemented by a simple link state routing (similar to OLSR). This service is used to transport control messages between the ordinary stations of the mesh network and the QoS manager. Additionally, it can be used to transport best effort traffic between arbitrary stations. As part of its normal operation, the link state routing that implements this basic service provides global information about the link states and hence a global view on the topology of the network. Since the QoS manager is part of the network, it can access this topology information through a cross-layer interface and can use it as input for its routing decisions.

All QoS flows are managed by the QoS manager. In order to establish a new QoS flow, a station sends a request to the QoS manager that contains the destination and flow specifications with maximum PDU size and transmit period. Based on a global view of the whole network, the QoS manager tries to find a feasible route that fulfils the flow specifications. If it cannot find such a route, the QoS manager rejects the request by sending a corresponding control packet to the requesting station. Otherwise, the QoS manager accepts the request. It sends control packets to all stations along the route to provide them with a flow identifier of the newly granted flow and a corresponding forwarding rule. These forwarding rules establish the route. After that, the QoS manager informs the requesting station that the flow was granted and that the route has been set up.

The global view of the QoS manager comprises three types of information: first, the topology information provided by the basic link state routing; second, information about all flows that have been granted and their corresponding resource reservations; and third, a *conflict model* that describes the mutual interference of the stations (the following subsections describe how this model is determined). The *path planner* uses this information for the local route planning.

3.2 Bandwidth in Wireless Mesh Networks

Meeting real-time requirements is achieved by reserving resources. However, the medium in wireless mesh networks is a resource with non-trivial properties. All devices share the bandwidth of a common channel with their neighbours. Therefore, the Wireless LAN standard (IEEE 802.11, 1997) defines the medium access protocol CSMA/CA. Without going into details the principle of operation is the following: Before a station is allowed to transmit, it checks if another station is already transmitting (carrier sense). In this case, it defers its own transmission. Further mechanisms are applied to reduce the probability of packet collisions, but these are not relevant for the following considerations.

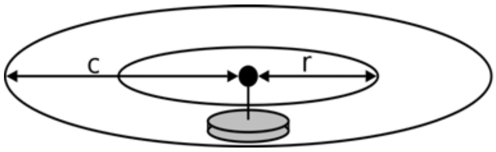

Figure 1: Carrier-Sense-Range

The area around a transmitting station can be divided into three significant parts (see fig. 1). In the receive/transmit range (distance < r) packets can be decoded and are delivered to the higher layers. A station in this area will receive all packets of the transmitter, as long as no collision occurs. The carrier-sense range (distance < c) is the area where the carrier signal is still received, but bit errors prevent the successful decoding of the packets. Determining this area is crucial for bandwidth planning, as stations in this area cannot transmit concurrently so that they share the bandwidth of the channel. Only stations outside the carrier sense range of a transmitting station do not sense the carrier and therefore are able to transmit concurrently. Due to the analogue nature of radio signals, the boundaries of these ranges are not clearly separated. Instead, a gradual transition between them can be observed.

For the purpose of resource management it is essential to know which stations are in carrier-sense range of each other. Thus, all approaches for giving bandwidth guarantees must include a method to detect interfering neighbours. The most simple assume that bandwidth is shared only with stations within the receive/transmit range, like (Xue and Ganz, 2003) or (Kuo, 2005). Others are based on the observation that the carrier-sense range is theoretically twice as large as the receive/transmit range and therefore include all two-hop neighbours, like (Chen and Heinzelman, 2006) or (Mahrenholz, 2006). However, these are only approximations and topologies often do not meet these assumptions. For such constellations, the admission control would still generate admission failures that cause network overload. Without further information one can only rely on the worst-case assumption that all stations are within carrier-sense range of each other. This very conservative assumption has the disadvantage of underestimating the available bandwidth, which results in a poor bandwidth utilisation in large networks.

3.3 Measuring the Mutual Interference

As explained above, information about the mutual interference among stations in the network is required. Since stations that are within the carrier sense range but not within the receive/transmit range sense the carrier but do not receive the messages. The source of a blocking carrier signal cannot trivially be identified. Instead, the calibration uses an indirect measurement approach based on the effect caused by sharing the bandwidth. During the calibration phase, a pairwise determination of the mutual interference is performed. The QoS manager coordinates this process, which consist of the following steps:

In the first step, every station tries to transmit a large amount of broadcast packets for a certain duration (4 second), which leads to an overflow of its interface queue. The neighbouring stations count the number of packets they receive and thus determine the number of physically transmitted packets. They report the number to the QoS manager, which stores the maximum of all reported values in its local database.

In the second step, the QoS manager selects pairs of stations that start transmitting like in the first step. Depending on the interference between the selected stations, a reduced number of transmitted packets per station will be observed. Stations that are not within the carrier sense range of one another do not share the bandwidth, so that each of them is expected to transmit about the same number of packets as during the first step. When stations are within the carrier sense range of each other, they share the bandwidth so that each station transmits only about half as many packets as during the first step. Our measurements have shown that real, heterogeneous networks do not always exhibit this idealized behaviour (Herms et al., 2007). In particular, asymmetric links and communication relations with ambiguous properties have been found.

The results of the calibration phase are stored in the local database of the QoS manager and are used for reservation decisions in the following phases.

The calibration phase takes a significant amount of time, as for n stations up to $0.5n(n\text{-}1)$ pairwise measurements are required, each about 4 s. Furthermore, the calibration is very sensitive to background traffic, which has to be stopped before the calibration phase. Thus, it cannot run simultaneously with the normal operation of the network. Instead, it is only started when the network has changed. Before such a re-calibration delivers the actual values, the conservative assumption of all stations interfering with each other is used.

3.4 Deriving the Conflict Model from Measurements

A typical way of modelling the interference in a mesh network is a conflict graph. Stations are described as nodes and edges indicate that the adjacent stations share bandwidth. Here, this means they are within the carrier sense range of one another. For modelling asymmetries and the gradual transitions between interfering and not interfering, the graph is directed and has edge weights between 0 and 1. A weight of 1 means, both stations share their bandwidth completely. A 0 indicates that the bandwidth not shared. Intermediate values express a gradual sharing of the wireless medium. The conflict graph is represented by a weighted adjacency matrix $C = c_{(n,n)}, c_{i,j} \in [0,1]$. The bandwidth consumption of the nodes is represented by a vector $\beta \in [0,1]^n$. This representation allows verifying if the bandwidth limit is not exceeded by checking the constraint $C\beta \leq 1$, as suggested in (Gupta, 2005). When this is not true for the potential adding of a flow, the corresponding reservation is rejected. The entries in the matrix C are derived from the measurements in the calibration phase.

To this end, we utilize the fact that the constraint must be fulfilled for the subgraph consisting only of the two nodes i and j. Thus, for each pair the following constraint $\tilde{C}\tilde{\beta} = \begin{bmatrix} 1 & c_{i,j} \\ c_{j,i} & 1 \end{bmatrix} \begin{pmatrix} \tilde{\beta}_i \\ \tilde{\beta}_j \end{pmatrix} \leq \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ must be true. During the measurement process the used bandwidth of the two stations $\tilde{\beta}_i$ and $\tilde{\beta}_j$ is maximal. It can therefore be assumed that the constraint is satisfied in its bounds, i.e. it is fulfilled with equality. In this case, the values of $c_{i,j}$ and $c_{j,i}$ can be derived from the observed consumed bandwidths during one measurement step: $c_{i,j} = \frac{1-\tilde{\beta}_j}{\tilde{\beta}_i}$. This way, each pair of probed nodes allows determining two entries of the conflict matrix. As long as no calibration has been performed for a pair, a value of 1 is assumed. This is a pessimistic yet safe assumption, which means these two stations share their bandwidth.

4. SIMULATION

For our architecture we have developed a wireless mesh routing implementation, which as available as Open Source project AWDS (AWDS, 2007). It runs in Linux-based environments and can be deployed on commercial hardware that uses embedded Linux, e.g. OpenWRT (OpenWRT, 2008). Currently, this routing software only provides best effort communication services, which are used as basis of our QoS extensions. An abstraction layer between routing and runtime environment (Herms, 2005) allows running the protocols on Linux-based devices and on the NS-2 network simulator without modifying the source code (NS-2, 2008). The remainder of this section presents results from a simulation-based evaluation.

4.1 Necessity of Reservation

In this section, we give an example of the necessity of reservation. The scenario is a network with 20 stations randomly distributed over an area of 100x400 m² with a uniform distribution. The underlying shadowing propagation model was configured according to a typical outdoor setting, which results in a receive/transmission range of about 90 m. The routes have therefore a low number of hops, so that a calibration is not required.

In this scenario, every minute a randomly selected station tries to establish a new data stream. The data streams are broadcast addressed, meaning that they are received by all neighbouring stations. As broadcast packets are always transmitted with the fixed basic data rate, this helps preventing effects caused by the automatic rate adaptation of the radio interface. A data rate of 2 Mbit/s was chosen according to the IEEE 802.11b standard.

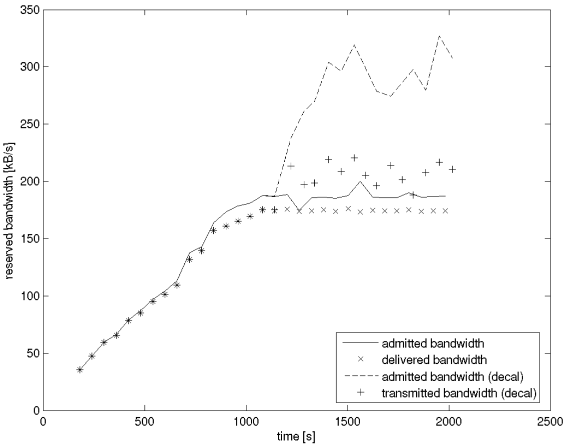

Figure 2: network behaviour under load

The order of reservations is fixed. Each request defines the source station, the packet size and the period of transmissions. In the scenario at hand, the packet size was 500 bytes (payload) and the period was uniformly random distributed between 1 ms and 100 ms. During simulation, the first 17 reservation requests were granted, all following were rejected because of a lack of bandwidth. To measure the delivered bandwidth, monitoring agents that report the amount of data received were installed on every station in the network.

The results are depicted in figure 2. The solid line represents the admitted bandwidth with reservation. It can be seen that the resource manager limits the admitted bandwidth. The delivered bandwidth nearly equals the admitted bandwidth, except for the typical packet loss. The dashed line represents the admitted bandwidth without a correct reservation. During the simulation, the reservation was enabled, but de-calibrated, so that 200% of the available bandwidth could be reserved. As this is the only difference between both simulations, the graphs are identical in the first part. From the point of time, where the reserved bandwidth converges to 100%, they start to diverge. With the de-calibrated reservation, an increasing discrepancy between reserved and delivered bandwidth can be observed. This indicates a considerable packet loss. Furthermore, both reserved and admitted bandwidth exhibit an unstable behaviour. This is caused by the loss of too many control packets, which results in the exclusion of stations from the network – they seem to be inactive. The consequence is that the bandwidth allocated to those stations is deemed to be available and is assigned to other requests. Of course, this unintended behaviour further sharpens the overload problem.

This example scenario underlines the necessity of a suitable mechanism to prevent congestion in mesh networks. Otherwise, the network can drift into a state of serious instability. When using real-time data streams, a reservation is unavoidable, as congestions will result in long delays and congestions control mechanisms like those of TCP cannot be used for real-time flows.

4.2 Network Calibration

For networks with a large number of hops, the calibration described above improves bandwidth utilization. In this section, we give a simple example to illustrate this. The example consists of a regular topology (see figure 4) with 20 stations and a fixed distance (90 m) between adjacent stations.

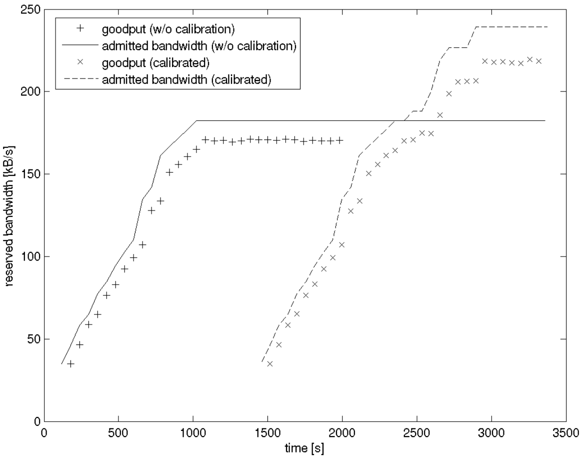

Figure 3: Utilization with and without calibration

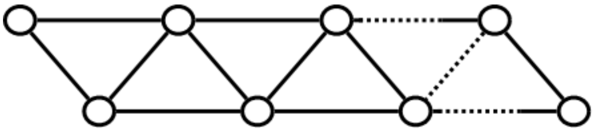

Figure 4: Example topology

Figure 3 depicts a comparison between admitted bandwidth and goodput in the example topology with and without calibration. The reservations in the calibrated network start after 1500 s, that is after the calibration process has been completed. The figure shows that admitted bandwidth and goodput of the calibrated network are above the respective values of the non-calibrated network. This is possible due to the concurrent use of the wireless medium.

5. CONCLUSION & FUTURE WORK

In this paper, we have shown that congestion is a serious problem in wireless mesh networks, especially when used in the automation context with its fixed-data-rate applications. The reservation mechanism presented in this paper tackles this problem. Furthermore, the calibration phase allows avoiding overly pessimistic worst-case assumptions and hence increases utilization of the available bandwidth.

The calibration phase determines interference among stations or the potential for concurrent transmissions respectively. Currently, it occurs once during network initialization and additionally whenever the network changes significantly. In both cases, it is necessary to pause running applications during the calibration in order not to influence the measurements. However, some applications (e.g. chemical plants in the process-industry) require uninterrupted network connectivity for long periods (years). During this time, changes in the network and especially in the environment are likely to occur (e.g. machines, installations, production materials, and people). These changes can possibly have the following negative effects. If two nodes were initially out of carrier sense range and are now within carrier sense range, this can lead to overload in the network if not detected. Additionally, if two nodes were initially within and are now out of carrier sense range, this can lead to poor medium utilization if not detected. Therefore, in order to avoid these negative effects in dynamic environments, additional measures are required to determine mutual interference between the nodes concurrent with the normal operation.

Please remember that direct measurement of the carrier sense range is not trivially possible because stations will only deliver packets from stations within the receive/transmit range and not from those between receive/transmit and carrier sense range. For that reasons, an indirect, active measurement approach has been proposed above. However, active measurements should be avoided during normal operations so as not to hinder ongoing tasks. One possible approach to determine the carrier sense range with passive measurements is using corrupted packets. Corrupted packets are incorrectly decoded frames, which are normally discarded by the driver. However, it is possible to instruct the driver to deliver such packets to the upper layers. If it is still possible to determine the sender of such packets, the information can be used to improve current knowledge about the carrier sense ranges.

We propose the following approach, which combines passive received signal strength (RSS) measurements and simulation to continuously estimate the carrier sense range of the stations and calibrate the resource manager.

In ongoing research, we are developing methods to calibrate a wireless propagation model from measurements in the existing network (Ivanov et al., 2007). The current application of these methods is a coverage planning that maintains the availability of the WLAN under changes of the environment. However, their capability to predict the RSS of the stations in all areas of the environment can also be used to estimate which stations are out of the receive/transmit range but within carrier sense range because both ranges can be characterized by a corresponding RSS value. The model that is used for these predictions is continuously adjusted to the reality by inter-node RSS measurements. In this way, changes in the infrastructure or the environment are automatically reflected in the model. These measurements are passive and therefore run online without conflicting with applications. Both alternative methods might have inaccuracies: the information in the corrupted packets might not be decoded and the simulation method might have prediction errors. Even though, they will bring additional information that will help to preserve QoS guarantees or increase network utilization in online mode.

REFERENCES

AWDS (2007), Project Homepage, online http://awds.berlios.de

Braden, R.,Ed. et al. (1997), Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification, http://tools.ietf.org/html/rfc2205

Cisco Systems (2007), Cisco Documentation, „VoIP Call Admission Control“, http://www.cisco.com/univercd/cc/td/doc/cisintwk/intsolns/voipsol/cac.htm

Chen, L. & Heinzelman, W. (2005) QoS-aware Routing Based on Bandwidth Estimation for Mobile Ad Hoc Networks IEEE Journal on Selected Areas in Communications, 23, 561 – 572

Gupta, R. (2005) Quality of Service in Ad-Hoc Networks Univiersiy of California, PhD Thesis, Berkley, 2005

Herms, A. and Mahrenholz, D. (2005), Unified Development and Deployment of Network Protocols, *in* 'Proceedings of Meshnets '05'.

Herms, A., Lukas, G., and Ivanov, S. (2007), Measurement-Based Detection of Interfering Neighbors for QoS in Wireless Mesh Networks, *in* 'Proceedings of Mobile Summit 2007'.

Herms, A.; Ivanov, S; Lukas, G. (2007b) Precise Admission Control for Bandwidth Reservation in Wireless Mesh Networks; in *Proceedings of The Fourth IEEE International Conference on Mobile Ad-hoc and Sensor Systems,* 2007, Pisa Italy

IEEE 802.11, (1997) ‚IEEE Standard for Information technology - Telecommunication and information exchange between systems - Local and metropolitan area networks - Specific requirements. Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs)’

IEEE, (2005), ‘IEEE 802.15.1 Wireless MAC and PHY Specifications for WPANs’ http://standards.ieee.org/getieee802/download/802.15.1-2005.pdf

Ivanov, S., Nett, E., and Schemmer, S. (2007), Planning Available WLAN in Dynamic Production Environments, 7th IFAC International Conference on Fieldbuses and Networks in Industrial and Embedded Systems, November 7-9, 2007, Toulouse, France, to appear.

Kuo, Y.; Wu, E. H. & Chen, G. (2005) Admission Control for Differentiated Services in Wireless Mesh Networks Meshnets '05

Mahrenholz, D. (2006) Providing QoS for Publish/Subscribe Communication in Dynamic Ad-hoc Networks University of Magdeburg

Nagle, John, (1984) Congestion Control in IP/TCP Internetworks, Ford Aerospace and Communications Corporation, January 1984, http://tools.ietf.org/html/rfc896

NS-2 (2008), project homepage, online http://nsnam.sf.net

OpenWRT (2008), project Homepage, online http://openwrt.org

Xue, Q. & Ganz, A. (2003) Ad hoc QoS on-demand routing (AQOR) in mobile ad hoc networks Journal of Parallel and Distributed Computing, 63, 154-165