Synergy Of Ad-Hoc And Infrastructure Based Networking

Diplomarbeit

der Philosophisch-naturwissenschaftlichen Fakultät

der Universität Bern

vorgelegt von

Isabel Steiner

2006

Leiter der Arbeit:

Prof. Dr. T. Braun

Forschungsgruppe Rechnernetze und Verteilte Systeme (RVS)

Institut für Informatik und angewandte Mathematik

Betreuer der Arbeit:

Marc Danzeisen
Forschungsgruppe Rechnernetze und Verteilte Systeme (RVS)
Institut für Informatik und angewandte Mathematik

Abstract

Heterogeneous networking has evolved as a natural consequence of the various existing technologies and is getting more and more important. Heterogeneous networks combine existing network technologies such as General Packet Radio Service (GPRS), Universal Mobile Telecommunications System (UMTS), Wireless Local Area Network (WLAN), etc. to benefit from the resulting capacity, which is of course much larger compared to scenarios where just one technology is used. Many devices already contain hardware to access more than one network technique and additional network cards support the use of multiple technologies. The handover between the different networks can be done seamlessly, i.e. the sign-off from the previous and the sign-in to the desired network is not visible for the user. The ability to be always online has many advantages such as periodical synchronization of the e-mail inbox, availability of instant messaging, Voice over Internet Protocol (VoIP) programs (e.g. Skype), and permanent access to the Internet.
Additional to the access to infrastructure-based networks like GPRS, UMTS, and WLAN, mobile devices offer the possibility to establish ad-hoc peer-to-peer connections to other devices without using an access network. This can be supportive in many situations, e.g. whenever no network connection is available, accessible or affordable. However, in some cases it would just be too slow and sometimes also expensive to share data over the Internet although two peers are close enough to build up a direct connection. The establishment of such ad-hoc links can be done in an automated way.
An important concern in the context of mobile networking is the energy-awareness of all involved algorithms. Therefore, a new energy-saving mechanism has been developed for heterogeneous mobile networks. It is called "on-demand mode" and prevents a device from consuming energy for network maintenance although the wireless link is not used at the moment, i.e. no data is sent or received by a device. The on-demand mode wakes up a device over the network in case some data has to be transferred and puts it into sleep mode once again when transmission is finished. Additionally to the energy saving qualities, the on-demand feature is able to free network resources.
Ad-hoc peer-to-peer and on-demand enabling software depicted in this work are not on the market already and before starting development it would be interesting to make some academical investigation to quantify the advantage resulting of a seamless heterogeneous environment including handovers between peer-to-peer ad-hoc connections and additional benefits of on-demand networking. The goal of this work is to analyze the overall performance increase due to additional ad-hoc connections and to measure the energy consumption decrease resulting from the on-demand mode.
To analyze the benefits of heterogeneous networking from a user's as well as an operator's perspective, a suitable simulation tool is required in order to measure the different aspects of interest. While the user is interested in higher throughput and longer battery lifetime, the operator's requirements are more available network capacity per user due to lower overall charge of the network and revenue due to new services. A simulation tool suitable for the planned measurements must support various network technologies and mobility patterns. A module to integrate seamless handover technology and the two additional features – ad-hoc and on-demand ability – could then be programmed supplementary. The fact that no such tool is available today led to the decision to build up a custom simulator from

scratch fitting the request for heterogeneity and seamlessness. Thus, the Heterogeneous Network Simulator (HNS) has been designed within this work for the specific needs in heterogeneous environments. For the results, the center of attention is on the application layer behavior, e.g. total duration of data transmissions and battery consumption. Further, the overall connectivity of the nodes will be of interest. The simulator supports different mobility models such as Random Waypoint Mobility, Reference Point Group Mobility, Manhattan Grid Model, Gaussian Mobility, abstractions of current network technologies such as GPRS, UMTS, and WLAN as well as a mechanism to perform handovers between these technologies. Additionally, the ad-hoc mode is modeled to enable data exchange over a single-hop WLAN peer-to-peer link. The on-demand mode enables the sleep functionality to free network resources and to save energy.

The evaluation focuses on estimating the added value of peer-to-peer ad-hoc links in heterogeneous environments and the global benefit resulting from on-demand mode, i.e. the ability to avoid the use of broadband communication when not required in order to save resources and power. The saved power is estimated based on preceding energy consumption measurements on mobile devices. Various test scenarios have been set up to simulate the different scenarios with several seeds. The results show that allowing peer-to-peer ad-hoc connections increase the throughput and decrease the power consumption notably. On-demand mode as additional energy- and resource-saving approach further decreases the power consumption. Together, the simulated features have a significant positive impact on the throughput and energy consumption.

Contents

1 Introduction, Terms and Related Work **1**

2 Wireless Network Technologies and Heterogeneous Networking Protocols **3**

2.1 Access technologies in wireless networks . . . 3
2.2 IEEE 802.11 . . . 3
2.3 GSM – Global System for Mobile Communications . . . 4
2.4 UMTS – Universal Mobile Telecommunications System . . . 6
2.5 UWB - Ultra Wideband . . . 7
2.6 Summary of wireless network technologies . . . 7
2.7 CAHN - Cellular Aided Heterogeneous Networking . . . 8

3 Mobility and Traffic in Wireless Network Simulation **11**

3.1 Mobility models . . . 11
3.1.1 The Random Waypoint Mobility model (RWP) . . . 11
3.1.2 The Gauss-Markov model . . . 13
3.1.3 The Reference Point Group Mobility model (RPGM) . . . 14
3.1.4 The Manhattan Grid model . . . 15
3.2 Network simulation tools . . . 16
3.2.1 NS-2 . . . 16
3.2.2 QualNet . . . 16
3.2.3 Opnet . . . 17

4 Design and Functionality of Heterogeneous Network Simulator **18**

4.1 Introduction . . . 18
4.2 Suitability of existing network simulators for our project . . . 18
4.3 Prestudies with NS-2 . . . 18
4.4 Requirements and Limitation of the Simulator . . . 20
4.5 Implementation . . . 21
4.5.1 Design . . . 21
4.5.2 Development process . . . 24
4.5.3 Configuration . . . 24
4.5.4 The Event Scheduler . . . 25
4.5.5 Handover decisions and bandwidth allocation in HNS . . . 26
4.5.6 Mobile nodes . . . 27
4.5.7 GUI . . . 28
4.5.8 Mobility . . . 29
4.6 Evaluation and Verification . . . 31
4.6.1 Logging of events . . . 33
4.6.2 Logging of base stations and access points . . . 33
4.6.3 Logging of nodes . . . 34
4.7 Analysis and Presentation of Results . . . 34
4.7.1 Evaluation relevant measurements . . . 34
4.7.2 Evaluation of results . . . 35

5 Simulation and results **37**
5.1 Introduction 37
5.2 Simulation results obtained from prestudies and prototyping 37
5.2.1 Mathematical Approach 37
5.2.2 First HNS Approach 42
5.2.3 Intermediate simulation results to evaluate the implementation of HNS 44
5.3 Simulation of user relevant aspects 51
5.3.1 Varying node density 52
5.3.2 Varying session density 54
5.3.3 Varying bandwidth ratio 55
5.3.4 Varying coverage of infrastructure technologies 56
5.4 Simulation of user and operator relevant aspects 56
5.4.1 User benefits 58
5.4.2 Operator benefits 66
5.5 Evaluation 73

6 Conclusion **75**
6.1 Summary 75
6.2 Future Work 75

7 Acknowledgments **77**

8 References **78**

A Integration of Swisscom Mobile Unlimited **79**

B The HNS configuration file **80**

C Implementation examples **82**
C.1 Java implementation of the Wireless LAN model 82
C.2 Java implementation of the UMTS model 83
C.3 Java implementation of the GPRS model 84

D Figures and Tables **86**

1 Introduction, Terms and Related Work

This work aims at analyzing the positive impact of the synergy of ad-hoc and infrastructure-based networks in terms of different prospects. The underlying technologies to facilitate such interaction are given in today's networks and allow high performance. This is attractive but not always intuitive to work with for most users. New solutions proposed by researchers and vendors take care of this aspect, thus the first steps toward seamless connectivity at anytime, anywhere and for everyone are being accomplished. Existing technologies already provide a useful infrastructure for heterogeneous networking. The Medium Independent Handover specification (MIH) - a future standard currently developed by the IEEE, called 802.21 - will provide enhancements for the interoperability of both IEEE 802-based and non-802 networks by adding an additional layer between Layer 2 and 3 of the ISO/OSI model.
The interest in heterogeneous networking has grown rapidly in the last few years. Appropriate seamless networking solutions respect the user's need regarding configuration and usability. Already existing products try to meet these requirements and provide a first standard of seamless networking to non-expert users. Continuing the efforts and following this idea by integrating more features and automated configuration will surely lead to big acceptance and therefore widely spread appliance. This also follows the trend toward tools for self-configuring ad-hoc connections, which hide complexity from the user.
Some terms often used in this document will be defined in the following paragraphs.

Ad-hoc peer-to-peer link: This is a direct infrastructure-less connection of at least two mobile devices. The association between the two nodes is supposed to be self-establishing and secured.

Heterogeneous network: Various kinds of heterogeneity do exist in combination with mobile networks. In matters of devices, heterogeneity would mean the mixture of different mobile devices whereas in terms of applications it could be a combination of different protocols and applications within the same network. Here, heterogeneity describes the manner of the network environment, i.e. a mix of different wireless technologies. Devices and applications are not further specified; they can be diverse or similar.

Handover: There are different forms of handovers between networks. Horizontal handovers depict a change between networks of the same kind and are already widely used. Vertical handovers are much more difficult to perform because they imply a change in access technology. When not stated otherwise, in this document a handover is always a vertical one. If the association to the target cell is established before deassociation of the previous cell this is denoted as *soft handover*, otherwise the term *hard handover* is used.

Seamless handover: In this document, seamless handovers refer to vertical handovers which are preserving the same network context, i.e the same Internet Protocol Address. Such a handover is rather fast performed compared to network changes where the whole context is torn down and reestablished again. Although the term *seamless* suggests a handover with no interruption at all, in the real world there can be noticed some breaks.

Network simulation: In academic research on networking a simulator is an abstraction of a real network to a program. Simulators are highly complex and do often simulate almost all aspects of a network from physical to application layer. This is not the case in this work. We simplify a network in big parts and just focus on some aspects by replacing the complex lower functionality by empirical or statistical models.

The goal of this work is to show the benefits of having seamless handovers in heterogeneous environments with an additional possibility to switch to non-infrastructure-based ad-hoc connections. Measurable criteria will be the enhancement of the mean throughput, the decrease of session duration and the resulting load reduction at the base station. In a second evaluation the power savings enabled by on-demand ability of the nodes will be examined based on empirical studies of networking hardware power consumption. An implicit requirement for the success of such evaluation is the availability of an appropriate simulation tool, which was planned to be an existing network simulator with possibility for custom extensions. However, preferred network simulators, namely Network Simulator (NS-2), Qualnet and Opnet, turned out to be unserviceable in the context of this work. Thus, a completely new simulator has been developed, which took, together with the evaluation of existing tools, the major part of the effort.
Regarding the planned simulation based implementation and evaluation of the mentioned features there seems to be no related work at the moment. One reason could be the missing standards in this field. The Mobile Independent Handover standardization process of the Institute of Electrical and Electronics Engineers (IEEE) is in a draft state only and will maybe become the state of the art in this field. However, until now, existing solutions are proprietary approaches developed by vendors and operators, which hide the architecture of their products. Relating to the simulation of heterogeneous networks there exists an interesting approach of George F. Riley from the Georgia Institute of Technology dating from 2000 [1], which introduces a dynamic simulation backplane to achieve interoperability of multiple different network simulators. The tool enables the use of different simulation instances for specific tasks in a simulation scenario. For our needs, this solution is not applicable because it is not able to deal with changing topologies, i.e. nodes, which have to change their attachment from one simulation instance to another during the simulation.

The remaining document is structured as follows: Chapter 2 first introduces wireless networking technologies, namely GPRS, UMTS and IEEE standards for WLAN. In the second part of the chapter, handover decisions in heterogeneous networks as well as bandwidth allocation models in wireless networks are discussed. Chapter 3 gives an overview of mobility and traffic patterns most commonly used in network simulation and presents popular simulation tools, such as NS-2, Qualnet, and Opnet. In Chapter 4 the implemented Heterogeneous Network Simulator is documented and its qualities and limitations are stated. Chapter 5 describes the setup of the simulations including used parameters and patterns and presents the results and interpretation. The conclusion and an outlook on future work on this topic conclude the document.

2 Wireless Network Technologies and Heterogeneous Networking Protocols

Heterogeneous networks benefit from various technologies. The more different networking technologies are available, the more total network capacity and coverage are offered. The following sections describe the network technologies used in the simulations. These are mainly WLAN, UMTS, GPRS and GSM. Of course there could be many more technologies considered to be part of heterogeneous networks such as UWB (Ultra Wide Band), EDGE (Enhanced Data Rate for GSM Evolution), Bluetooth, infrared, etc. However, this work focuses on the technologies that have already been deployed within existing heterogeneous networking hardware. The second part of the chapter covers handover decisions and bandwidth allocation in wireless networks.

2.1 Access technologies in wireless networks

2.2 IEEE 802.11

The 802.11 standard as described in [2] has been developed by a working group of the Institute of Electrical and Electronics Engineers (IEEE). It specifies the family of Wireless Local Area Networks (WLAN) with the underlying technology of Ethernet and CSMA/CA for medium-access. The basic specification of 802.11 applies for WLANs. It operates in the 2.4 GHz band and provides transmissions of 1 or 2 Mbps. A first security level is implemented by frequency hopping spread spectrum (FHSS) or direct sequence spread spectrum (DSSS) with phase-shift keying (PSK[1]) modulation. Not only is the rapidly and pseudo-randomly switching carriers of frequency hopping spread spectrum signals more difficult to intercept but they also provide more robustness against noise and interference. To generate direct spread spectrum sessions some pseudo-noise sequence is added to the signal at phase modulation.

802.11a: An extension to the basic WLAN is the 802.11a standard. It provides a throughput of up to 54Mbps in the 5GHz band. Unlike the original protocol it uses the orthogonal frequency division multiplexing (OFDM) encoding scheme, which is based on frequency-division multiplexing (FDM) where multiple orthogonal signals are sent out at different frequencies. Each of the 52 used OFDM subcarriers can use BPSK[2], QPSK[3], 16-QAM[4] or 64-QAM encoding. As PSK is easier in implementation it is more widely used, for example in Bluetooth.

802.11b/g: The 802.11b standard defines another extension to the initial specification.

[1] Phase Shift Keying is a digital modulation scheme that transmits data by modulating the phase of a reference signal (the carrier wave).

[2] Binary Phase Shift Keying is the simplest form of PSK because only two phases are used to modulate the signal

[3] Quadrature Phase Shift Keying uses four phases to modulate the signal.

[4] Quadrature Amplitude Modulation is the encoding of information into a carrier wave by variation of the amplitude of both the carrier wave and a 'quadrature' carrier that is 90 ° out of phase with the main carrier in accordance with two input signals.

Complementary code keying (CCK)[5] can be used for modulation additionally to the 802.11a modulation techniques. Compared to the original standard, this allows higher data speeds of up to 11Mbps and is less fragile to multipath propagation interference. Even faster transmissions can be obtained with the 802.11g standard, which is operating in the same 2.4GHz band but providing speeds up to 54Mbps using Orthogonal Frequency Division Multiplexing (OFDM).

802.11e: This part of the 802.11 family introduces Quality of Service (QoS) for multimedia support to the existing 802.11 standards while maintaining full backward compatibility. QoS is of value in home and business environments where audio- and video-on-demand as well as voice over IP capabilities are getting more and more important.

802.11i: This standard finally adds more security through the Advanced Encryption Standard (AES) protocol. This block cipher algorithm is known as the successor of 3DES (Data Encryption Standard). The most important advantages are fast encryption and decryption in hard- and software and little memory usage.

2.3 GSM – Global System for Mobile Communications

The Global System for Mobile Communications (GSM) is surely the most popular standard for mobile telephony in the world. It was first commercially applied in 1991 in Finland as the first standard of the second generation (2G) of mobile networks. It differs from all its preceding standards in such that both channels – signaling and voice – are digital. As hardware is not addressed in the GSM specification, every operator is able to deploy equipment from arbitrary vendors due to simple interoperability allowed by the open standard. Further, international roaming services permit the subscribers to use their mobile device in many countries. Most GSM Networks operate at 900 and/or 1800 MHz with few exceptions. At this frequency, mobile phones connect to the cellular network by searching for cells in their vicinity. Signals are modulated by QPSK technology with enhancements from Gaussian Modulation.
The key elements of the GSM network design are structured in three main parts as shown in Figure 1. Base Transceiver Stations (BTS) receive and transmit radio signals. Up to 100 Base Transceiver Stations are under control of one Base Station Controller (BSC), which provides the intelligence behind. It allocates the radio channels, controls handovers from BTS to BTS and it acts as concentrator for low capacity connections toward the Mobile Switching Center (MSC). The MSC is logically placed in the Network and Switching Subsystem, which is at the same time the core of the GSM network. The MSC provides circuit switched delivering of calls to subscribers, mobility management by handling handovers from BSC to BSC and GSM services to mobile phones being in the served area. Integrated or very tightly linked with the MSC there is the Visitor Location Register (VLR). This temporary database holds the roaming subscribers in this area. The data stored in the VLR includes IMSI (International Mobile Subscriber Identity, the subscriber's identity number stored on the Subscriber Identity Module), MSISDN (Mobile Subscriber Integrated Service Digital Network, the subscriber's mobile phone number), the services the subscriber is al-

[5] Complementary Code Keying operates with a pair of finite bit keys in which the number of identical bits is equal to the number of different bits (example: [1,0,1,0], [1,0,0,1]).

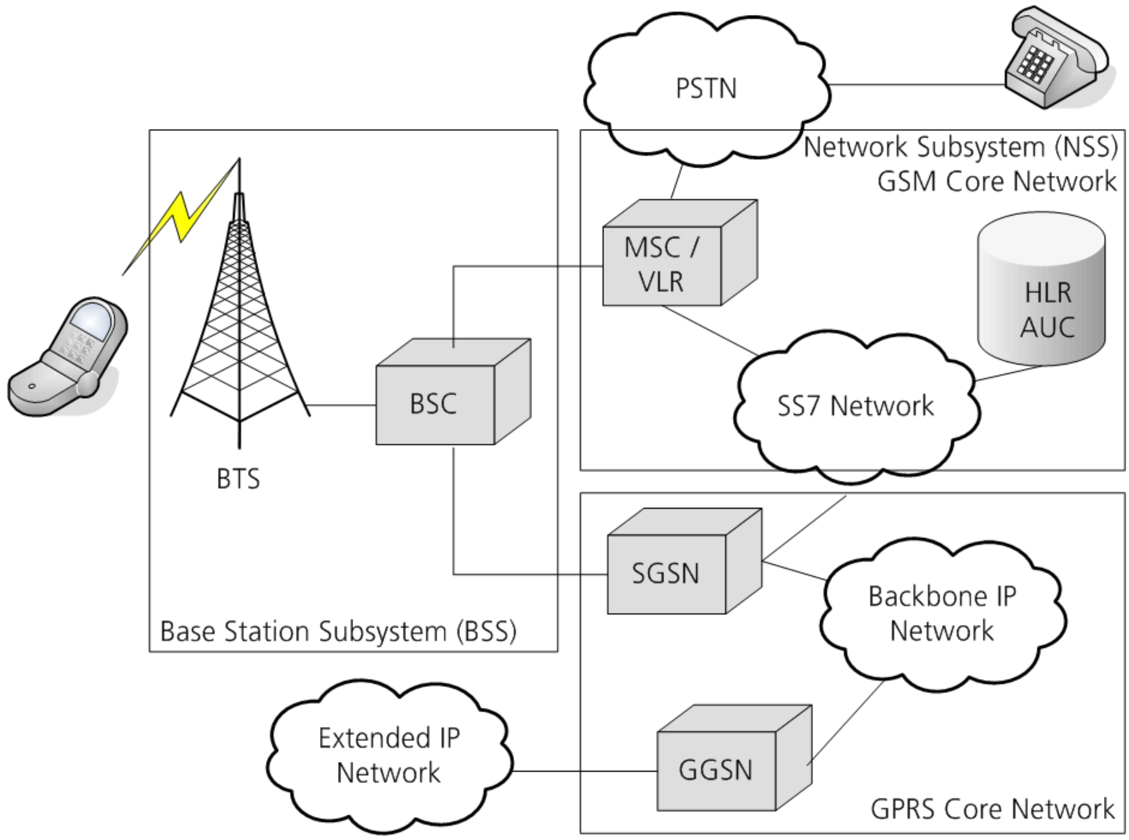

Figure 1: Overview of the GSM

lowed to access, Home Location Address of the subscriber and authentication data. The VLR informs the Home Location Register (HLR) about newly arrived users in the cell, tracks subscribers within the servicing area, allows or disallows services for the subscribers and deletes records if subscribers are inactive for a too long time or moving out of the region. The HLR is a central database that stores the entries of all mobile phone numbers, which are allowed to access the GSM core network. The entries consist of the SIM data, i.e. IMSI as unique identifier, subscribed services, current location, etc. The HLR provides data to the VLR and SGSN (Serving GPRS Support Node) when a user roams in for the first time, mediates between the Gateway MSC and the Short Message Service Center (SMSC) to allow incoming calls or messages and removes subscriber data as soon as subscribers roam away. The third logical part of GSM Network is the General Packet Radio Service (GPRS) core network, which provides functionality for the GPRS network. In contrast to the circuit switched GSM network, the GPRS network provides mobility and session management and transport for IP packet services. The main elements discussed here are Gateway GPRS Support Node (GGSN) and Serving GPRS Support Node (SGSN). The GGSN acts as a Gateway and detunnels data from inside the core to send out normal IP packets and vice versa. The SGSN interworks with the connected radio network (BSC) and the GSM core. Different from the code division multiplexed GSM channels, GPRS is time division multiplexed. By using unused slots of the GSM, higher data rates can be achieved. Due to all this enhancements, GPRS is denoted as 2.5G (second and a half generation).

For a summary of the characteristics of the discussed technologies refer to Table 1.
An important key feature of the GSM is the Subscriber Identity Module (SIM). Stored on this smartcard also known as SIM card there is the user's subscription information and personal data of the user, e.g. phone book. From a technical point of view, the SIM is a microcomputer equipped with Central Processing Unit (CPU), Read Only Memory (ROM), Random Access Memory (RAM) and Input/Output functionality. The SIM Serial Number (SSN) can uniquely identify each SIM.

2.4 UMTS – Universal Mobile Telecommunications System

Unlike the GSM technology that is part of the 2^{nd} generation of mobile communications, UMTS is a 3rd generation technology whereas HSCD (High Speed Circuit Switched Data) and GPRS (General Packet Radio Service) build the the 2.5^{th} generation. For the time being the UMTS standard has been implemented into the core network of GSM/GPRS, which required some adjustments to be performed. The UMTS terrestrial air interface consists of FDD (Frequency Division Duplex) and TDD (Time Division Duplex). For FDD, the frequency band has to be split up into a band for the upload and one for the download with a minimum guardband between the two bands to avoid interference, which also introduces waste of bandwidth. TDD has the advantage of variable bandwidth allocation for upload and download.
FDD is an older scheme that was best suited for applications that generate symmetric traffic such as voice. In return, TDD is best suited for bursty, asymmetric traffic, namely Internet, and other data centric services because it is more flexible in dynamically reconfiguring the allocated bandwidths for up- and downstream than FDD. TDD also uses the spectrum more efficiently because there is no guardband required between transmit and receive channels.
Compared to GSM/GPRS, this results in higher capacities, lower transmitting power and better connection establishment at a maximum bandwidth of 1.23 Mbps. There is also less inference due to the frequency channels that are 25 times larger than the GSM ones. Table 1 shows the interrelation of the discussed wireless technologies.
The UMTS access network consists of Radio Network Controllers (RNC) and Node-Bs [3]. The RNC is responsible for the routing of connections between the Node-B and the core network as well as for the surveillance of the Node-B's activities. It also provides a relocation function to optimize the routing and to support hard handovers. The Node-B has to provide send and receive functions in one or more cells, depending if it consists of one single antenna or of multiple sector antennas. CDMA networks provide a soft handover function without interrupting the data stream. If a mobile node comes closer to the edge of a Node-B where it receives signals from other Node-Bs it is able to establish multiple connections and then release the initial one.
UMTS enabled devices such as mobile phones and PCMCIA cards for Laptops are already widely distributed and inter-technology handover services between GRRS an UMTS are being integrated.

2.5 UWB - Ultra Wideband

Ultra Wideband (UWB) radio communication technique got its name from using very large fractal spectrum in the frequency band. Different from all other radio frequency communications, UWB does not use a specific carrier but operates with modulated high frequency low energy pulses of less than one nanosecond of duration. The amount of spectrum occupied by a UWB signal, i.e. the bandwidth of the UWB signal is at least 20% of the center frequency. Thus, a UWB signal centered at 2 Gigahertz occupies 400 Megahertz and the minimum bandwidth for a signal at 4 Gigahertz would be 800 Megahertz. These properties allow achieving data rates of up to 1 Gigahertz per second on distances of less than 10m. Although this sounds very promising and there is a chance that UWB will replace other short distance technologies and even cables soon, there are still some concerns. Since very short duration pulses are transmitted in this large bandwidth spectrum, some fear that there will be too much interference with other technologies. This could be avoided either by using only low transmission power or by restricting UWB in certain spectrum. Another difficulty with UWB rises at reception of the signal, which is based on time-correlation of pulses. Because of the low duty cycles and the short duration of pulses it is difficult to synchronize sender and receiver. Nevertheless, semiconductor companies and experts are confident that UWB will come on the market soon. Once all the standardization and regulation aspects would have been finished, UWB could potentially replace short distance peer-to-peer communication and increase the bandwidth of ad-hoc peer-to-peer connections by a factor of 200 or more. Since there are no concrete indications of when and how UWB will be allowed in Europe, the simulation parameters for peer-to-peer ad-hoc connections in this work are based on both, conventional WLAN properties and UWB.

2.6 Summary of wireless network technologies

Technology	Frequency Range (GHz)	Number of users per channel	Peak data rate
WLAN 802.11a	~ 5	127	54 Mbps
WLAN 802.11b/g	~ 2.4	127	11/54 Mbps
GSM	$\sim 0.9/1.8/1.9$	8	14.4 Kbps
GPRS	$\sim 0.9/1.8/1.9$	8	115 Kbps
UMTS	$\sim 0.9/1.8/1.9$	15-50	1.23 Mbps
UWB	N/A	N/A	1 Gbps

Table 1: Overview of the technologies (Values for Europe)

Table 1 gives an overview on some general parameters of the discussed technologies. These values build the base for the technology models used in simulation. Parameters such as the maximum number of users per cell, coverage radius and throughput of the different technologies as well as the handover behavior from one to another technology are relevant for the simulation described in Section 4.5.5.

In a heterogeneous network, not all of the described technologies offer the same performance and thus, each of them fits for different tasks. To periodically check e-mails for example, GPRS would perform well enough as long as no big messages have to be synchronized. To browse the web, in contrast, GPRS would not be the preferred technology to use, WLAN or UMTS would fit much better. Hence, in a heterogeneous network it is not sufficient to just choose a technology, but also the benefit for every user depending on his preferences needs to be optimized. This is where handover decisions play an important role, which is also discussed in the next chapter.

2.7 CAHN - Cellular Aided Heterogeneous Networking

A mobile ad-hoc network is a collection of wireless terminals that can be deployed rapidly. Integrating ad-hoc networks with a well-established cellular network can improve communication and security and enrich the cellular services. In CAHN it is proposed that the entire network architecture uses the cellular network to facilitate the mobile ad-hoc networking. All signaling goes via the cellular network while the data is delivered through the ad-hoc links. Future wireless technology aims at providing an umbrella of services to its users. Ad-hoc networks have become attractive because of their potential in commercial applications. In ad-hoc networks, the issues of quality of service (QoS) and security are complicated because of the lack of reliable methods to distribute information in the entire network. The integration of heterogeneous wireless technologies can improve the network performance, thereby meeting the demands for QoS. The CAHN concept proposes a novel integrated architecture to improve ad-hoc networking. The architecture follows the idea of "out-of-band signaling" (over a cellular network) enabling an ad-hoc network to improve the quality of network control and management [6].

Auto-configured ad-hoc mode in mobile networks Ad-hoc WLAN connections can be built up as soon as two mobile devices are insight of each other. To prevent the auto-configuration process of establishing ad-hoc connections with strangers, there must be an identification process amongst peers. If the users of the end-devices being configured are sitting side by side they could use a shared secret (e.g. PIN) as it is used to secure Bluetooth applications. However, this manual key management is not secure nor convenient. Nodes within reach of direct communication should have the ability to switch to ad-hoc mode with a certain peer without being bothered with complicated configuration decisions. For this case it is proposed to use the mobile phone number (aka MSISDN) as an identifier and USSD to initiate a secure ad-hoc connection for peers as it is proposed in [7].

1. Connection request from Alice to Bob's MSISDN over USSD or by SMS, including the IP address of Alice's device
2. Location of Bob's GSM device by paging
3. GSM paging response
4. Connection request of Alice delivered to Bob's device
5. Transfer of the connection request from Bob's phone to his computer

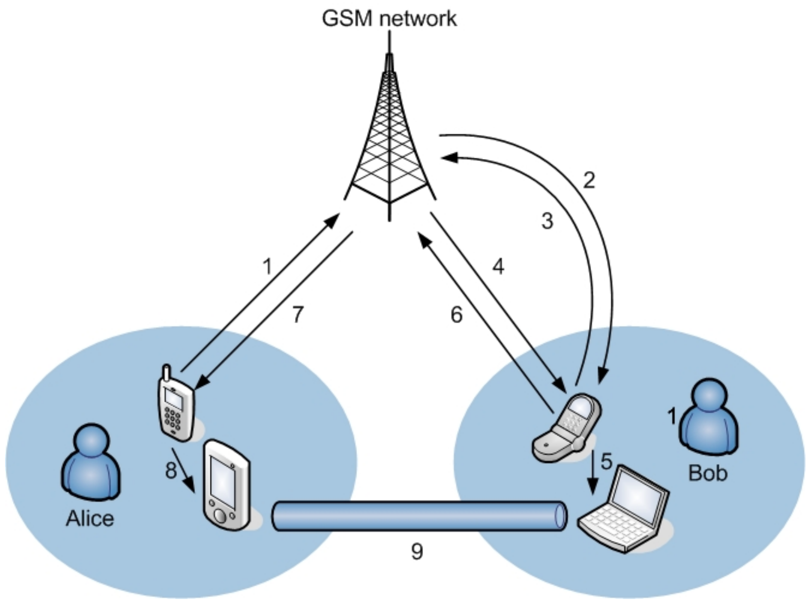

Figure 2: CAHN connection establishment

6. Connection response including IP address of Bob's computer as well as security and connection parameters to Alice's MSISDN

7. Connection response from Bob to Alice

8. Transmission of the connection response to Alice's end device

9. Secured link establishment between Alice's and Bob's end devices

Allowing ad-hoc peer-to-peer connections in heterogeneous networks is only beneficial if the connection is as secure as those in other technologies. CAHN is responsible for setting up, managing and terminating ad-hoc connections. As documented in [8] the CAHN protocol fulfills this request for security and also provides some mechanism to automatically establish the data link between peers hiding complexity from the user. Having data and signaling plane separated from each other is the main idea of CAHN. The signaling is considered to take place over a cellular network providing high coverage and also being responsible for the authentication of the users via their mobile phone number (MSISDN). The link establishment proceeds according to Figure 2. Based on the CAHN protocol, ad-hoc connections in heterogeneous networks could be established and put into operation seamlessly. The realization is possible in different ways: For example the ad-hoc peer-to-peer connection could be initiated based upon personal phone book information and established as soon as the peers have entered the shared key. Seamless handovers from another technology to a peer-to-peer connection and vice versa can then be enabled using the Mobile IP Route Optimization. This functionality is called SMACS (Smart Multi Access). The SMACS layer is logically placed above the CAHN layer and is introduced to handle end-to-end IP sessions. Therefore, heterogeneous IP sessions can be routed either using infrastructure based technologies or using direct ad-hoc links provided by CAHN. An Adapter is used to

adapt CAHN messages for underlying technologies, whereas a Connector configures specific security parameters, IP address and frequency.
Some definitions that are relevant in the context of this work are given in the following paragraphs:

Always-on mode: The default behavior of most of today's mobile devices is to either be always online (always-on) or switched off completely. Whereas in the switched-off state there can not be any communication at all, in the switched-on state the device is ready to communicate at every time.

On-demand mode: The idea of the on-demand mode is to put networking devices to sleep when no data is being sent or received. This is heavily decreasing the power consumption of a mobile device because networking hardware is relatively energy consuming. Whenever there is a need to communicate, which is detected by the server side, the device is woken up, synchronized, and put into sleep state again.

Session: This denotes the context of connection during the transmission of a data volume. A session begins with the initiation of a connection and ends with the last packet of the data volume to transmit. For example, a session can be the transmission of an email, the download of multimedia content or the upload of a document to some shared folder at the workplace.

3 Mobility and Traffic in Wireless Network Simulation

3.1 Mobility models

For the abstraction of mobile networks a mobility pattern according to the setup of the simulation scenario must be chosen. Many of these patterns have been implemented in the well known and often used network simulators, which are discussed in Section 3.2. Some of them are just acting randomly with no limitations whereas others follow more restricted rules. An overview of the different mobility patterns is given in Figure 3. The mobility patterns integrated in the developed simulator are described in the next sections. They fall into the three categories Random Models, Models with temporal dependency, and Models with spatial dependency. Parameters, which are important for all mobility models explained in more detail subsequently are the following:

- Number of nodes: The number of nodes which should be moved by the mobility model
- Size of the simulation area
- Scenario duration: The whole duration of the scenario, inclusive the initial cutoff phase
- Duration of initial cutoff phase: Defines how many additional seconds at the beginning of the scenario should be skipped. This is important for the random distribution of the nodes.
- Seed: The random seed for the movement behavior. This makes scenarios reproducible.

The specific parameters for each mobility model is explained in the corresponding sections. They correspond to the parameters used in the BonnMotion implementation [14].

3.1.1 The Random Waypoint Mobility model (RWP)

This is probably the most popular and commonly used mobility pattern. It is a simple model, following stochastic rules and straight forward in implementation. The following steps represent the node's moving pattern:

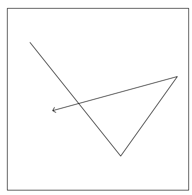

Figure 4: RWP

1. A node randomly chooses a destination point (waypoint) in the area and moves with constant speed on a straight line to this point.
2. After waiting a certain pause time, it chooses a new destination and speed, moves constantly with that speed to the destination, and so on.

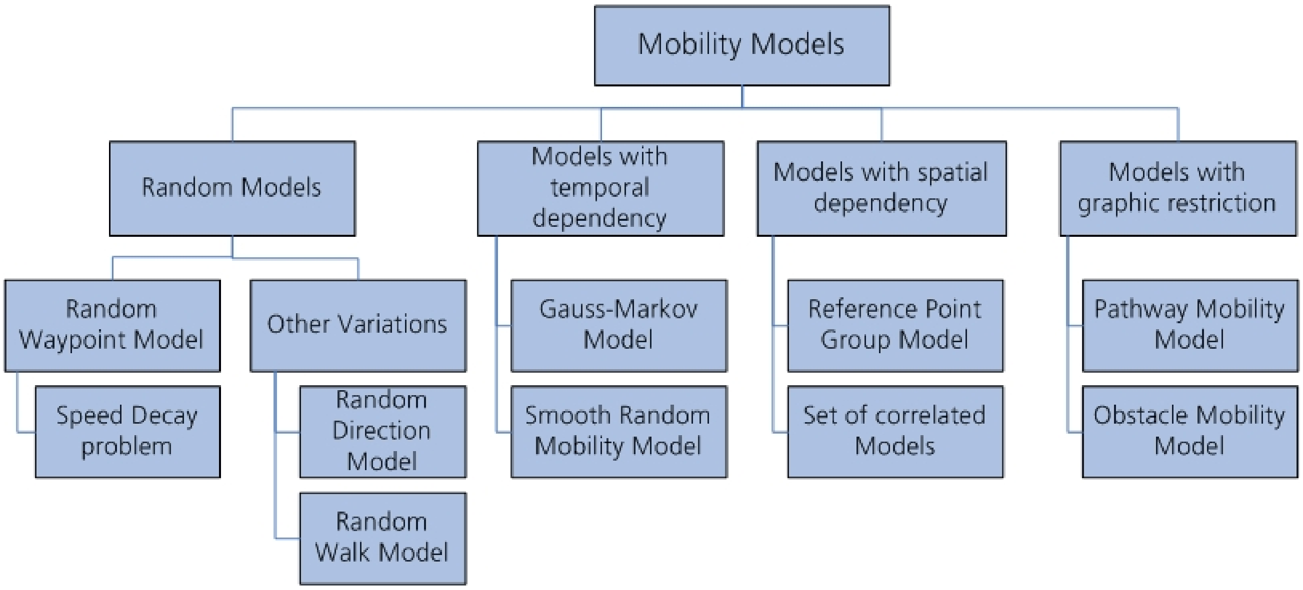

Figure 3: Overview of mobility models

3. The destination points are supposed to be uniformly randomly distributed in the simulation area.

However, for several stochastic reasons, the distribution of the nodes in a simulation area is independent of the speed and non-uniform as shown in Figure 5.

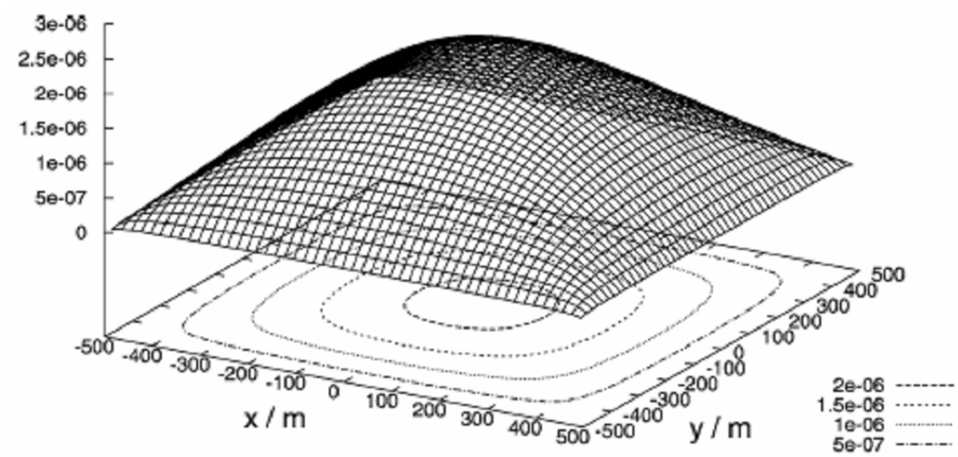

Figure 5: RWP mobility node distribution [13]

The probability to find a node at the border of the simulation area goes to zero. This is caused by the fact that nodes choose a non-uniformly distributed direction angle at the beginning of each movement period. The density function of the direction angle is determined by the shape of the simulation area and the starting waypoint as described detailed in [13]. The BonnMotion mobility generator [14] used in the presented simulator also provides the simulation of attraction points additionally to the normal RWP movement

behavior. Using attraction points, locations that are frequented more often by users can be modeled, e.g. cities (larger scenarios), train stations, access points or shopping malls (smaller scenarios).
By default, the parameters, which are configurable in the RWP Mobility Model additionally to the general ones are the following:

- Minimum speed: The maximum speed, which nodes can choose randomly to travel the simulation area.
- Maximum speed: The minimum speed, which nodes can choose randomly to travel the simulation area.
- Maximum pause time: The maximum duration, which nodes can choose randomly before moving to the new destination.

A derivation of pure random mobility is the introduction of attraction points, which can be specified as coordinates.

3.1.2 The Gauss-Markov model

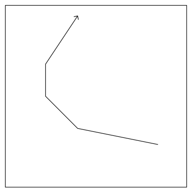

Figure 6: Gauss

In probability theory, a stochastic process has the *Markov property* if the conditional probability distribution of future states of the process given the present state depends only on the present, i.e. it is conditionally independent of the past states (the process progress). Derived from a normal distribution (which is often called Gaussian distribution) a *Gaussian function* is specified by a mean and a covariance function. If the mean function is x then the covariance function $C(x, x')$ represents the expected covariance between the value of y in function of x and x'. The actual function y at point x is assumed to be a single sample from the Gaussian distribution.
The Gauss Markov process for mobility modeling is equipped with a memory to provide a time-correlated velocity model, instead of a pure uncorrelated Markov process. Thus, additionally to the minimum and maximum speeds a memory degree in [0,1] has to be specified, which is further explained later in this section. The procedure for this mobility behavior can be summarized as follows:

1. A node is assigned a certain speed and direction at the beginning of the simulation. For a fixed duration, the node travels through the simulation area according to these values.
2. After this fixed time the node chooses other destination coordinates and speed, not just randomly as in RWP mobility model, but with respect to the previously chosen values. The node's speed v_n is set to

   $$v_n = \alpha v_{n-1} + (1-\alpha)\overline{v} + \sqrt{(1-\alpha^2)} v_{x_{n-1}}$$

and the node's direction θ_n is set to

$\theta_n = \alpha\theta_{n-1} + (1-\alpha)\overline{\theta} + \sqrt{(1-\alpha^2)}\theta_{x_{n-1}}$

where v_n is the speed and θ_n the direction of the node at the time interval n, $0 \leq \alpha \leq 1$ is the tuning parameter, which defines the degree of motion randomness. $\overline{v}$ and $\overline{\theta}$ are constants representing the desired mean speed and direction, and $v_{x_{n-1}}$ and $\theta_{x_{n-1}}$ are random variables from a Gaussian distribution. Totally random motion (Brownian motion) can be obtained by setting α to 0, linear motion is obtained by setting α to 1.

This model is maybe more appropriate to the behavior of humans compared to the RWP procedure because their future location and speed are not completely unknown but can be predicted from the parameters already gathered from the way they have traveled so far. Important parameters to specify for the Gauss Markov Model are the following:

- Maximum speed of the node
- Minimum speed of the mode
- Speed standard deviation
- Angle update frequency
- The angle standard deviation

3.1.3 The Reference Point Group Mobility model (RPGM)

The mobility behavior in the RPGM model is not characterized by the random path of individual mobile nodes but the movement in groups. The idea of having a set of nodes moving together in approximately the same direction is implemented with reference points, which hold the group together.
As the movement pattern of the nodes is based on the Gauss-Markov model, the procedure for the path decision is the following:

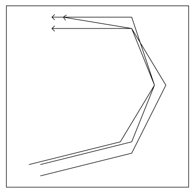

Figure 7: RPGM

1. From the last reference point the group chooses a new one according to Gauss-Markov model, e.g. with a higher probability for directions that do not differ too much from the last direction.
2. After that, a speed is chosen according to a Gauss-Markov process. Together with a destination inside the radius around the chosen reference point, this procedure results in having groups traveling through the simulation area, held together by the specified radius and choosing a time- and destination-correlated path according to the defined memory value discussed in Section 3.1.2.

3. When a group of nodes crosses another group, the individual nodes can decide to change to the other group with a predefined probability with respect to the minimum and maximum group sizes, which are also specified before simulation.

The following parameters should be specified for this model:

- Maximum distance from a node to its group's center
- Average group size
- Standard deviation from the average group size
- Probability with which nodes change to an other group, which is crossing the way

3.1.4 The Manhattan Grid model

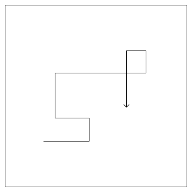

Figure 8: Manhattan

This mobility pattern tries to model the behavior of humans in a city. The topology of the city is assumed to be rather modern, with straight streets crossing each other orthogonally. The resulting grid is then traveled by the mobile nodes, which decide at each junction whether they want to change direction or not, according to a given turn probability. Their velocity is determined by a minimum and a maximum speed as well as a pause time. The density of the grid is defined by the number of blocks in both directions of the two-dimensional simulation area. As most European cities are not built up this way, the Manhattan Grid model is not used in the simulations described later in this work although it is integrated in the simulator.

The following parameters specify the behavior of the Manhattan Grid model:

- Number of blocks on the x-axis direction
- Number of blocks on the y-axis direction
- Distance interval after which the node possibly updates its speed
- Probability for the node to turn at a crossing
- Probability for the node to change its speed after each distance interval
- Mean speed of the node
- Minimum speed of the node
- Standard deviation of normally distributed random speed
- Probability of the node to pause given the node does not change its speed after the given distance interval
- Maximum pause time of the node

3.2 Network simulation tools

Many existing network simulators permit to set up and run simulation scenarios that take into account the whole ISO/OSI stack and therefore should behave almost like real networks. One goal of network simulation is to predict the state of a system or a network during a certain period of time. This can be done by modeling the actual system and then evaluating its behavior. As a continuous evaluation of the system state is a difficult and time-consuming task, a common approach is to evaluate the behavior of the system only at discrete moments in time.

In a *discrete event simulation* events are sorted by the time they occur and processed chronologically. Most networking problems can be treated by discrete event simulation, as the state of the system only changes due to certain events, such as the start or the end of the transmission of packets. In this type of simulation, the simulator maintains a queue of events sorted by the time they should occur. The simulator reads the event list and schedules succeeding events as each event is processed. It is not necessary that the simulation is executed in real time. It is more important to extract relevant information, in order to allow a statistical evaluation of some performance metrics after the end of the simulation process.

3.2.1 NS-2

NS-2 9 is a discrete event simulator designed for academic and industrial networking research. It provides substantial assistance for the simulation of network behavior with full support for the lower layers to examine delay, jitter, packet loss, etc. Many basic concepts have been implemented over time such as different routing algorithms, transport protocols, media access protocols, queuing algorithms, mobility models and many more. The simulator is split in two parts, one implemented in Tcl[6] and the other in C++. The basic advice is to use Tcl for configuration, setup and one-time activities and to use C++ to do anything that requires the processing of each packet of a flow or whenever the change of an existing C++ class in ways that have not been expected is necessary, e.g. the particular change of an antenna pattern.

NS-2 began as a variant of the REAL network simulator in 1989 and has evolved substantially over the past years. Many contributors have included their new modules. The big acceptance of NS-2 in industry as well as in the academic world results from the fast evolution and also the fact that there is no purchase fee.

3.2.2 QualNet

In contrast to NS-2, QualNet [10] is commercially distributed by Scalable Network Technologies (SNT). The latest version includes an embedded mode to accept real network traffic and supports Policy Based Routing, SIP (Session Initiation Protocol), physical models such as microwave, Fast Rayleigh Fading and also a weather model for tropospheric fading. A three-dimensional graphical user interface for extremely realistic visualizations of communication networks gives the additional real-world-touch. The implementation is in

[6]Tcl: The Tool Command Language is a script language most often used for prototyping and testing.

C++ and well documented in most parts. The simulator provides similar features as NS-2, which most of the time are released later but are also more stable. The high purchase fee is somehow justified since QualNet provides a fast and very accurate support service.

3.2.3 Opnet

Opnet [11] is a highly evolved simulator, which provides professional graphical user interfaces on the three modeling domains into which the simulator is split up, i.e. one domain for the network, one for the node and one for processing. The network domain describes the general network topology in terms of sub-networks, nodes, links and geographical context. The node-internal architecture is handled in the node domain and treats flows between nodes such as functional elements whereas the processes (protocols, algorithms, applications), specified by finite state machines and extended high-level language, are described on the processes domain level.
Like the QualNet simulator, Opnet has to be purchased and is therefore more often used in industry than at universities.

4 Design and Functionality of Heterogeneous Network Simulator

4.1 Introduction

As documented in the preceding section, NS-2 and other existing simulators already provide complex functionality and many features to simulate mobile networks. However, existing simulators entail several problems regarding heterogeneity, which lead to the consequence to build up a simulator that fits the needs in terms of heterogeneity. The reasons for not choosing an existing simulator but building up a new one are explained in the first section. After that, the requirements of the newly developed Heterogeneous Network Simulator (HNS) and the desired functionality are listed. The last section discusses implementation issues such as the design of the whole program, the implementation of specific modules, the configuration of the simulator, and possibilities of analyzing results.

4.2 Suitability of existing network simulators for our project

The network simulation tools, which have been discussed in Chapter 3.2, are widely used to simulate isolated aspects of specific network topologies. Most academic network simulations focus on the measurement of packet delay, packet delivery rates, jitter etc. in homogeneous wired or wireless networks. In contrast, dealing with a heterogeneous topology where multiple existing communication technologies can be used simultaneously is much more complex and not implemented in existing simulators. All the source code of NS-2 is available and various modules exist in order to simulate UMTS, GPRS, GSM, WLAN with Mobile IP and even Bluetooth. Therefore, one could assume that it would be feasible to reuse all this code and to implement a module enabling seamless handovers among them. Unfortunately, the possibility of integrating the various network modules into the same scenario turned out to be rather complicated and even impossible without the huge effort of implementing new modules and changing big parts of the existing code. However, the idea of using NS-2 is further investigated, which is documented in the next section.

4.3 Prestudies with NS-2

This section shows all steps performed before rejecting NS-2 as the basic skeleton for the planned heterogeneous simulations. At the beginning, several Wireless LAN networks are deployed over an NS-2 simulation area and multiple mobile nodes moving according to the Random Waypoint Mobility model are created. Constant Bit Rate (CBR) traffic is set up between pairs of nodes. The results of this test run show that one mobile node can only be attached to one access point. The nodes are not capable to register at any other access point than their own.

The main functionality used for our investigations are provided by Mobile IP [5], which enables seamless handovers. It is either available as an additional module for the version NS-2.1b6 or built in from the version 2.26. With both versions as a basic system, the required features are implemented one by one and the newly gained configurations have are verified iteratively. Unfortunately, the GPRS and UMTS modules are designed to be

used isolated and the patch to make the extra modules work comments out pieces of code belonging to the Mobile IP package. One reason could be that it is not intended by the developers to use UMTS and GPRS together with 802.11 networks simultaneously. With a reasonable effort it would be possible to find a solution for this problem; however, the basic functionality of Mobile IP is necessary for the planned evaluations.
Another test-run shows that mobile nodes have to be in their Home Agent's (HA) range at the beginning of the simulation; otherwise, they will never register with some Foreign Agent (FA). Therefore, a new test is set up with an initial configuration, where all nodes are in the range of their HAs at the beginning of the simulation. Before starting traffic generation, the simulation is run for 500 seconds in order to distribute the mobile nodes in the simulation area. Now, the packets are sent by all nodes, but they still do not always reach the destination node. Further investigations demonstrate, that the packets from mobile node to mobile node are only delivered, if both are registered at their HAs. However, the packets are always delivered if the receiver is a wired node. These findings lead to the next test scenario, where all packets are forwarded to a wired node, which relays them to the mobile destination node. To a certain degree this seems to work out and solve the problems mentioned before. Unfortunately, additional prestudies reveal anomalies in the behavior of mobile nodes reentering a HA's cell after having traveled outside the HA range. Therefore, some supplementary tests are made to qualify the problem. The results point out that the behavior is reproducible and shows up every time a mobile node reenters its HA's cell and wants to reestablish a connection after having traveled the simulation area without having had a binding to some infrastructure for a too long time. Two test scenarios, as shown in Figure 9, have been performed to show the drawback of the Mobile IP implementation in NS-2, which is therefore not suited for our needs.

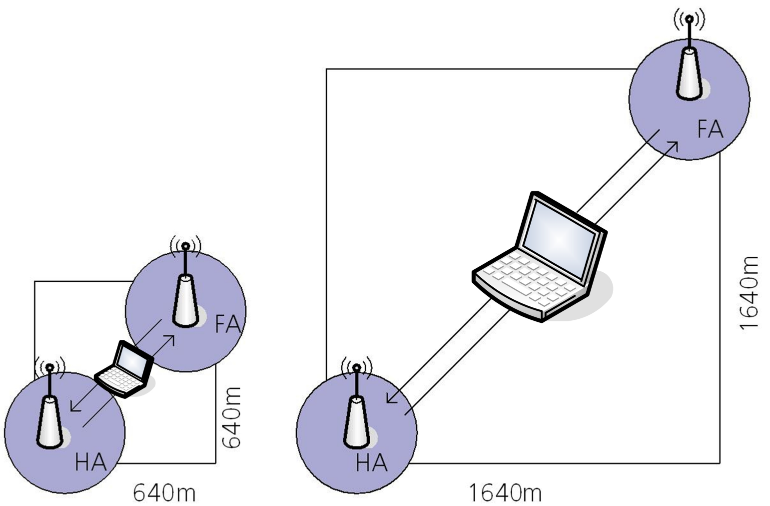

Figure 9: Two differently scaled test scenarios

In the small test scenario a field of $409'600$ square meters with a HA in the lower left and a FA in the upper right is set up. A mobile node starting at the HA travels to the FA

and back sending packets with a constant rate to a wired node located nearby the HA. In this scenario, all packets are reaching the destination node except for the packets sent in the non-covered area in the middle of the field. However, the node registers with the FA and finally also with the HA again, when returning. This is not the case for the large test scenario. Here, everything works fine until the mobile node has to register with the HA again, on its way back. One reason might be a timeout on the HA, which rejects the mobile node after a while. Another possible cause could be too small queues, which just discard sent packets after a long time without connectivity. This would mean for the simulations that a node doesn't have a realistic possibility to use WLAN when the coverage is sparse. A lot of further investigation and many fixes to the existing code would be necessary just to decide about the feasibility of the development of a heterogeneity module for NS-2. This is an unknown parameter in terms of time investment before even starting with the real implementation of an additional heterogeneous layer.
A real alternative wold be the implementation of a backplane as proposed in [1]. The idea plans to have several instances of NS-2, each of them dedicated to one wireless network part. Concretely, there would be four implementations of NS-2 for each of the technologies GPRS, UMTS, WLAN, and peer-to-peer connections over WLAN. An additional proprietary instance programmed in either C, C++ or Java would manage the simulation mission and dispose the subtasks to the single instances. This is an interesting implementation challenge and the only unsolved problem is the management of connections between two nodes using different technologies, e.g. one node being in the range of a GPRS base station sending to a node, which has only WLAN available. Different instances of NS-2 running at the same time would maybe solve the problem. Some channels in the backplane in order to pass packets from one instance to another would serve as a message passing system. Apart from the introduced clock synchronization problems, the implementation of such a backplane would be a huge and complex job going beyond the scope of one diploma thesis. As a further alternative, Opnet and Qualnet are also examined. The problems with Qualnet are exactly the same and furthermore, for Qualnet, not all the source code is available and the degree of possible modification is limited.
Thus, after a long and unsuccessful evaluation phase on simulating heterogeneous networks in NS-2, we decide to build up a new simulator with less lower layer complexity than NS-2 but providing the required functionality to handle heterogeneity. Like this, the effort is not necessarily smaller but much easier to calculate and the risk of being unsuccessful is reduced.

4.4 Requirements and Limitation of the Simulator

Because in the context of this work it was not possible to cover all aspects of a network simulation, a focus had to be set on important features and some functionality had to be simplified. As the results of the work are supposed to show connectivity benefits from a user's point of view, lower layer characteristics such as detailed propagation models, routing protocols, queuing, and medium access layer protocols were less important than heterogeneous end-to-end connectivity properties and multiple network technologies usable simultaneously. An additional problem when modeling heterogeneous networks is the lack of standardization of heterogeneous networking protocols. Thus, no model of handover-techniques among different technologies, bandwidth allocation etc. are available to abstract from. The decision to build-up a simulator has introduced many limitations compared to a real simulation, which especially concerns the lower layers. In return, many features which are missing in common simulators have been implemented, e.g the possibility to use all implemented network technologies at the same time, the session based setup and evaluation ability, etc.

Requirements The simulator should support nowadays network technologies and provide a function to handle seamless handovers between them. Additionally, it should be able to detect situations where links could be built up spontaneously between nodes via ad-hoc connections. An important part of the work would be to implement an event-queue to handle the connection requests as well as other occurring events and a graphical "real-time" output to indicate mobility, connections and handovers. For the post-processing of the simulation several log files should be generated in order to evaluate the desired parameters with appropriate graphs. The following sections describe how these requirements have been fulfilled in the new simulator.

4.5 Implementation

4.5.1 Design

As discussed in Chapter 3, well-known network simulators are implemented in C++ or C with support of scripting languages for the definition of the simulation scenarios and for the evaluation of results. The Heterogeneous Network Simulator (HNS) is programmed in Java, which is well suited for an implementation with this level of abstraction because it is highly object-oriented and comprehensible and thus easy to extend with new models. The simulator works event-driven and session oriented. This means, instead of treating single packets, whole streams are processed and split into pieces, whenever the network environment changes, e.g. new sessions are initiated, and handovers are processed. Regarding mobility, HNS is compatible with NS-2 mobility scripts, which are produced by the BonnMotion mobility generator [14]. It provides the simulator with Reference Point Group Mobility (RPGM), Random Waypoint Mobility (RWP), Manhattan Grid Model, and Gauss-Markov model, all further described in Section 3.1 and is directly built-in to the program. More information on mobility in HNS can be found in Section 4.5.8. The simulation parameters can be configured via the main configuration file that defines also basic values such as seeds, positions and ranges of base stations and access points, and traffic models. The configuration file is read in at the beginning of the simulation process and according to the traffic models, events are generated and stored in the event queue. The same thing happens for every mobility change. The rest of the information in the configuration script is stored as Java objects and attributes.
Before the simulation starts, the scenario can optionally also being built up graphically in a GUI. When the simulation is initialized, an Event Scheduler thread processes the events in the queue and synchronizes with the graphical interface to display movements and sessions in the GUI, which is also described in Section 4.5.7. For the evaluation, log files are generated at runtime. The main simulation and evaluation tasks are shown in

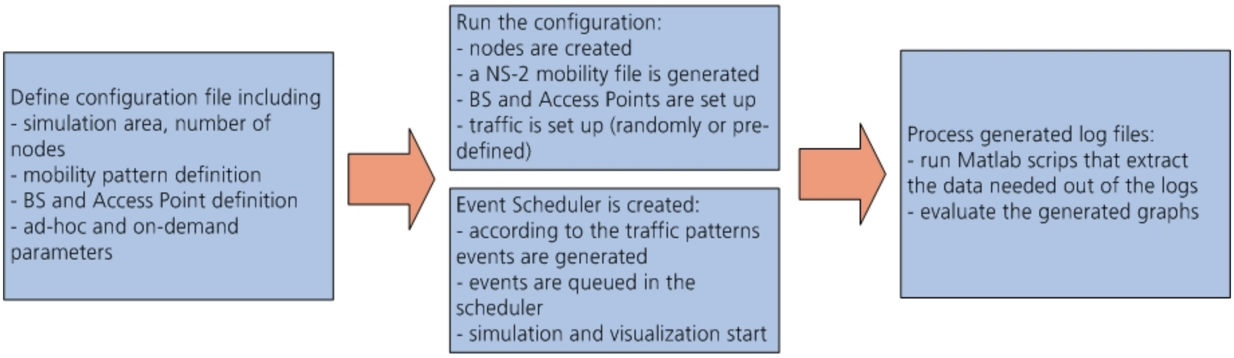

Figure 10: Simulation procedure with pre- and post-processes

Figure 10.
In the pre-processing stage, the entire network is configured offline, which is described more detailed in Section 4.5.3. From the moment the simulation has started, all the events are stored in log files in order to make the evaluation in a post-processing stage.

Figure 11 gives an overview of the package organization in the simulation program. The more detailed class diagram is presented in Figure 12. The general architecture splits the program into the packets view, control, util, and simulator. Their functions are also described in the Figure 11.

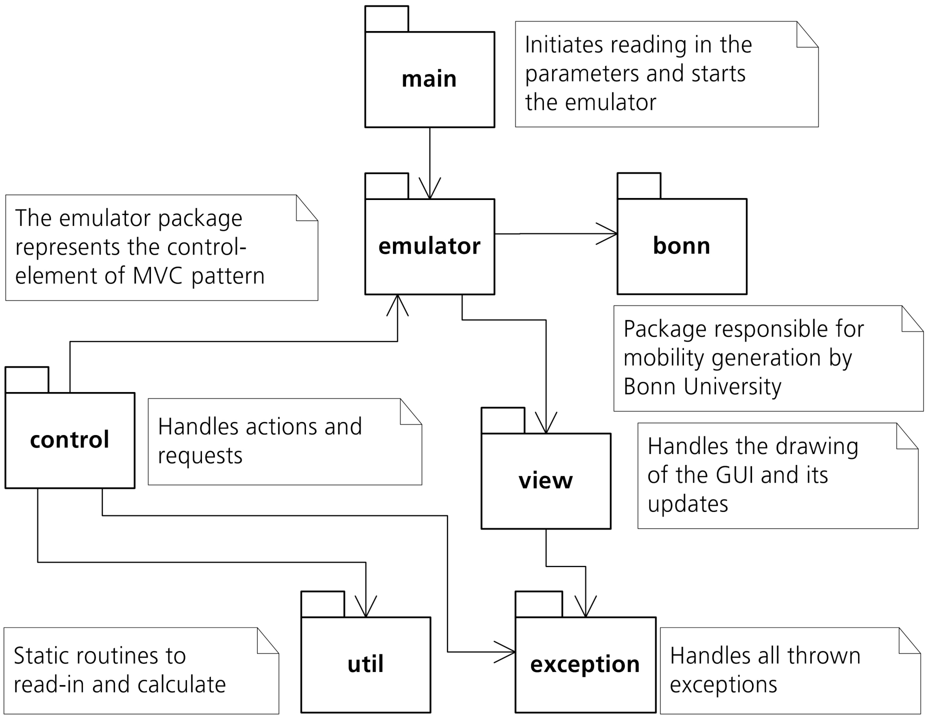

Figure 11: Overview of packages in HNS

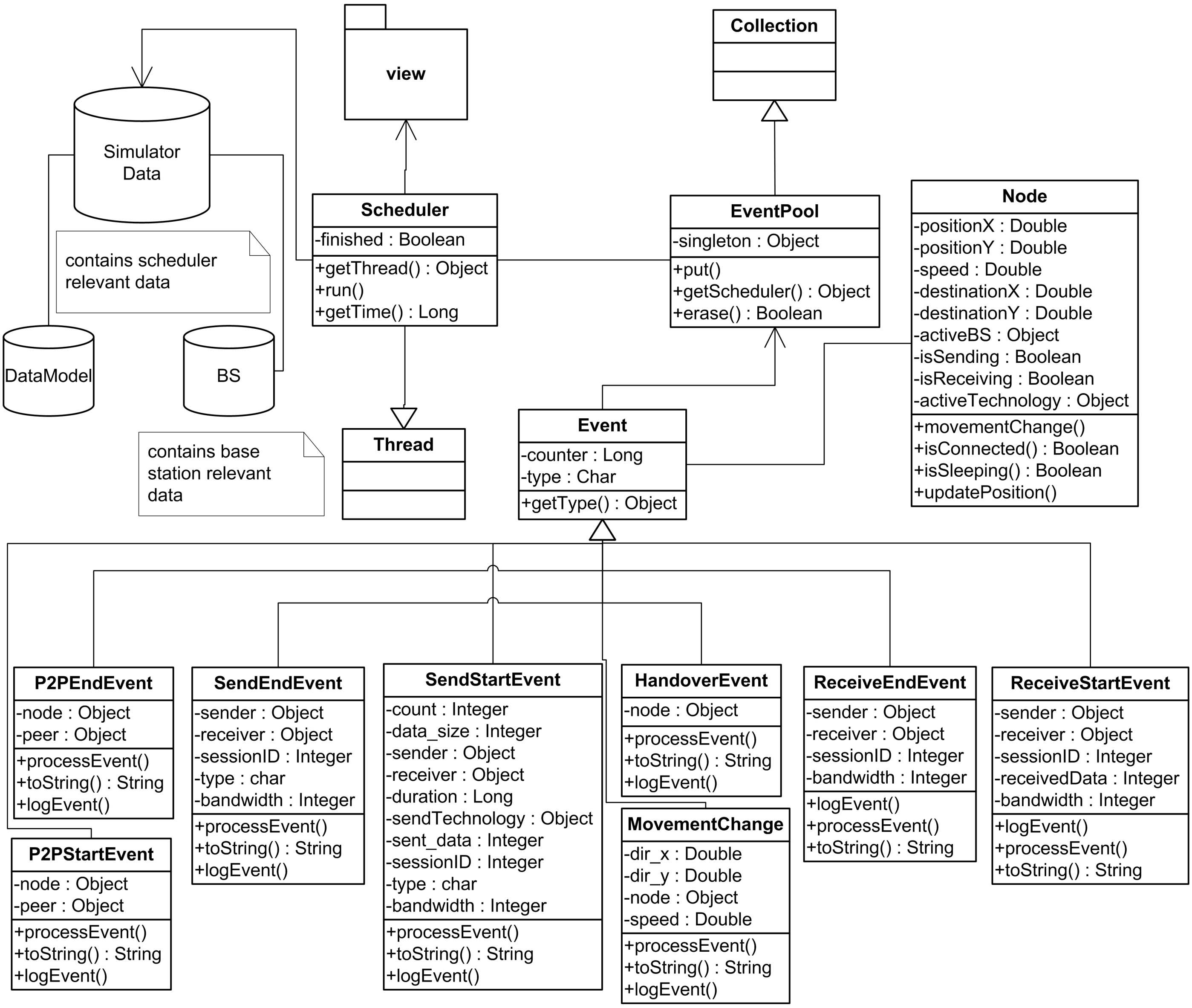

Figure 12: Simplified class diagram of HNS

4.5.2 Development process

After discarding the idea of using an existing simulator and extending it with the new features and capability of heterogeneity, the choice of an appropriate programming language is an important step for the success of the new simulator. As Matlab provides a huge library of useful functions, and many existing simulators are programmed in Matlab, it is the first choice. To evaluate the suitability of Matlab some mathematical skeleton is implemented to show the basic idea of what should be the final results of the simulations. This is documented in 5.2.1. This small experiment shows that Matlab is really useful to realize the mathematical aspects of network simulation. The only missing parts are a suitable graphical interface in order to make realtime graphics and some threading possibility. As the calculations to be made during simulations are not that complex, it is more convenient to choose a real high level programming language such as C++ or Java. The later is chosen since it fulfills many important requirements: Java is simple to extend with further modules, which could be necessary to extend the simulator later. Additionally it provides large graphic libraries as well as the possibility to implement threads, which is very useful for the performance of the simulator itself and also for the simultaneous visualization.
The very first implementation performs a simulation in two phases: A first run of the simulation without sessions to determine the handover points. In the second phase, the real simulation, where the handover points are put into the Event Scheduler as static occurrences. This turns out to be not dynamic enough because the handover of nodes does not only depend on available technologies and nodes, but also on the ongoing traffic. Therefore, beaconing is introduced and the pre-run of the simulation is taken away. Instead, during the simulation, nodes are detecting their environment in pre-defined intervals. This introduces additional overhead - as in real wireless networks - but in return the choice of the access network is a lot more accurate, involving knowledge of current load at available base stations.
The first version of the simulator is strictly time driven. A clock is running and increased by steps of milliseconds. Events in the pool, which are matching the current time are then performed, such as start sending, stop sending, change direction and speed. This time based type of simulation turns out to produce much overhead. For a scenario of 4000 seconds of duration, $4'000'000$ queries to the event pool have to be performed, many of which are unnecessary and just increasing the simulation time. This leads to the decision to make event based simulations. The benefits are shorter simulation times caused by higher performance of the simulator. The same scenario, which is calculated with Matlab is also simulated with the HNS with some small adaptions. The results are also presented in Section 5.2.2.

4.5.3 Configuration

All simulation parameters are specified in a configuration file that is read in by the simulator at the beginning of the simulation process. The three big parts that have to be defined are *mobility behavior*, *base station and access point locations* and *traffic generation*. For the mobility specification the syntax of BonnMotion [14] can be applied without changes. As the simulator accepts all mobility files that follow the NS-2 syntax, user-defined scenarios can be used, without following a specific mobility pattern. After reading in the mobility file, all the mobility changes are queued as events in the Event Scheduler.

```
MovementChange movementChange =
  new MovementChange(node, destX, destY, speed);
eventScheduler.put(Time.getCurrentTime(),movementChange);
```

The base stations and access points are defined in terms of location, technology and range in the configuration script and stored in a list of base station objects.

```
Vector baseStations = new Vector();
...
BaseStation baseStation =
  new BaseStation(xCoordinate, yCoordinate, radius, technology);
baseStations.add(baseStation);
//store the base stations in the data model
DataModel.setBasestations(areas);
```

To change propagation models or technology parameters, the Java code has to be adapted. The technologies are represented by classes which are sub-classes of a Technology class, defining the required parameters. This is described in more detail in Section 4.5.5.
Traffic can be set up either manually or by help of a random traffic generator that generates random sessions with a maximum of data to transmit and a maximum of peer-nodes according to a specified seed in order to make simulations reproducible. In case of manual traffic, sessions have to be defined in terms of data amount to send, a start time to begin the transmission and the peer nodes to send the data to. Before simulation start the traffic is generated and added to the event queue.

```
SendStartEvent sendEvent =
        new SendStartEvent(sendingNode,receivingNode,
        dataSize, EventType, uniqueSessionID());
eventScheduler.put(starTtime, sendEvent);
```

There is also a Constant Bit Rate (CBR) traffic generator implemented in HNS. Additionally, the ad-hoc and on-demand features can be enabled or disabled. For the handover decision three criteria can be specified in the configuration script. The network selection can be done either according to the best signal, the lowest costs or the best available bandwidth. The script language is not a particular one but the basic structure is taken from Qualnet. The parameters are read in and either stored in static classes holding basic values and configuration settings or they are processed directly and put into the Event Scheduler in case an event is triggered by the given parameter. An example script can be found in Appendix B.

4.5.4 The Event Scheduler

As the whole simulation is event-driven, a queuing system is necessary to schedule the queued events. The events in HNS are bundled by a Java TreeMap, which implements the

Collection interface. A TreeMap is a collection of value/object pairs which are sorted by their value. The most obvious thing is to fill up the TreeMap with time/event pairs, which then are sorted by time. This provides an efficient way for the scheduler to take the next event and to process it. The events are of different types and inherit from SchedulerEvent so they can be treated all in the same way by the scheduler. The first events at the beginning of each simulation are SendStartEvents and MovementChangeEvents, but the EventPool is normally growing during simulation because each send event triggers a ReceiveStartEvent on the other side, which is scheduled with a slight delay. When a handover occurs at either the sender or the receiver, the sending node stops sending for a short fixed time x and schedules a new SendStartEvent with the remaining data to send at time $t + x$. Thus, with every interrupting event (handover, peer-to-peer, movement change), the event queue is modified, i.e. events are rescheduled, deleted or added and the simulation proceeds. As already mentioned in Section 4.5.1, the event queue does not treat single packets but session streams.

Because of the multiple threads running at the same time in the simulator, the simulation is not time-correlated. One aspect to consider is the occurrence of multiple events at the same time. As this is not possible because of the obligatory uniqueness of values in collections, events occurring at the same time as events stored in the TreeMap have to be delayed by one millisecond. If this happens frequently, the simulation would be falsified. To control this effect, the scheduler throws a SimulatorException if too many events occur at the same time. A change of seed would help to resolve the problem; however, this case never happened until now.

4.5.5 Handover decisions and bandwidth allocation in HNS

In order to enable seamless connectivity in a heterogeneous network, some mechanism is necessary to decide about the right moment to handover to a neighboring or overlapping cell that best fits the user's needs. Of course, there are many factors, which possibly influence this decision as for example Quality of Service (QoS) issues, the strengths of available signals, costs and type of accounting,and load balancing among neighboring and overlapping cells. Thus, a handover decision is based on an algorithm, which weights the influencing factors according to the user's preferences with respect to the operator's settings. Obtainable services offering seamless connectivity are based on per-volume accounting strategies and thus, the mobile node is generally interested in having always a sufficient bandwidth whereas the operator's attention is lying on balancing the load and at the same time providing a sufficient bandwidth to the user, e.g. providing higher bandwidth to users of video streaming services and lower bandwidth to a user just checking e-mails periodically. Such considerations are not yet implemented in today's access networks; instead, all customers are treated in the same way. If this process is not controlled or influenced by the user, today it works as follows:

A mobile node is always assigned to the cell, which provides the best bandwidth. In the best case this is Wireless LAN. If no access point is in reach, UMTS or GPRS are chosen. The bandwidth sharing within the corresponding cells is not exactly the same for cellular networks as for Wireless LAN. A fair access point allocates timeslots equally among the competing nodes and the access point itself. As a rough approximation, the bandwidth for a node is then$\frac{m}{n}$, m being the total capacity offered by the access point and n being the number of competing resources. Hence, the full bandwidth provided by the access point can be exhausted. This is not always the case for UMTS. Although a common base station is able to supply 1 Mbps at maximum, a single user can only profit from at most 384 Kbps, which corresponds to a third of the maximum bandwidth. For n users and a maximum capacity of m, the upload bandwidth (bw$_{up}$) and the download bandwidth (bw$_{down}$) for one mobile node can be calculated as follows:

$$bw_{up} = \begin{cases} 64, & \frac{m}{n} \geq 64 \\ 0, & otherwise \end{cases}, \quad bw_{down} = \begin{cases} 384, & \frac{m}{n} \geq 384 \\ 128, & 384 > \frac{m}{n} \geq 128 \\ 64, & 128 > \frac{m}{n} \geq 64 \\ 0, & otherwise \end{cases}$$

In GPRS, both the nodes and the base stations are commonly using coding scheme 4 (CS-4), offering theoretically 21.4 Kbps for each slot. Whenever a node enters a GPRS cell which has more than one up- and one downlink TDMA slot available, the node can acquire up to 2 uplinks and up to 4 downlink slots, resulting in a maximum upload data rate of 42.8 Kbps and a downlink data rate of 85.6 Kbps. These values may vary from operator to operator depending on the vendors. In a real GPRS network, these values are considerably lower, because the CS-4 can only be used for nodes being very close to the base station. More realistically, nodes use the CS-3, delivering actual speeds of about 14.4 Kbps per slot. The bandwidth allocation in GPRS is similar to the one in UMTS regarding the discrete behavior. The maximum number of users is denoted by x whereas the maximum cell capacity is once again represented by m:

$$bw_{up} = \begin{cases} 42.8, & \frac{m}{x} \geq 42.8 \\ 21.4, & 42.8 > \frac{m}{x} \geq 21.4 \\ 0, & otherwise \end{cases}, \quad bw_{down} = \begin{cases} 85.6, & \frac{m}{x} \geq 85.6 \\ 42.8, & 85.6 > \frac{m}{x} \geq 42.8 \\ 21.4, & 42.8 > \frac{m}{x} \geq 21.4 \\ 0, & otherwise \end{cases}$$

Considering these theoretical bandwidth values it can be stated that UMTS does not always deliver better bit rates than GPRS, therefore, for the simulations, the technology providing the best bandwidth at a given moment is chosen to make the handover decisions. For the bandwidth allocation, the simulator implements the presented models for the three technologies UMTS, GPRS and WLAN. The Java classes for the different access network technologies implement mainly two functions: one to get the current upload for a specific node at a specific base station. The code can be found in the Appendix in Sections C.1 to C.3.

4.5.6 Mobile nodes

In the simulator, mobile nodes are uniquely identified by their identifier. Further important attributes are position, destination and speed. The ability to specify device-dependent properties for each mobile node is generally implemented but not used in the simulations described later in this work. A node's position is updated through a function each time an event occurs. The new position can easily be determined through the current position, the direction and the time of the last update. Beaconing lets the node detect base stations, access points and one-hop peers. Before a session initiation to another node it is verified if the destination node is a one-hop-peer. In the positive case the session is established directly with the peer node. Otherwise the infrastructure network to which the node is currently attached is used as default gateway. All the nodes are stored in a global vector to provide the ability to search for and modify them at any time.

4.5.7 GUI

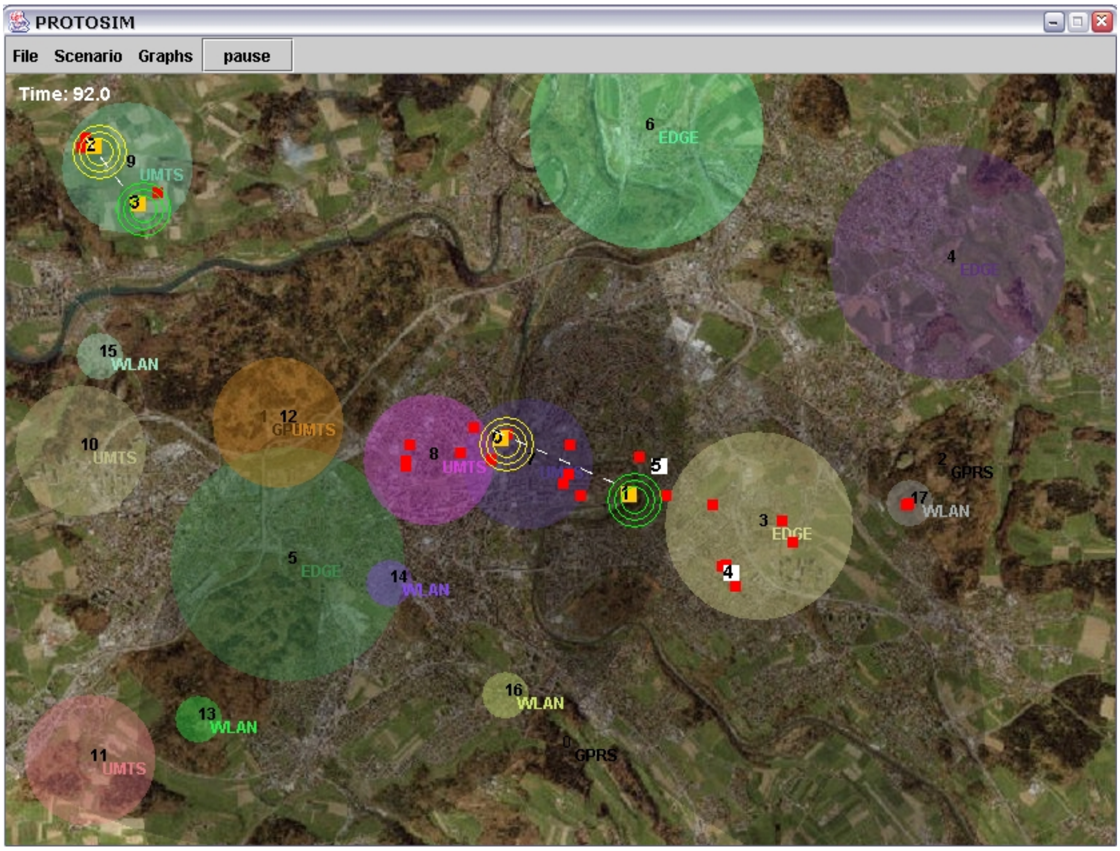

Figure 13: The Graphical User Interface of HNS

The GUI visualizes what is going on during the simulation. It is a part of the implementation from the beginning and also helps to debug the simulator in case of visible misbehavior of the network.
The GUI shows the points of interest: First of all, the base stations and access points are represented with technology, location and range. A map can be loaded as background image to visualize the geographical topology. The nodes' movements are shown at runtime but as the simulator is event driven, the nodes do not always move continuously, especially if there is no event for a long time. The logical connections between nodes are indicated by straight dashed lines between them and also sending and receiving activity can be observed. Handover points are shown as red dots. A screenshot of the GUI is shown in Figure 13.

For larger or multiple simulations it does not make sense to visualize the simulation process. For this reason there is also the possibility to call a batch job on a folder structure which then processes every configuration file in the folder and stores all results without any graphical output.

4.5.8 Mobility

As mentioned in previous sections, the whole mobility generation is done by the application of the Bonn University in Germany [14]. Their Java code provides mobility file generation for NS-2 for the previously described models. Since it would be quite boring and also fault-prone to always produce mobility files first and then specify in the basic configuration where they are located in the system, the BonnMobility code is fully integrated into the simulator and the mobility behavior can be specified in the configuration file. A pre-process in the program then takes every line of the generated NS-2 mobility file and defines the appropriate events which afterwards are fed into the Event Scheduler. At runtime, whenever a change of direction occurs, the new values for direction and speed are adapted by the mobile node. The positions of all the nodes are updated at every event. Whenever nodes are in sight of each other, i.e. when they are closer to each other than a random value in a specific interval, ad-hoc connections are established spontaneously.

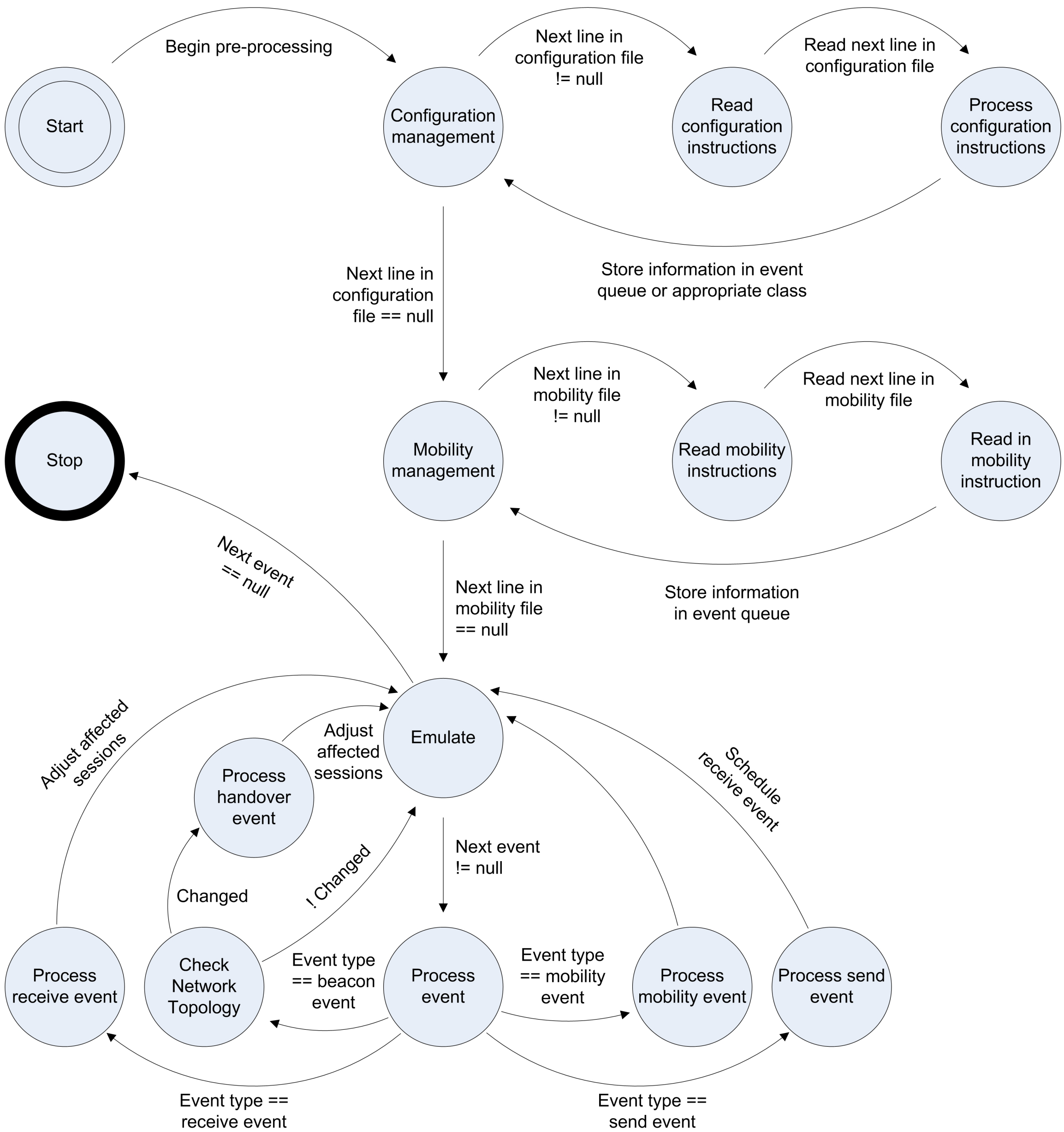

Figure 14: State Diagram of HNS

If a handover is indicated during a transmission the flow is adapted to the new attachment-point's technology. Possible handover possibilities can be detected through beaconing. If another base station or access point fits the handover decision criteria the new binding is established and the old one released. To prevent an overhead through handovers, this is only done if the last handover is more than a few seconds old.
Sections 4.5.3 to 4.5.8 are also reflected in the state diagram in Figure 14. The states *Configuration management* and *Mobility management* represent the pre-processing phase. Once the configuration and mobility files have been read in, the *Simulate* state is reached and the events in the queue are processed as long as the Event Scheduler is not empty.

4.6 Evaluation and Verification

In order to verify the simulations there are different mechanisms of control. Some values have to be observed during simulation, others have to be verified based on the log files at the end of simulation. The mobility part is not verified as the code is taken from the University of Bonn without changes. First of all, it is interesting to observe how the Heterogeneous Network Simulator behaves compared to NS-2. In the early stage of development some basic connectivity measurements are done with two similar setups and exactly the same mobility file for both, the newly built simulator and NS-2. The scenario is based on an area of 5000 by 5000 meters, on which different grids of access points are laid out reaching from a topology of 4 by 4 (as shown in Figure 15) to 8 by 8 nodes.
As all access points have a range of 250 meters, the five topologies with different node densities correspond to total area coverages of 12.6 to 50%, which is represented by the dark blue line in Figure 16. Ten mobile nodes, which represent five communicating pairs, are traveling the simulation area according to the Random Waypoint Mobility script. These pairs are observed for 800 seconds and the single node's total on-time, i.e. the time they have been in the range of an access point as well as the theoretical connectivity between the pairs of nodes are measured. The decision whether a node has connection or not is based on connectivity. Whereas in HNS the antenna pattern of all wireless technologies is circular, in NS-2 the access points have antenna patterns, which are not absolutely circular and interference is influencing connectivity. The range of the access point is for both set on 250 meters. The connectivity of HNS is based on pure geometrics, which has been verified by a simple test with custom mobility, where a mobile node had to travel a straight line with exactly 50% of connectivity from a geometrical point of view. With HNS precisely the same percentage was measured, which leads to the conclusion that the model works as designed. The results for the connectivity comparison for NS-2 and HNS are shown in Figure16.

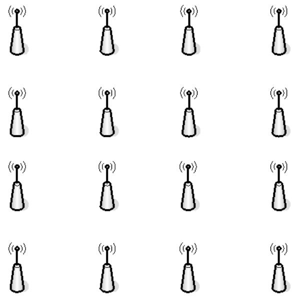

Figure 15: Example grid for 4 by 4 nodes

The on-time graphs of the single nodes for both, HNS and NS-2 correlate very much with respect to the total coverage values, although NS-2 shows slightly lower on-time values than HNS. This surely results from the physical model of NS-2 which is not mathematically calculated but simulated with more real-world aspects such as interference and multi-propagation. The behavior of the connectivity graphs can be interpreted in a similar way.

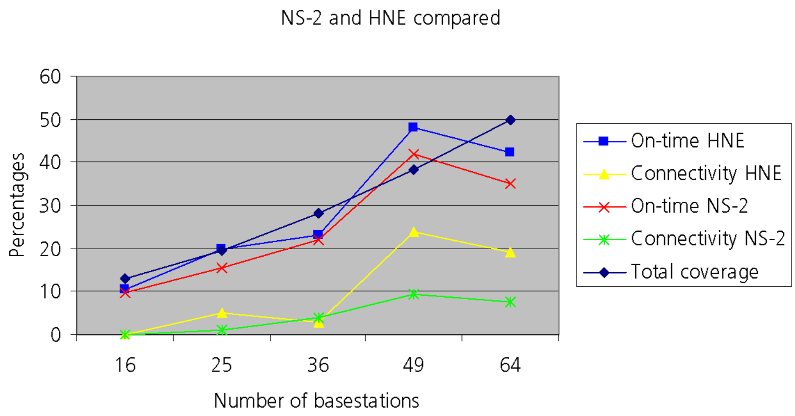

Figure 16: Measured connectivity of NS-2 and HNS compared

Whereas the connectivity of the node pairs in HNS is one fourth of the on-time of the nodes, the same nodes in NS-2 are only connected during one sixth of the total on-time. This low connectivity of only 7.6% in the NS-2 simulation scenario with 50% coverage is caused by the straight Random Waypoint paths which sometimes only touch the borders of access points on a whole segment. This leads to long distances without connection, which can be observed in Figure17. It depicts the 7 by 7 access points scenario. The paths of the ten nodes are shown by straight lines, which in many cases do not cross the range of access points for a long distance. The lover connectivity in NS-2 is surely caused by the more realistic propagation models as well as interference, which also explains the low value for 64 base stations.

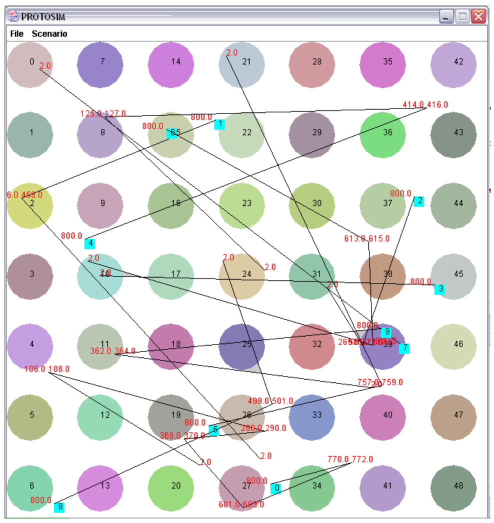

Figure 17: Example of test scenarios

For test reasons, the way of the mobile nodes is documented with straight lines as well as time indications at each mobility change. The colored circular areas refer to infrastructure-based networks, whereas the intermediate white areas depict network wholes.
To check the right behavior of the simulator's different technology models and handover strategies, the same simple setup with sparse node density, simple traffic pattern but this time with a custom mobility pattern is simulated. After a short simulation time, all the output is recalculated by hand and then compared to the output generated by the simulator. To validate the right behavior in bigger simulations, the state of the simulator is examined as often as possible to detect misbehavior triggered by failures in threads, missing synchronization, and other unexpected behavior.
The collection of the results is done at runtime. All events and many processes produce a log, which is written into several specific files. There are three main types of log files described in the following sections.

4.6.1 Logging of events

The log entries consist of the following values:

Event type	Session ID	Sender	Rec	Start time	End time	Thp	Techno logy	Dura-tion	Next HO	Data sent

Event type:	can be *send start*, *send end*, *receive start*, *movement change*, *ad-hoc session beginning* and *ad-hoc session termination*
Session ID:	uniquely identifies the session
Sender:	unique node number of the sender
Rec:	unique node number of the receiver
Start time:	start time of the event in milliseconds
End time:	end time of the event in milliseconds (if applicable)
Thp:	specifies the throughput in kbps
Technology:	refers to the used network technology. It can also be $NoConn$ if there is no technology available or $P2P$ if it is an ad-hoc session
Duration:	indicates the length of the session in milliseconds (if applicable)
Next HO:	specifies if a handover will occur during the current transmission
Data sent:	represents the amount of bytes that is transmitted (if applicable)

Obviously, not all the fields are defined for all the events. Many values can be extracted out of this log to show the throughput over time, the duration of the sessions, the activity of the nodes, etc.

4.6.2 Logging of base stations and access points

Every time a node attaches to a base station or access point or disconnects from it, a log is generated with the following fields:

Time	BS ID	BS range	# of users	Event type	Node numb

Time: specifies the current time in milliseconds
BS ID: uniquely specifies the node's attachment or detachment point
BS range: shows the size of the cell in meters
# of users: number of users attached to a base station
Event type: determines whether it is a register or a sign-off request
Node #: specifies the node registering at or signing off the BS

For a more global evaluation of the activities at the base stations and access points, the total load of the base station is calculated at the end of the simulation.

4.6.3 Logging of nodes

To quantify end to end connectivity, the logging of some node specific behavior is necessary:

Node #	Total ms	Total bytes	Ms of outage	outage count	Ms of delay	delay count	P2P ms	P2P bytes	Tech1 ms	Tech2 ms

Node #: specifies the node which is logged
Total ms: summarizes the time spent in sending
Total bytes: summarizes the bytes sent in total
Ms of outage: specifies the total time, the node is blocked while sending, i.e. a connection breaks during transmission and is continued later
Outage count: counts the times of connection outage
Ms of delay: summarizes the total time the node has to wait before the transmission begins
Delay count: counts the number of times the node has been delayed
P2P ms: summarizes the time the node has been transmitting infrastructure-less
P2P bytes: summarizes the total bytes sent infrastructure-less
Tech1 ms: specifies the total time the node has been using technology one (e.g. WLAN, GPRS, ..)
Tech2 ms: specifies the total time the node has been using technology two (e.g. WLAN, GPRS, ..)

4.7 Analysis and Presentation of Results

4.7.1 Evaluation relevant measurements

The whole post-processing of the log files is done mainly with Matlab, Excel, or GNUplot. The logged data is processed in Matlab and several scripts help to extract the desired data and generate graphs and statistical values. The data analyzed in simulations are the following:

On-time evaluation The logging of active periods of a node is important to determine the benefit of the on-demand mode. This information can be taken out of the base station log file.

Base station activity The log file of the base station activity holds enough information to produce a graph that shows the registered nodes at each base station at any time of the simulation.

Mean throughput per session Iterating on the log file of the events and summarizing the throughputs during the subsessions results in a graph which shows the mean throughput for every session.

Energy consumption during a session The energy consumption during a session is estimated based on empirical values (see Appendix A) and literature references. The logs of on-time is used calculate the energy consumption of a node at the end of simulation

Mean duration of the sessions An important performance indication is the session duration. The session duration is an important factor for the user's experience of a fast and seamless session.

Delay One measurable criteria of the quality of end to end connectivity is the delay. It is the duration a node is waiting to initiate a session but is blocked due to lack of resources in the network.

Block rate The block rate represents the occurrence of blocked session starts.

Outage Another indicator for the quality of a session is the time a node's communication is interrupted after the session has already been started.

Drop rate The drop rate represents the occurrence of interrupted sessions.

Efficiency The indicator for efficiency is defined to be the total amount of data sent over a peer-to-peer connection divided by the amount of bytes totally sent. This quantifies the impact of additional ad-hoc connections compared to transmissions over infrastructure-based technologies only.

Network load The network load determines the overall load of the network.

4.7.2 Evaluation of results

The development of the simulator happens in different phases and results in several intermediate states of evolution. During this time, not only the code evolves but also the way to present the findings and to process the logged data into graphs. The manner of collecting, calculating and representing results is refined steadily for the final version.

From the beginning, the logging process is very much focused on sessions and many important aspects in order to qualify heterogeneous scenarios are developed during testing.
The evaluation of some intermediate results presented in Section 5.2.3 shows some important drawbacks. The session based simulation introduces some difficulties: It has to be avoided that many sessions do not finish within the simulation time. This could happen because of the normal distribution of session starts with too large data amounts. The measured values could therefore be misleading. The second big issue is the representation of the results. Although the simulations are session based and much data is logged per node, there is no evident reason to show the results per session or node. However, this type of graphs does make sense for prototype evaluation and testing purposes because processes can easily be tracked and irregularities can be found.

5 Simulation and results

5.1 Introduction

The evaluation is split up in three parts. The first section collects some important simulations and results, which are performed during the design and implementation process in chronological order. It documents the intermediate steps, which have an influence on the final simulator. The second evaluation has a focus on the users scope and on the potential benefit of the ad-hoc and on-demand mode without being limited to specific communication technologies. Like the second part, the third part is also simulated with the final version of HNS. It investigates on both, user and operator scope in existing networking environments. The degree of improvement is measured in terms of transmission characteristics, such as for example throughput and interruptions. Another intend of the evaluation is to measure the energy consumption of the nodes in case of the enabled on-demand and ad-hoc features. The two last sections are further separately treated in terms of user and operator benefits.

User benefit resulting from ad-hoc and on-demand mode
Users in wireless network environments are interested in reliable links providing good quality in terms of few interruptions and sufficient throughput. In addition, customer satisfaction has also beneficial impact for the operator.

Operator benefit resulting from ad-hoc and on-demand mode
An operator is primary interested in optimal resource management. As the revenue for data transfer is decreasing further because the billing is not organized by volume anymore, the operator is interested in offloading traffic to ad-hoc connections. Additionally, the operator benefits from load balancing to be able to serve as much users as possible with the fewest possible load on the infrastructure.

5.2 Simulation results obtained from prestudies and prototyping

5.2.1 Mathematical Approach

The first approach was done in Matlab, as mentioned in Section 4.5.2. The experiment is based on a simple real world example, which is shown in Figure 18, a formula is set up in order to quantify the benefits of the example scenario. Below the scenario diagram, a graph is showing the data volumes to be transmitted via peer-to-peer (blue) and over the infrastructure (red).

Scenario Alice and Bob meet at the train station in the morning and configure their laptops for a peer-to-peer communication. Bob wants to share the video recording of a business presentation with Alice. The amount of data to be transferred is 1.8 GB. As they board the train they have successfully established the connection and Alice begins the download. When they arrive at the destination station about six minutes later, the video transfer has not finished yet. But Bob and Alice do not have the same destination and therefore their peer-to-peer connection breaks. In the case of a seamless handover to another technology they could just stay connected to continue the download. When they meet again for lunch four hours later they re-establish the peer-to-peer connection and finish the transmission.

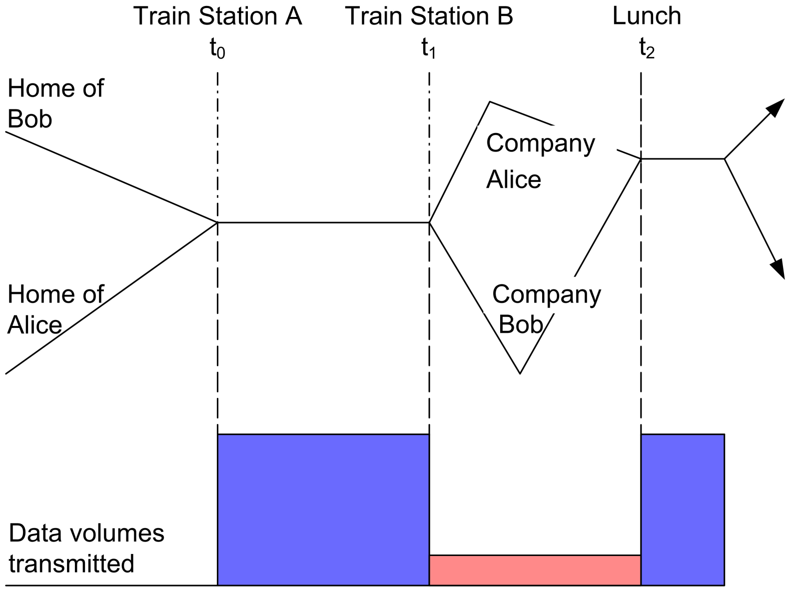

Figure 18: Business scenario calculation

The results presented in Table 2 show the benefits resulting from the seamless handover to an infrastructure-based technology when the peer-to-peer connection can no longer be maintained.

- Fixed parameters:
  - Time they meet at train station: $t_0 := 0s$
  - Time they leave the train and lose P2P connectivity: $t_1 := 400s = 6m\ 40s$
  - Time they meet again for lunch: $t_2 := 14400s = 4h$
  - Data to transmit: 1.8GB
- Neglected parameters:
  - Duration of establishing a peer-to-peer connection by hand
  - Duration of handovers
  - Retransmitted data due to the abrupt loss of the peer-to-peer signal in the case of non-seamless networking

The results prove that with the facility of seamless handover, at least 140 seconds can be saved, for the case that only GPRS is disposable. If there is even UMTS available, one could save 840 seconds, which equals 14 minutes. The most significant savings can be achieved by seamless handover to WLAN, which results in 1 hour 29 minutes less transmission time. In a second experiment, t_2 (time when Alice and Bob meet again for lunch) is incremented steadily to find the point of saturation mathematically, i.e. the point where the transaction is finished within the period with connection via infrastructure. The results are presented

Technology	Bandwidth	Duration without HO	Duration with HO	Improvement
P2P	5 Mbps			
WLAN	2.5 Mbps	16880 s	5360 s	68.25%
UMTS	0.3 Mbps	16880 s	16040 s	4.98%
GPRS	0.05 Mbps	16880 s	16740 s	0.83%

Table 2: Comparison of different handover technologies

in Figures 19 to 21. The bars depict the duration for the whole session in case of peer to peer only (white bars) and also with seamless handover to an infrastructure network (black bars). These are GPRS in Figure 19, UMTS in Figure 20, and WLAN in Figure 21. Apparently, in all three graphs there is always an improvement when performing the handover to an infrastructure-based technology. For the cellular technologies the benefit is rather little for small values of the interval. For all options of access technologies the black bars do stagnate at a certain point, i.e. from the point when the session ends within the interval [t_1, t_2].

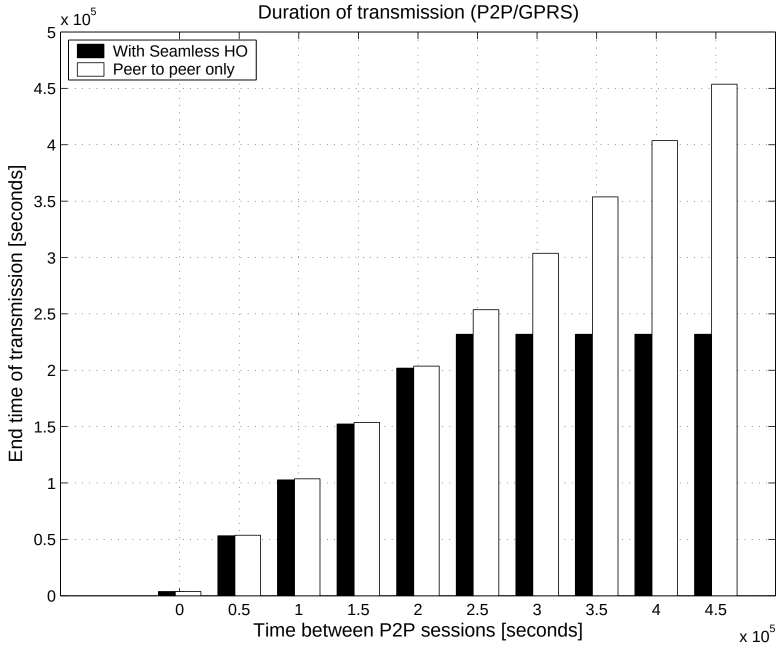

Figure 19: Seamless handover to GPRS versus P2P only

Obviously the duration to finish the transmission with GPRS is rather long with 250'000 seconds. This implies that after these almost 70 hours between t_1 and t_2 the transmission of the 1.8 GB has finished and it is not beneficial any more to switch to a peer-to-peer connection again. With UMTS, this critical point where a handover to peer-to-peer becomes useless is 40'000 seconds or 11 hours. Thus, UMTS is six times faster than GPRS. With WLAN, the critical point is reached at one hour already. This signifies that for this scenario,

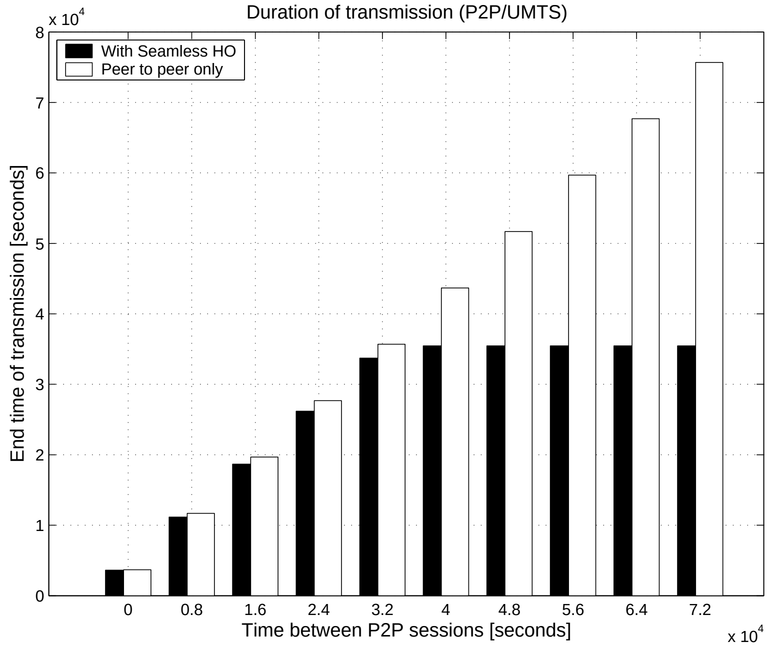

Figure 20: Seamless handover to UMTS versus P2P only

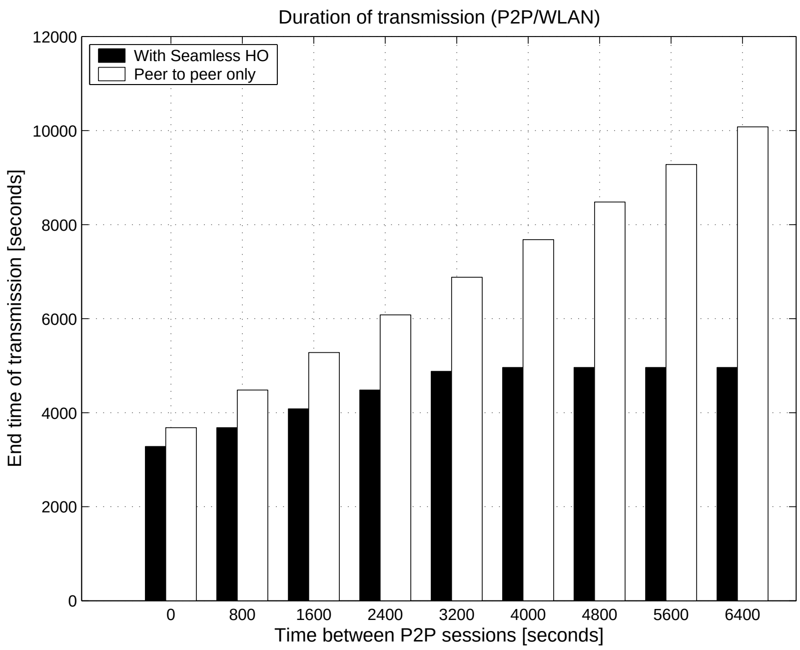

Figure 21: Seamless handover to WLAN versus P2P only

a seamless handover to WLAN after the first period of peer-to-peer would introduce great benefit and shorten the overall transmission time.

5.2.2 First HNS Approach

The scenario from Figure 18, which is calculated with Matlab is also simulated with the HNS with some small adaptions. The adapted version is once again presented in Figure 22.

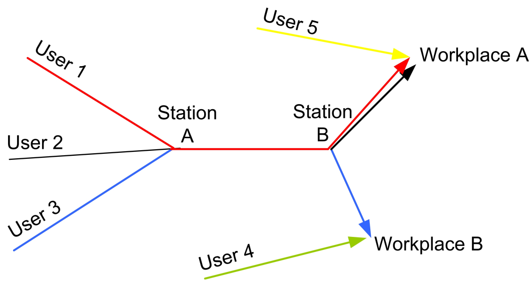

Figure 22: Scenario with five users

The detailed setup is described in the next paragraph and shown in Figure 23, which visualizes the scheme of Figure 22 in a map.

Scenario Five users go to work by train and bus. Three of them start in the upper right corner and go to the train station (Station A), where they all take the same train to the city center (Station B). There, two of them are heading north-west to their Workplace A (yellow square at the left) while the third is going almost straight to the north to his own Workplace B. Another user is living in the lower right corner, where he takes the bus to the Workplace B directly and the fifth is starting in the left center of the city where he takes the bus toward south-east (Workplace A). There are two regions covered with EDGE, one in the north-east and one in the city center, six UMTS base stations distributed randomly as well as four WLAN access points in the center of the city. The rest of the simulation area is covered with GPRS. The assumed bandwidths for these technologies are shown in Table 3.

Technology	Bandwidth
P2P	5 Mbps
WLAN	2.5 Mbps
UMTS	0.3 Mbps
GPRS	0.05 Mbps
EDGE	0.15 Mbps

Table 3: Bandwidths of technologies

The yellow points on the map are marks for the final destinations of the nodes, the red squares denote the handover points for either soft or hard handovers, or handovers from or to direct peer-to-peer links.
The traffic is set up randomly with a certain seed, so that the simulation is reproducible: Every node chooses a random number of peers in the range [1, total nodes], a smaller

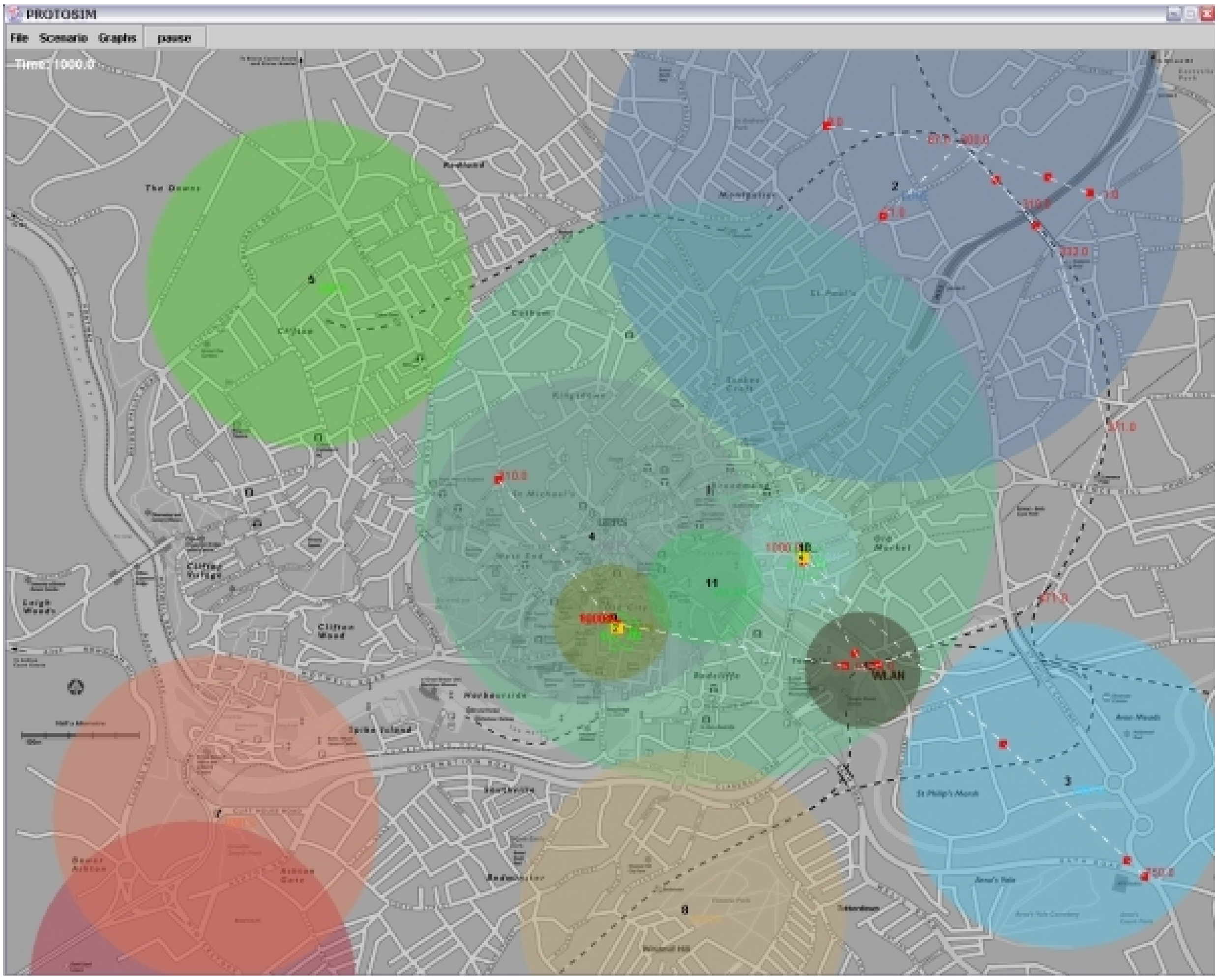

Figure 23: Scenario with 5 users and a random number of sessions

number of peers are chosen with bigger probability. The following code represents the selection of communication peers:

```
int numbPeers = 0;
double randD = rand.nextDouble();
for (double j=1; j<nodes.size(); j++){
  if(randD < 1.0/j)
    numbPeers++;
}
```

Then every node selects a start time for the traffic in the interval [1, simTime/2]:

```
randD = rand.nextDouble();
long startTime = Math.round((simtime/2.0)*randD*1000);
```

The amount of data is randomly chosen in the range [1KB, 0.5GB]. This data is sent to every chosen peer at the decided start time:

```
randD = rand.nextDouble();
int dataAmount = (int) Math.round(randD * 500000.0);
```

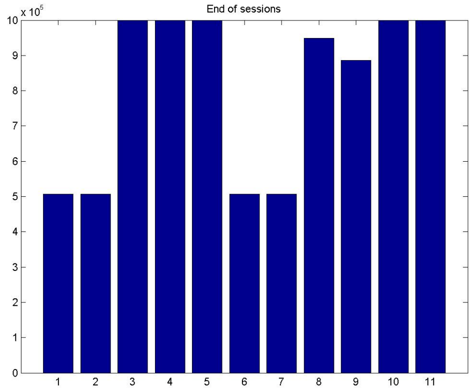

Figure 24: Scenario without handover to ad-hoc connections

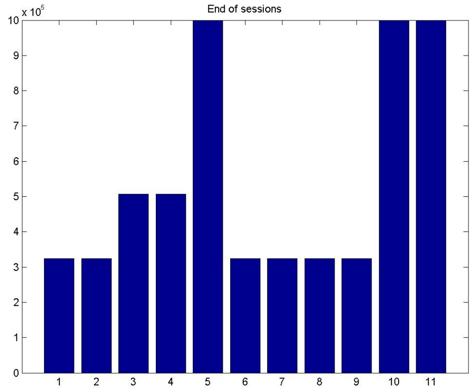

Figure 25: Scenario with handover to ad-hoc connections

The Figures 24 and 25 show the durations of the eleven randomly chosen sessions for this example in milliseconds. In the case of only infrastructure-based connections there are five sessions which do not end until the end of simulation. In the case of seamless handovers to ad-hoc connections whenever nodes get close enough to build up a peer-to-peer ad-hoc link there are only three sessions, which do not terminate before simulation ends. The transmission time for the randomly chosen sessions almost halves from 638.3 seconds in average to 367.5 seconds in average.

5.2.3 Intermediate simulation results to evaluate the implementation of HNS

The following results of an intermediate simulation series show some basic processes. The graphs are qualified both, formally (explanatory power) and technically in the following.
In order to measure throughput, on-time, load at the base stations, sent data, etc. with enabled and disabled ad-hoc and on-demand features in heterogeneous networks, the four scenarios are built up:

- Test Case 1: no ad-hoc mode, RGPM
- Test Case 2: ad-hoc mode, RGPM
- Test Case 3: no ad-hoc mode, RWP
- Test Case 4: ad-hoc mode, RWP

For the basic setup parameters that do not change over all the test cases we fix the simulation area to 3000 by 3000 meters, which in this case should represent a city center. Into this field

30 mobile nodes are placed. The data traffic is set up per node randomly with the same seed for all the test cases. For the chosen seed this results in 41 sessions being initiated in the 800 seconds of simulation at random start time. There are three WLAN access points with 250 meters of radius, two UMTS base stations with 700 meters of radius, and two GPRS base stations with 1500 meters of radius distributed over the simulation area. For the mobility models there is chosen a minimum speed of 8m/s and a maximum speed of 25m/s with a pause time of 20s. The intension here is to simulate public transport in cities. Reference Point Group Mobility (RPGM) and Random Waypoint Mobility (RWP) are chosen as mobility models. The parameters for each of them remain the same in the different scenarios. The definition of RGPM requires a value for the mean group size, which here is 6 and a standard deviation for this value, which we set to 3. The maximum distance to the group center is 55m meters in our simulations, which assures ad-hoc communications within the group.
First of all, the throughput is measured per session. This is shown in Figures 26 to 29.
The bars are indicating the mean throughput per session. The overall mean is shown by the straight line. In the case of group mobility the throughput for ad-hoc enabled traffic is almost five times higher than for ad-hoc disabled traffic. The RWP mobility model has a lower impact on the throughput; it could only be increased by 1.5 times for the ad-hoc enabled scenario. These immense differences between ad-hoc and non ad-hoc scenarios result from the relatively big group sizes and thus the chance of communicating with someone in peer-to-peer range. This form of evaluation of the throughput has the advantage that it gives an overview and at the same time contains some degree of detail. However, these graphs have some important disadvantages. First of all a direct comparison of the four scenarios is not convenient. It would make more sense to discard some level of session detail and to directly put the results in the big context, e.g. all four throughput averages in the same graph.
Another measurement value is the activity of the nodes in on-demand mode. Normal nodes are active over long periods, whereas in on-demand mode they can save energy when there is no connection. The logs ofr this aspect is done per node and also the evaluation is initially done per node. The result is shown in Figures 30 to 33.
On the one hand, this way of presentation of the on-time of the nodes gives a good overview of the simulation. It is clearly visible that the session starts are normally distributed and that there are numerous sessions, which do not finish. For just one session per node, like in this example, it is not suitable to have such representation of the results. However, for scenarios with multiple sessions per node it might be be interesting to see the activity of the nodes together with some information about the used access network. Additionally, there is a second fact to keep in mind when making simulations based on sessions: Almost one third of the initiated sessions were not able to terminate within the simulation period. For the evaluation of node activity as well as the duration of transmission, the obtained results would be more relevant if every session would end before the simulation ends. Therefore, in the final simulations, the parameters for session start times together with the data amounts to send are chosen in a way, which allows every session to finish if the coverage is 100%, like here.

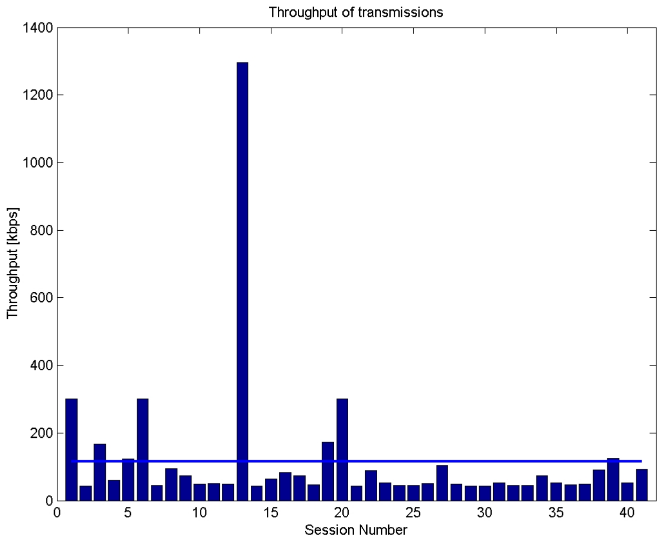

Figure 26: Test case 1: RPGM, no ad-hoc

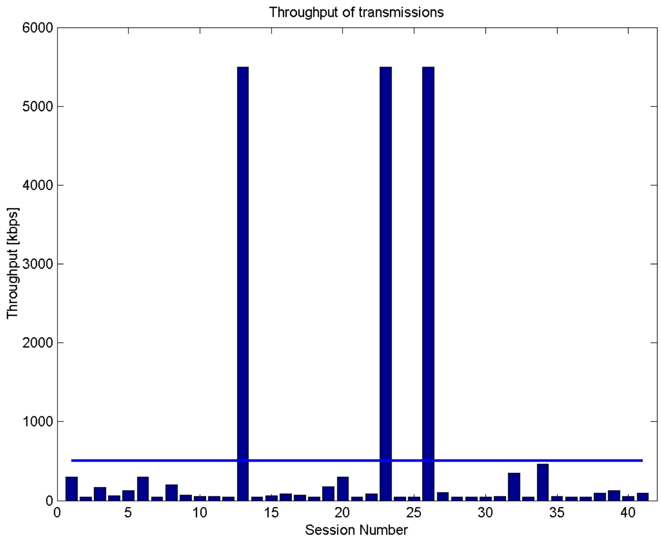

Figure 27: Test case 2: RPGM, ad-hoc

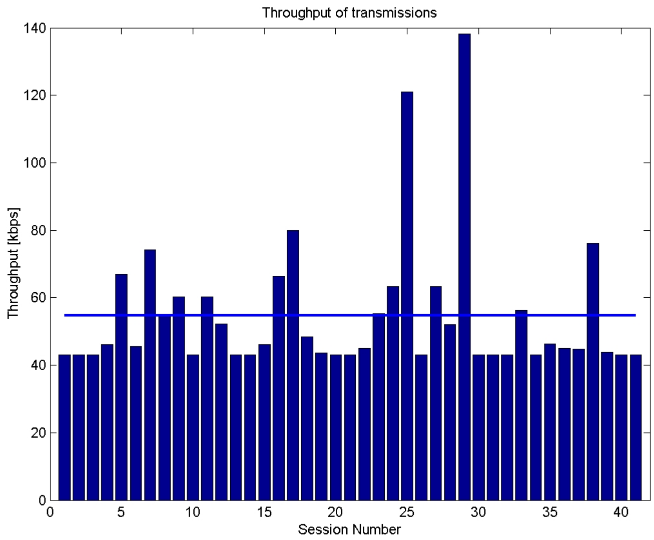

Figure 28: Test case 3: RWP, no ad-hoc

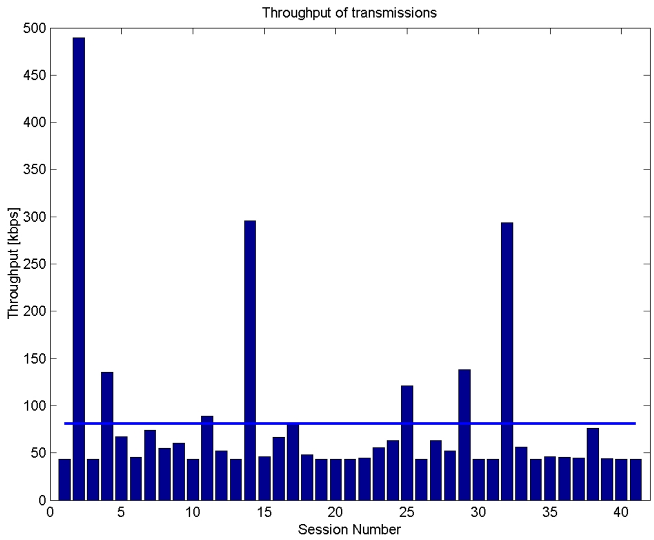

Figure 29: Test case 4: RWP, ad-hoc

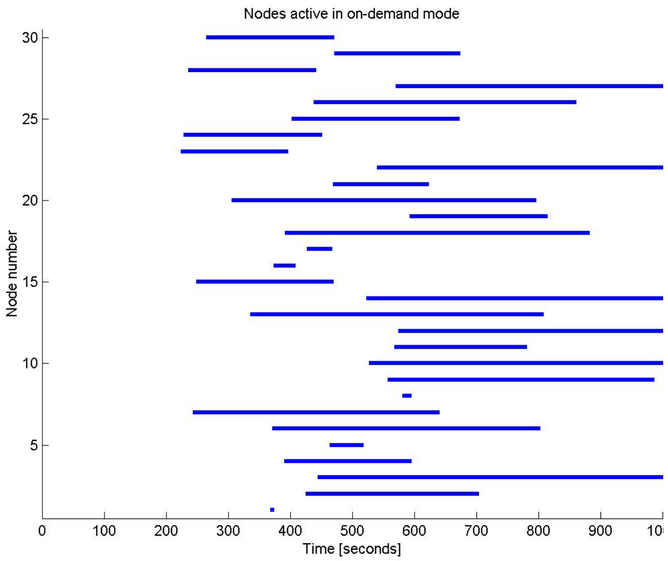

Figure 30: Test case 1: RPGM, no ad-hoc

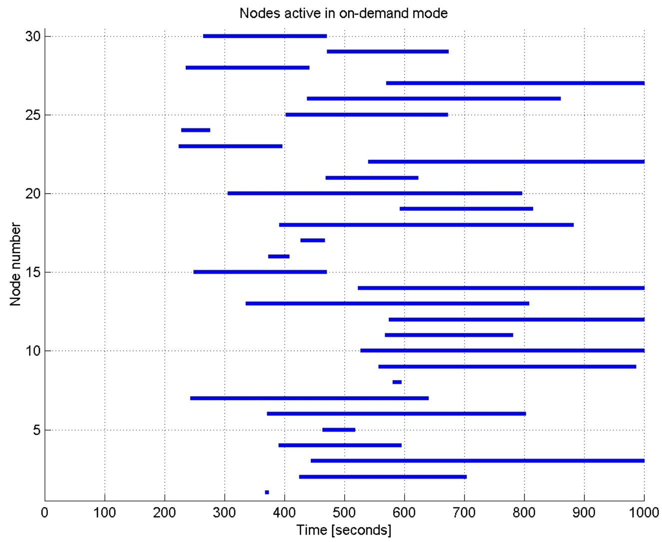

Figure 31: Test case 2: RPGM, ad-hoc

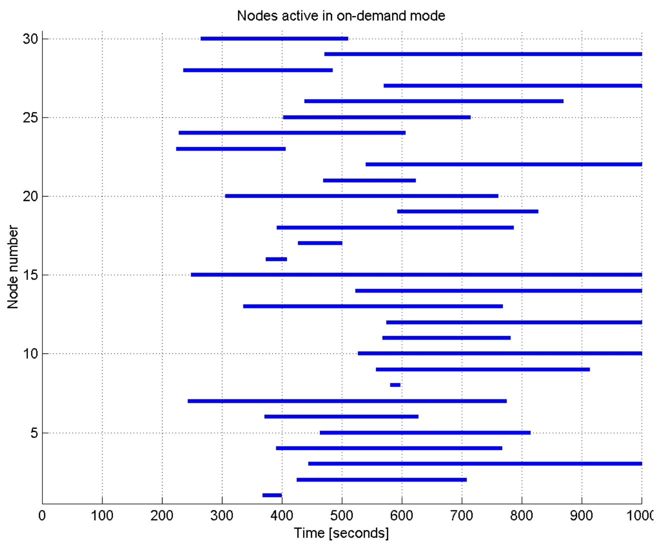

Figure 32: Test case 3: RWP, no ad-hoc

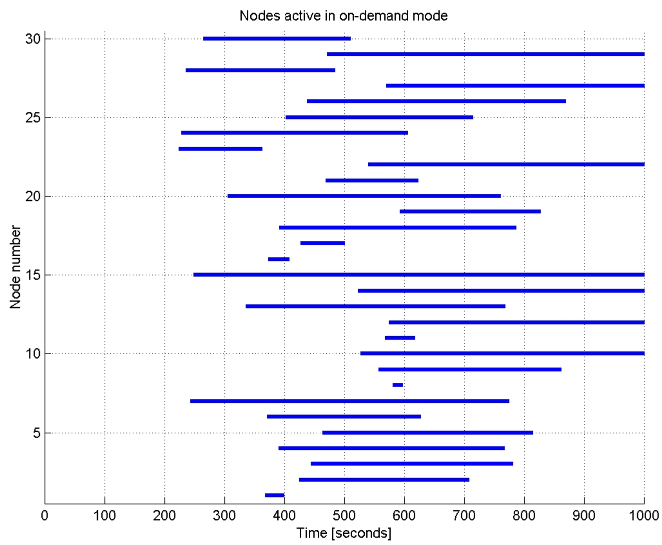

Figure 33: Test case 4: RWP, ad-hoc

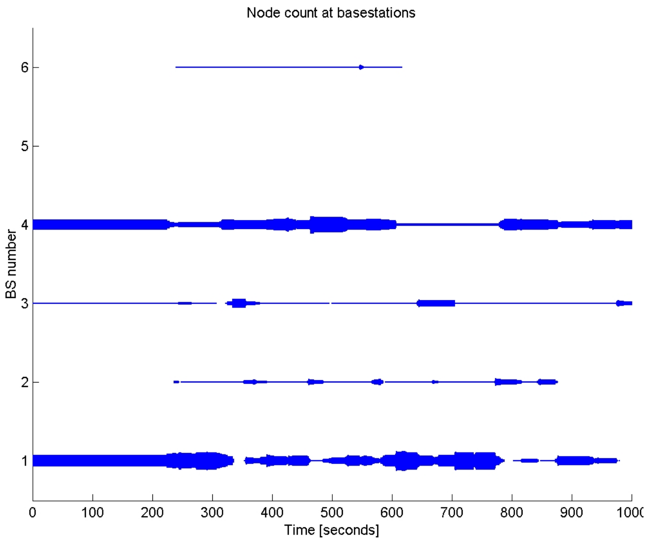

Figure 34: Test case 1: RPGM, no ad-hoc

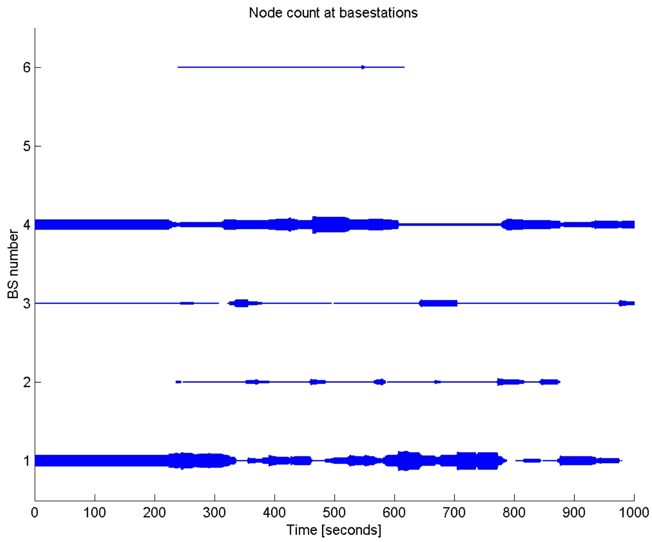

Figure 35: Test case 2: RPGM, ad-hoc

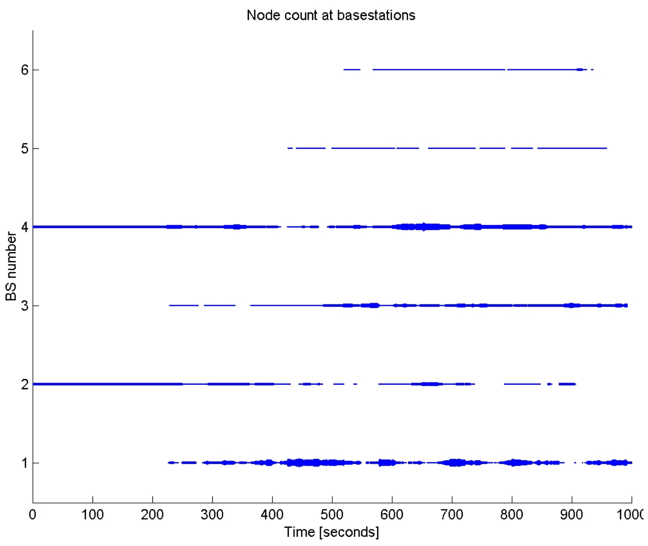

Figure 36: Test case 3: RWP, no ad-hoc

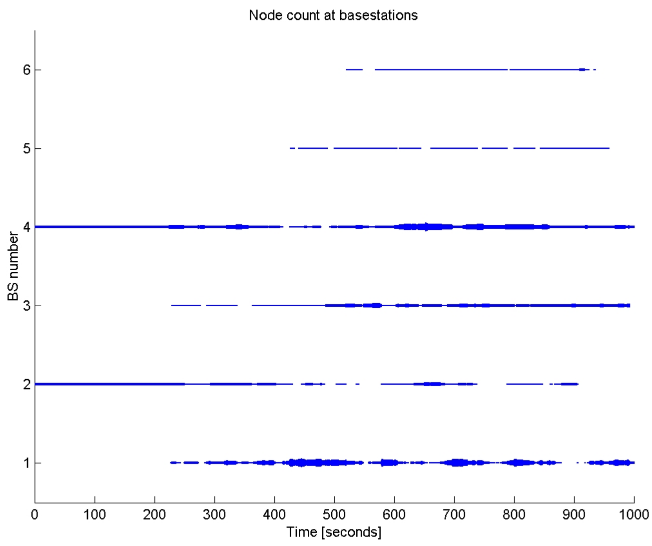

Figure 37: Test case 4: RWP, ad-hoc

Figures 34 to 37 show the number of nodes at every basestation over time. It turns out that the graphs do not show little differences between the setups, as for the two Reference Point Group Mobility cases. However, some main tendencies are well shown by this method. Looking for example at the distribution of nodes, it is clearly visible that the traffic is better balanced over the deployed technologies in the case of Random Waypoint mobility. This is somehow evident for the large group sizes which result in totally five groups traveling the simulation area in average and therefor use similar resources. On the other hand, with this scheme there is no way to read from the graphs how many resources are precisely saved in the infrastructure with the ad-hoc mode. Thus, the approach is more qualitative than quantitative and other ways have to be found to represent the results from an operator's perspective.
The last measurements presented in Figures 38 to 41 show the session based transmission time for the four setups. Again, the mean overall duration of transmissions is indicated by a straight line. In the case of group mobility, the graph shows that there are several sessions which can be reduced in duration due to the enabled ad-hoc mode. However, with Random Waypoint Mobility, there is no difference visible at all, comparing the two modes. Like for the measurement of the activity of the nodes, also here the number of unfinished sessions falsifies the results, because there are a few sessions with higher throughput in ad-hoc mode but they neither have any impact on the on-time of the nodes nor the duration of transmission in case of session ends at simulation termination. However, like for the other cases also here it will be helpful to have the global overview of the mean durations rather than the information about the single sessions.

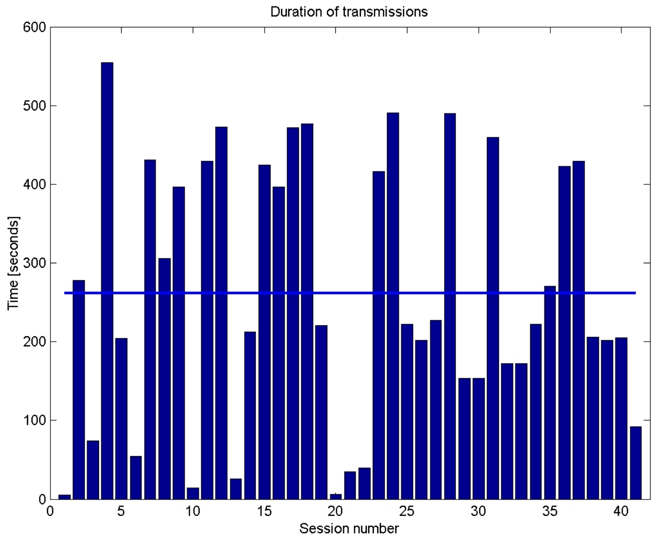

Figure 38: Test case 1: RPGM, no ad-hoc

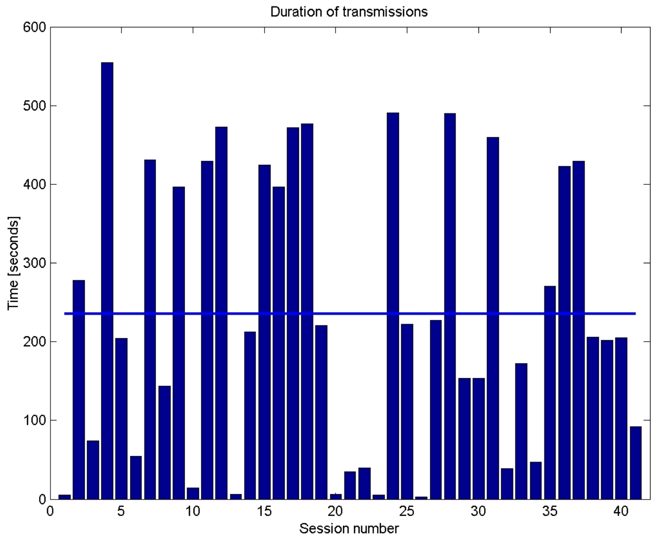

Figure 39: Test case 2: RPGM, ad-hoc

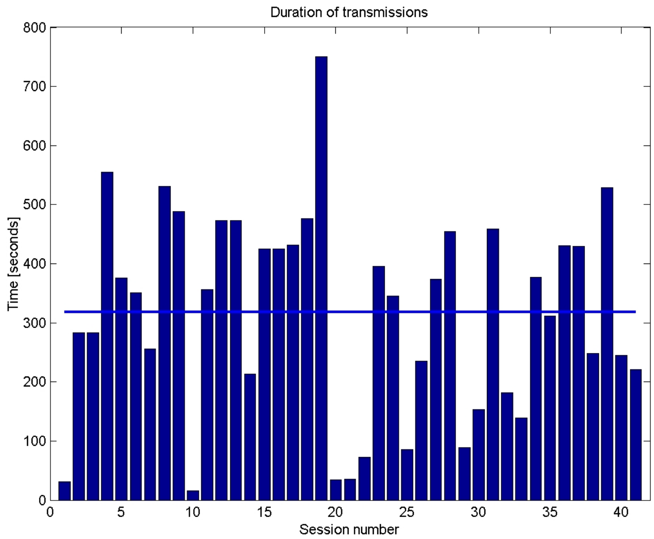

Figure 40: Test case 3: RWP, no ad-hoc

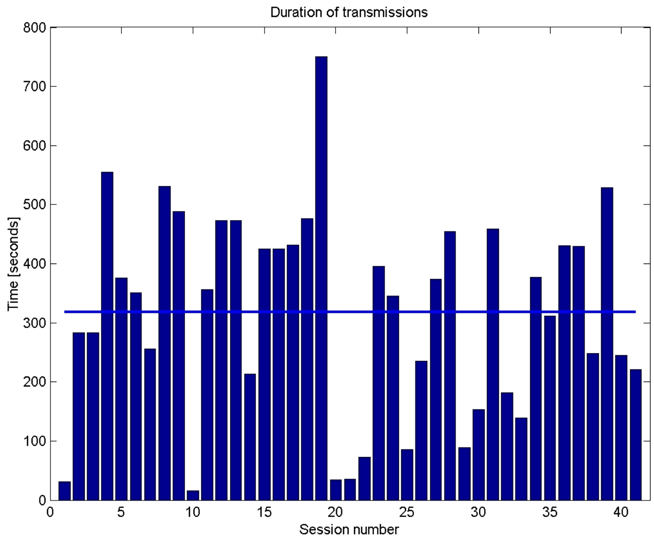

Figure 41: Test case 4: RWP, ad-hoc

5.3 Simulation of user relevant aspects

In the first simulation phase with the final version of HNS, the following parameters, which are introduced in Section 3.1, are fixed:

General parameters	
Number of nodes:	50
Scenario duration:	4600 s
Duration of initial cut off phase:	1000 s
Transmitted data per session:	10 KB to 100 MB (Randomly distributed)
Mobility parameters	
Minimum node speed:	1 mps
Maximum node speed:	15 mps
Maximum pause time:	30 s
RPGM group size:	4 (with a standard deviation of 3)
RPGM group change probability:	0.3
Standard values of variable parameters	
Simulation area size:	3000 m by 3000 m
Session density:	4 sessions per hour (Pareto distributed)

The nodes travel a given simulation area either by following the Random Waypoint (RWP) or the Reference Point Group mobility (RPGM) model. During the simulation multiple random sessions are established between pairs of nodes. The starting times of the sessions are Poisson distributed whereas the amount of data to be transmitted during the sessions follows a Pareto[7] distribution. The transmitted data is between 10 Kilobytes and 100 Megabytes per session and there are four sessions per hour for each pair of communicating nodes. The handover strategy of the nodes aims at always choosing the wireless technology with the highest bandwidth available, i.e. a vertical handover occurs whenever the currently used technology is no longer present or when an access technology providing higher bandwidth than the one in use becomes available. The resulting actual data transfer rate is equal to the minimum bandwidth of the technologies currently allocated by the two communicating nodes. The wireless technologies deployed in the simulation area are in this first simulation set inspired from the infrastructure-based technologies GPRS, UMTS and WLAN but primarily only defined by their bandwidth, range and coverage, which will be the tunable parameters for the different simulation scenarios. For the energy consumption estimations, the relative values shown in 4 have been used based on [15] and [16].
The four simulation scenarios are set up by varying the node density, the number of sessions, the ratio of the bandwidths of the available wireless technologies, and their coverage.
As the maximal distance from a node to its group center is 50 meters, which results in a probability of 87.5% for the nodes within the same group to communicate directly. This is because of the randomly chosen value for direct communication between 50 and 150 meters

[7]Pareto distribution: The probability density function of this distribution is $\frac{kx_m^k}{x^{k+1}}$, where x_m is the minimal possible value for a session start of node x.

	Receiving	Sending	Idle	Sleep
WLAN	1	2	1	0.05
GPRS	3	4	1	0.05
UMTS	2	3	1	0.05

Table 4: Used relative energy consumption values

and the totally circular antenna patterns of all technologies. Furthermore, the narrowest bandwidth available has a coverage of 100% of the simulation area. A second technology provided is 10 times faster and covers 50% of the simulation area. The third technology is 100 times faster than the first but with only 5% coverage. These values provide a reasonably rough approximation, considering today's deployed technologies such as GPRS, UMTS, and WLAN. The base stations and access points are distributed randomly all over the simulation area. The chosen bandwidth for ad-hoc peer-to-peer communication is assumed to be 10 times faster than Wireless LAN 802.11b standard, which would be achieved with UWB. The transmission range for the ad-hoc technology is randomly set to a distance between 50m and 150m for each of such communication.
The four scenarios are simulated for the four cases when nodes have each of the investigated features of ad-hoc node-to-node communication and on-demand features enabled/disabled, i.e., either ad-hoc or on-demand mode enabled, neither ad-hoc nor on-demand enabled, or both modes enabled. The average session duration is evaluated and the energy consumption is estimated in order to quantify the benefits of the ad-hoc and on-demand mode. The following sections focus on the benefits resulting from ad-hoc and on-demand modes from a user's point of view, i.e. throughput and energy consumption. All simulation results are given with a 95% confidence interval.

5.3.1 Varying node density

The first scenario evaluates the impact of the node density on the performance by varying the side length of the square of the simulation area from 1000 to 10000 meters.
The larger the simulation areas are, the lower is the probability that two peers can communicate directly in ad-hoc mode thus, the benefit of the ad-hoc feature is reduced with growing square size. This behavior is reflected in Figure 42. As the throughput does not differ for enabled on-demand feature it is not considered for the throughput evaluations.
For small areas, the RPGM model results in high probability that peering nodes come close enough to benefit from the high data rate ad-hoc link. For the smallest simulation area this results in average throughput increase from 200kbps up to 800kbps. If the ad-hoc feature is disabled, the throughput is quite constant for all simulation areas, which is expected since the relative coverage for the different technologies is independent of the size of the area. The consumed energy per node for the different simulation areas is represented in Figure 43.
Regarding energy consumption, approximately 20% can be saved if the on-demand feature is enabled. Additional 20% can be saved if the system switches to ad-hoc links in case of small communication distances. This is mainly due to the increased throughput, which in

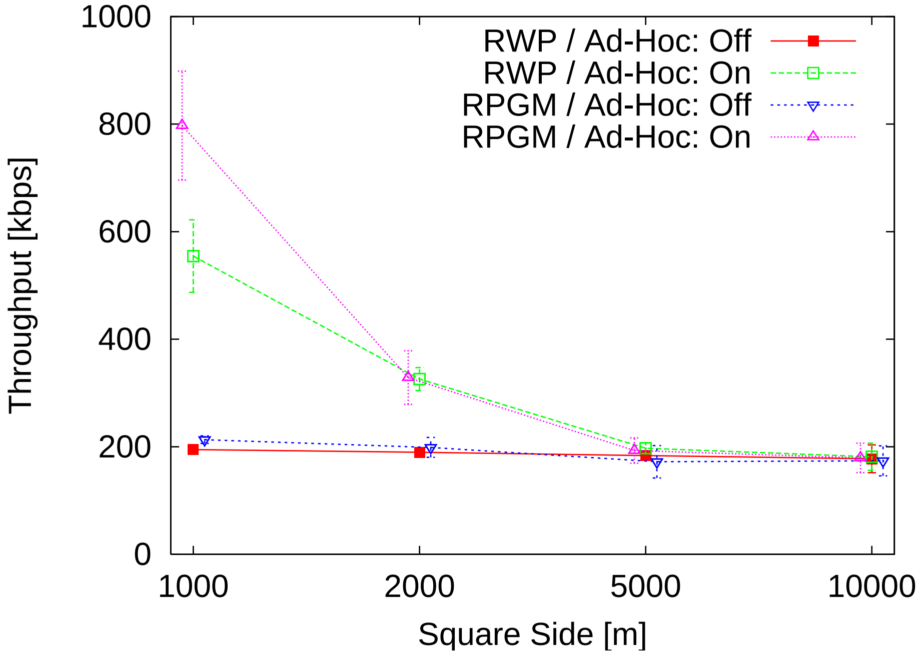

Figure 42: Throughput for varying node density

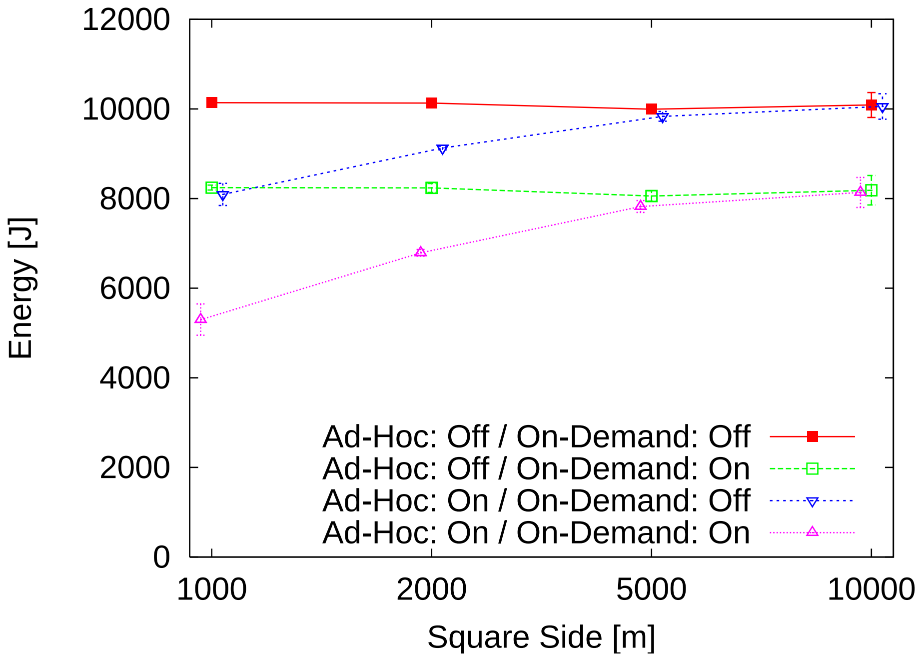

Figure 43: Energy consumption for varying node density

turn results in shorter transmission durations.

5.3.2 Varying session density

For very high session densities, nodes are constantly sending and/or receiving data anyway, such that the on-demand feature is not really beneficial anymore. In case that nodes send or receive data only rarely, the on-demand mode puts them into sleep mode until a new session is either initiated by themselves or by another node. Nodes without on-demand capabilities have to remain idle with all networking interfaces turned on in order to be reachable for incoming sessions. Today's devices show a significant difference in energy consumption for idle and sleep nodes. Thus, the on-demand feature reduces notably the use of battery power. In this simulation scenario, the session density is varied from 1 to 40 sessions per hour and communicating pair. In Figure 44, the impact of the session density on the average throughput of the sessions is shown.

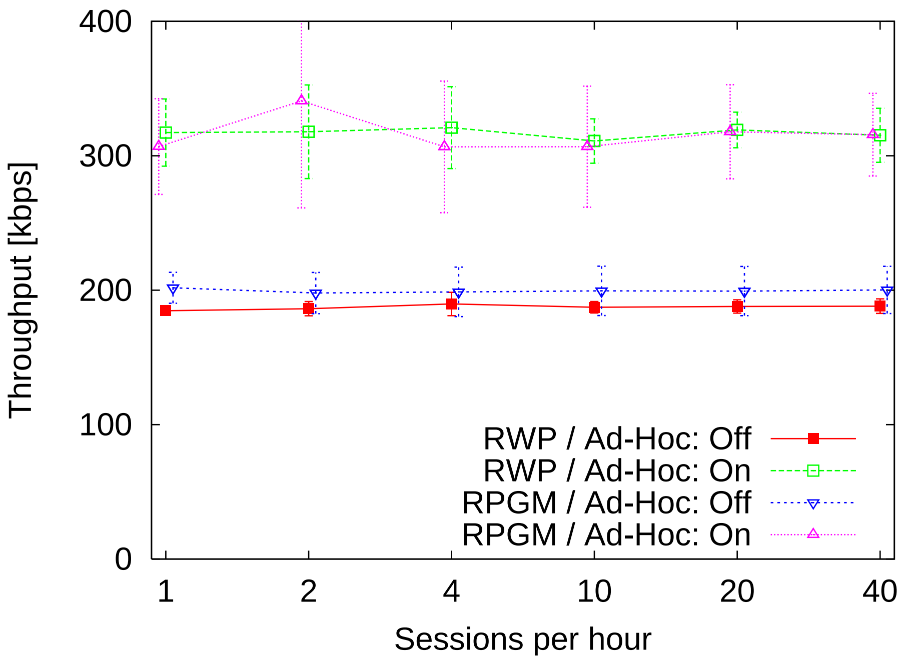

Figure 44: Throughput for varying session density

Since here the simulation area is set to 2000 by 2000 meters, the improvement is the same as the one depicted in Figure 42.

The throughput resulting from the RPGM scenario is not as significantly different than the throughputs of the RWP scenarios, which is due to the small group size chosen for the RPGM. With bigger group sizes the probability that two communicating nodes are within the same group and therefore able to use ad-hoc links is smoothly higher. However, choosing the group size too big is not reasonable either.

Figure 45 shows the energy consumption values for the different session densities. The potential energy savings depend very much on the session density, since the devices do only save energy if no session is ongoing. Again, the increased average throughput with enabled ad-hoc mode is decreasing the average session duration and thus the energy consumption.

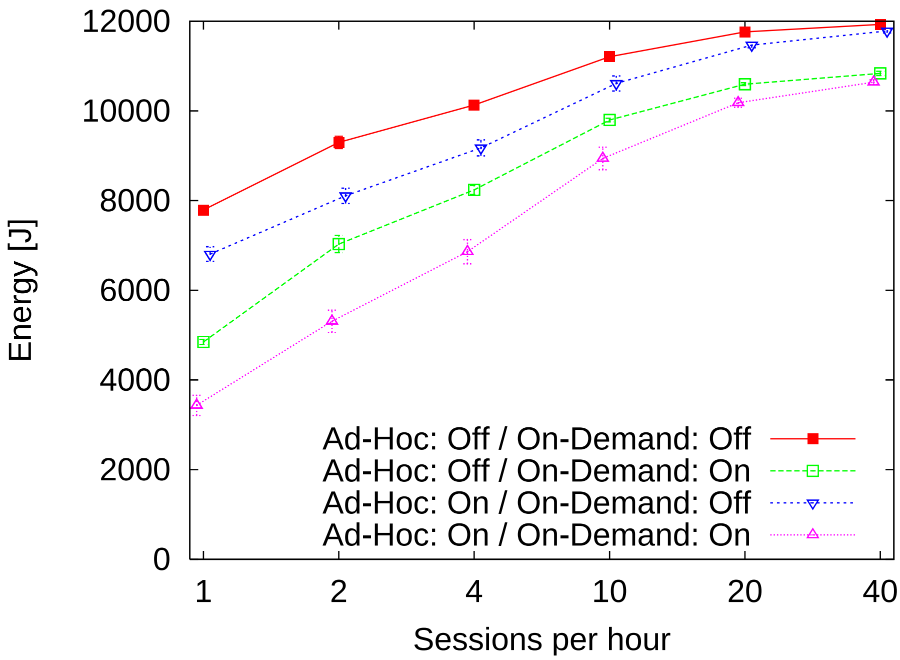

Figure 45: Energy consumption for varying session density

5.3.3 Varying bandwidth ratio

In this section, three infrastructure-based and one infrastructure-less technology for node-to-node links are tuned in terms of bandwidth for this scenario. The first kind of technology provides almost full coverage but has only limited bandwidth such as GPRS, EDGE, or also satellite networks. The second kind of technology represents 3G wireless networks such as UMTS, which provide higher bandwidth, but are not yet as widely deployed as 2 and 2.5G networks, and more often within urban areas. The third kind is represented by wireless broadband technologies. They are commonly not area-widely deployed, but at specific locations only, such as 802.11b in so-called Hotspots. Additionally, nodes can communicate directly without any infrastructure in direct ad-hoc mode. This can be done with technologies such as WLAN or UWB. The bandwidth ratio of these four kinds of technologies may vary strongly due to the number of currently active users, SNR (Signal to Noise Ratio), user and operator policies, etc. Two scenarios are set up with different bandwidth ratios for the used technology. The narrowest bandwidth technology of the four (e.g. GPRS) gives the reference value for the three others. In the first scenario the second technology (e.g., UMTS) provides twice the bandwidth of the first technology, the third (e.g., WLAN) 20 and the ad-hoc (e.g., UWB) 1000 times more than the first technology, i.e. 1:2:20:1000. The second scenario is simulated with a technology bandwidth ratio of 1:10:100:1000.
Figure 46 shows the average throughput for the two bandwidth ratio setups. The first scenario, having a very high difference between the data rates offered by the infrastructure-based and the ad-hoc communication technology, is very much profiting from the ad-hoc feature whereas the throughput increase in the second scenario is only about 20%. With these scenarios, the average throughput can be increased by the factor of 4. However,

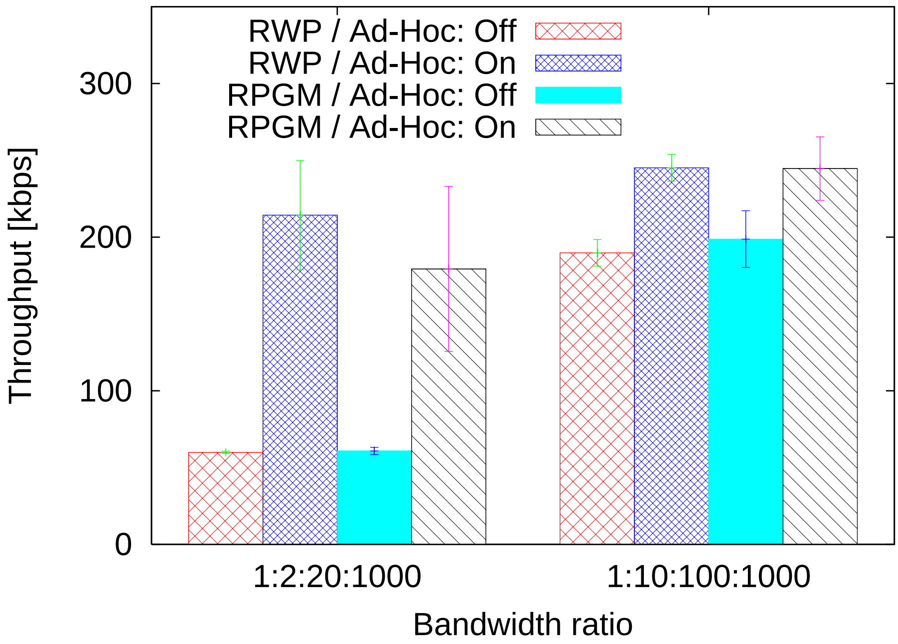

Figure 46: Throughput for varying bandwidth ratio

assuming further development of high bandwidth ad-hoc technologies like UWB offering data rates of up to 1Gbps, the capability of seamlessly switching to ad-hoc links could improve the average throughput even more.

5.3.4 Varying coverage of infrastructure technologies

In the last scenario, the performance in terms of throughput and energy consumption is analyzed for two different coverages of infrastructure-based technologies. In the first case, the coverage of the first, second, and third technology is 50%, 25%, and 1%, which can be considered as low coverage whereas in the second case the coverage is 100%, 80%, and 10% of the whole simulation area, which can be depicted as high coverage. The impact of the variation of the relative coverage is shown in Figure 47

It can be stated, that for low coverage of infrastructure-based technologies the added value of ad-hoc link usage is higher than for highly covered areas. In the first scenario the probability of having no or only very narrow band connection is rather high. Thus, even if the chance of having direct communication via the ad-hoc links is small as well, the impact on the session throughput is so much bigger once it occurs. Regarding the energy saving potential, the ad-hoc feature has nearly no effect compared to the on-demand capability for low coverage values, which is shown in Figure 48. For high coverage, the ad-hoc mode is more beneficial, but still outperformed by the added value of the on-demand feature.

5.4 Simulation of user and operator relevant aspects

The second simulation set further analyzes the influence of both peer-to-peer ad-hoc mode and on-demand ability on the quality of communications. The three access technologies are

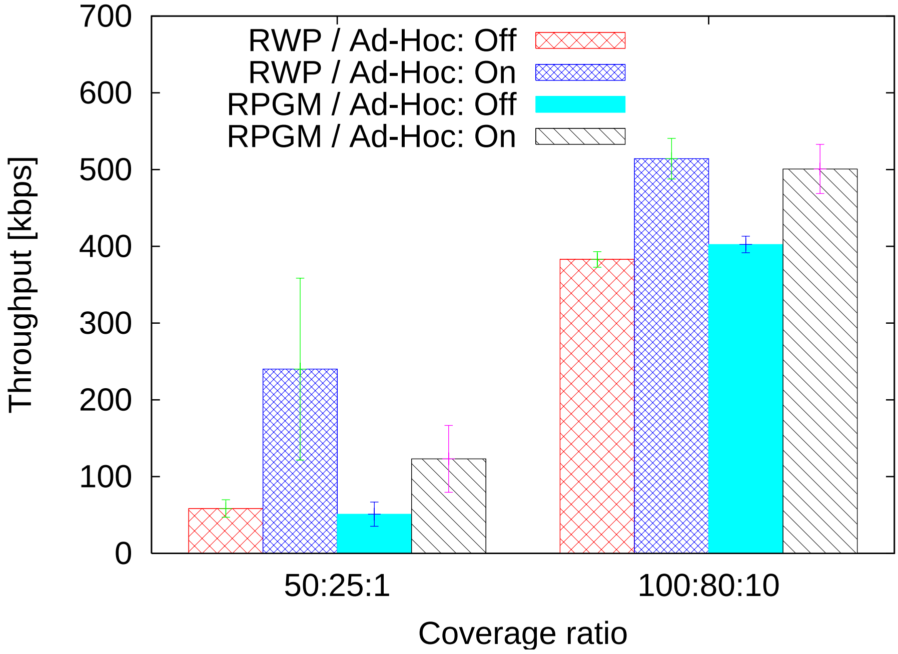

Figure 47: Throughput for varying coverage

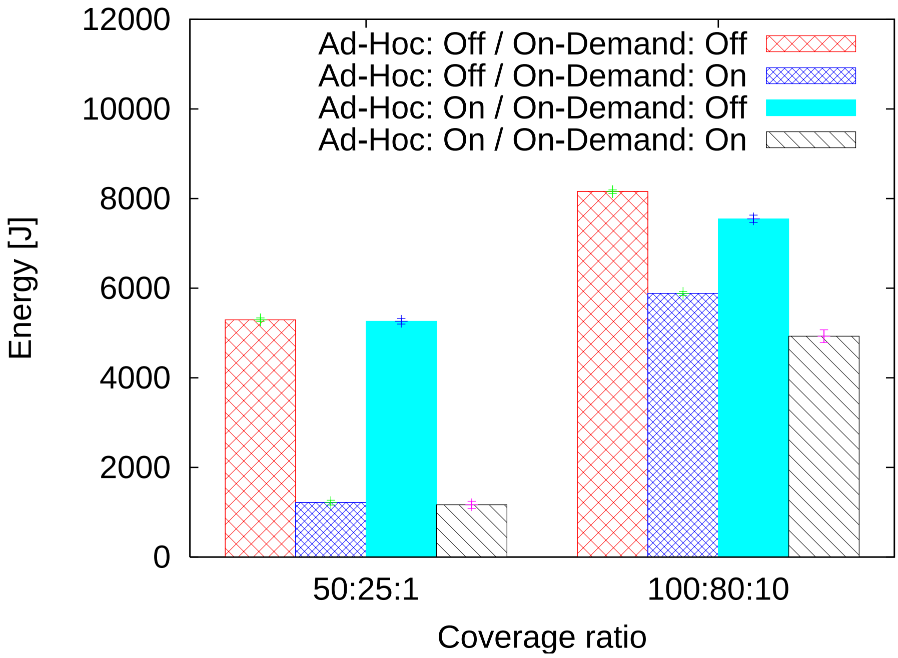

Figure 48: Energy consumption for varying coverage

modeled after GPRS, UMTS and WLAN with limited capacities and adaptive data rates delivered to each node, depending on the actual load of the cell or access point. The data rates obtainable by WLAN are defined as $1/(n+1)$ of the 11 Mbps provided by a standard 802.11b access point, where n refers to the number of active nodes attached to that access point. The capacity available in GPRS cells is adapted according to the total number of nodes in simulations. There are two GPRS cells per simulation, covering the area fully and each providing one slot at least to 50% of the total users. Assuming that the nodes are always exactly equally distributed, this would result in serving all nodes with a minimal data rate of one TDMA slot for up- and downlink. However, a non-uniform distribution, which surely occurs from time to time with both mobility models, RWP and RPGM, would result in blocked or dropped sessions. The bandwidth allocation in the simulation for both UMTS and GPRS works as described in Section 4.5.5.
The parameters used for this simulation set are the following:

General parameters	
Simulation area size:	2000 m by 2000 m
Scenario duration:	4600 s
Duration of initial cut off phase:	1000 s
Transmitted data per session:	10 KB to 100 MB (Randomly distributed)
WLAN coverage:	10%
UMTS coverage:	80%
GPRS coverage:	100% (Capacity limitations mentioned above)
Mobility parameters	
Minimum node speed:	1 mps
Maximum node speed:	15 mps
Maximum pause time:	30 s
RPGM group size:	4 (with a standard deviation of 3)
RPGM group change probability:	0.3
Standard values of variable parameters	
Session density:	2 or 8 sessions per hour (Pareto distributed)
Session density:	2 or 8 sessions per hour (Pareto distributed)
Number of nodes:	25, 50, 75, 100, 125, 150, 175, and 200

5.4.1 User benefits

Session block probability

A session block occurs every time a node is not able to start the session because of missing networking resources. The node is then waiting and continuously trying to start the data transmission until resources become available in the heavily loaded cell or the node is moving to another cell with available capacity. The node is taking any other available networking technology if the preferred technology is not available. Thus, the network is overriding the node's preference if the overall network performance can be increased. Figure 49 to Figure

52 show the session block probability for the RWP and the RPGM mobility model for both, 2 and 8 sessions per hour.

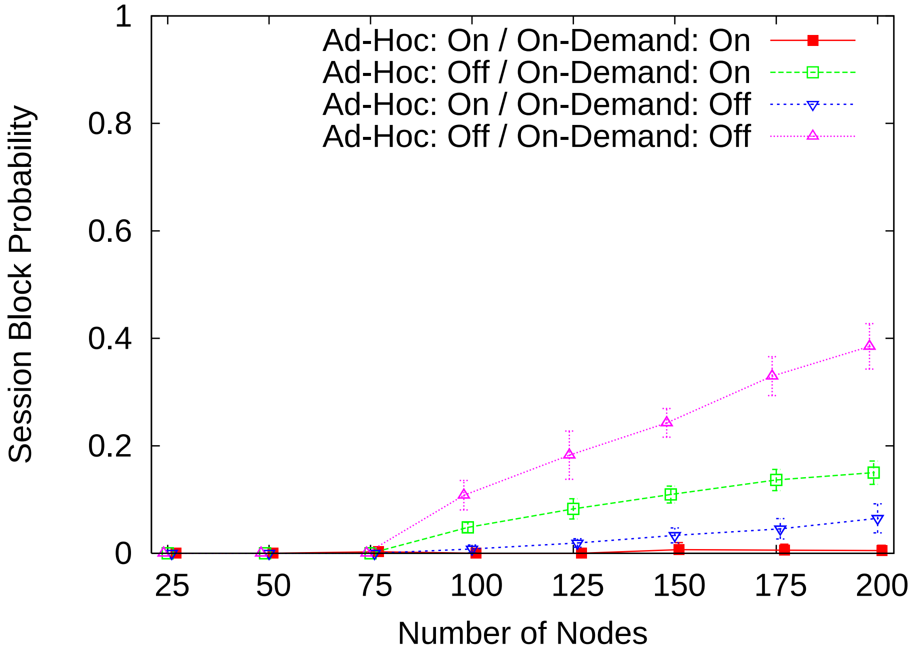

Figure 49: Session block rate: RWP with 2 sessions per hour

Figure 49 shows that for 2 sessions per hour, the network is able to serve approximately 75 nodes without blocking any session start. The ad-hoc mode has the biggest positive impact on the session block probability. Together with the on-demand feature, the blocking probability can be kept very low for up to 150 nodes. With increasing session density, the influence of the on-demand mode on the session blocking probability is also considerably reduced, which can be seen in Figure 50.
Figure 51 shows that if the RPGM mobility model is used, the same scenario results in an increase of the session block probability. This is mainly due to higher risk of resource shortage if nodes move together in groups and start sessions at the same time. The group mobility also decreases the impact of the on-demand mode on the session block probability, but increases the benefit of the ad-hoc mode at the same time.
If Figure 52 and Figure 50 are compared, the positive influence of the RPGM on the ad-hoc feature is clearly visible. The session block probability is reduced by approximately 18% for enabled ad-hoc mode.

Session drop rate and outage

Session drops result from connections, which are broken during communication. Nodes which come into the range of congested cells even without the possibility of a handover to GPRS drop their sessions. The node is then waiting for network resources to continue the session.
For the different scenarios, the number of drops per session is analyzed. Depending on the duration a node has to wait for new network resources, the overall session duration is

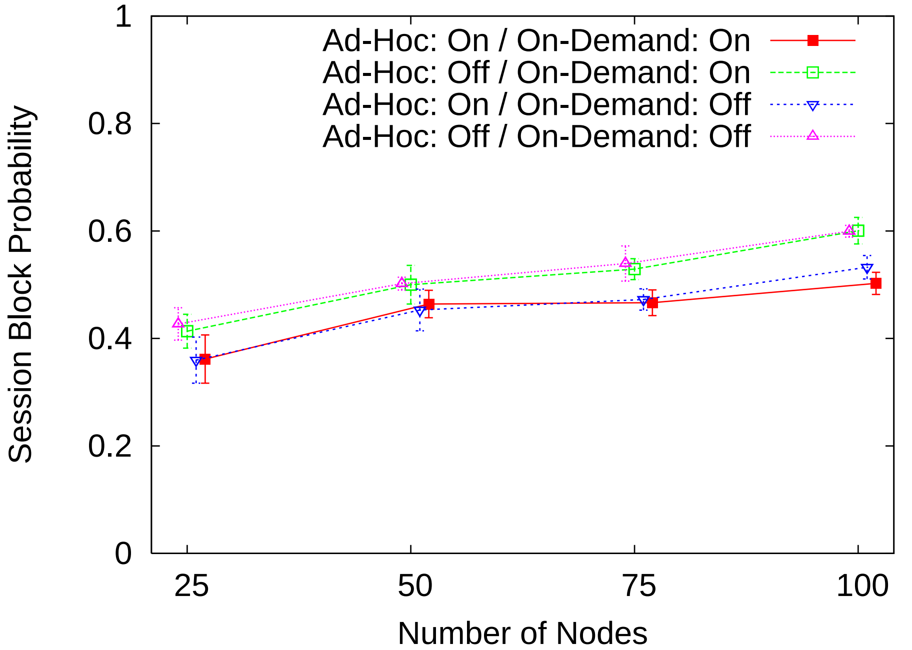

Figure 50: Session block rate: RWP with 8 sessions per hour

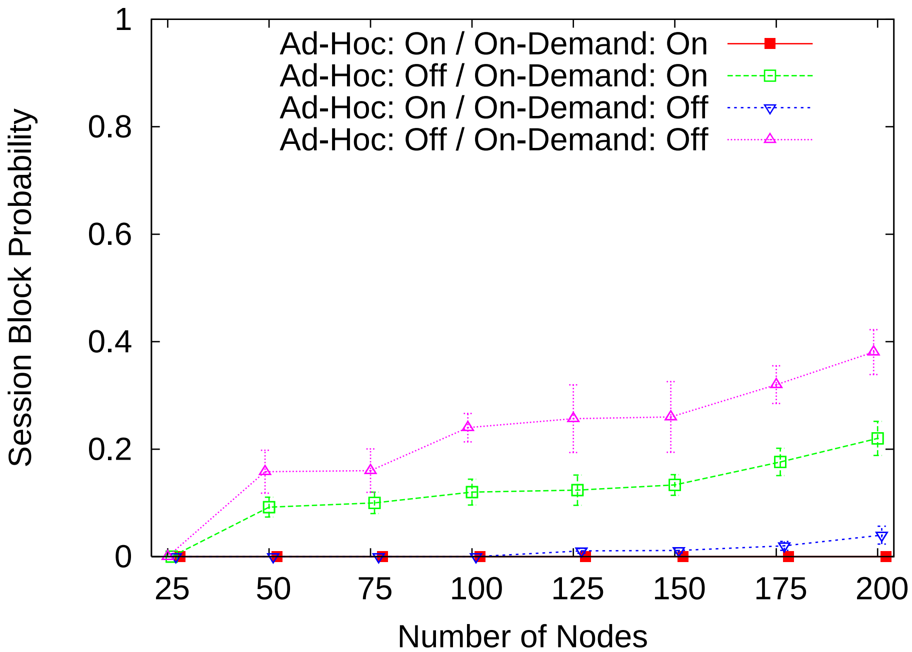

Figure 51: Session block rate: RPGM with 2 sessions per hour

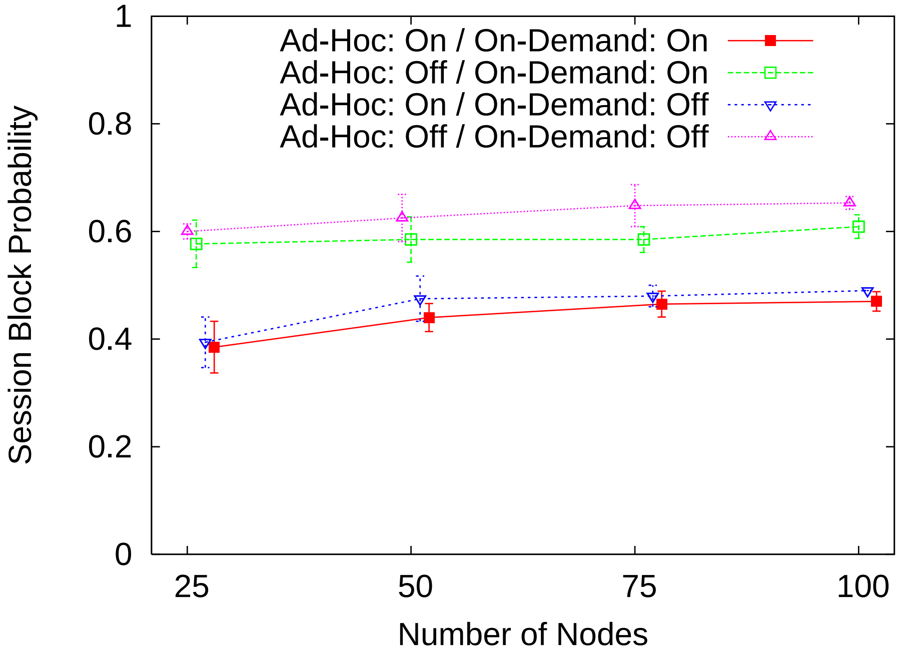

Figure 52: Session block rate: RPGM with 8 sessions per hour

increased. This duration where the node is not able to transmit any data is called outage. The outage ratio represents the time a node spends waiting for network resources after the session gets dropped with respect to the overall session duration. Increasing network load results in higher session block probabilities and session drop rates. This also increases the outage ratio and the nodes have to wait even longer before they receive network resources again. During a session outage, no drops do occur, which results in decreased number of drops per session again. If the network load is rather high, some sessions do not terminate during the simulation period and thus have fewer drops. This results in the fact that the network is not able to serve additional communication requests and thus further decreases the number of drops per session. This behavior is also represented by the following figures. Figure 53 shows the number of drops that occur during a session for the RWP mobility model, each of them starting 2 sessions per hour and having UMTS and WLAN coverage of 80% and 10%, respectively. The GPRS coverage is still 100% like for all other simulations. It can be shown that if more than 75 nodes are using the network, the sessions get continuously dropped. If more than 125 nodes are present, the average number of drops per sessions reaches its maximum. For more than 125 nodes the network load is so high that the requests for new sessions are just blocked and no more session drops occur.
With enabled on-demand mode the drops can be avoided completely due to the freed network resources between the sessions. The ad-hoc feature is considerably reducing the drop rate but is not able to avoid all session drops. Figure 54 shows the outage ratio for the same scenario. Without the help of the ad-hoc and on-demand mode, the outage ratio is increasing with the number of nodes. The outage starts to increase simultaneously with the session block probability as discussed before. With 8 sessions per hour, the drop rate also increases. Here, 25 nodes are already sufficient to load the network heavily, which results

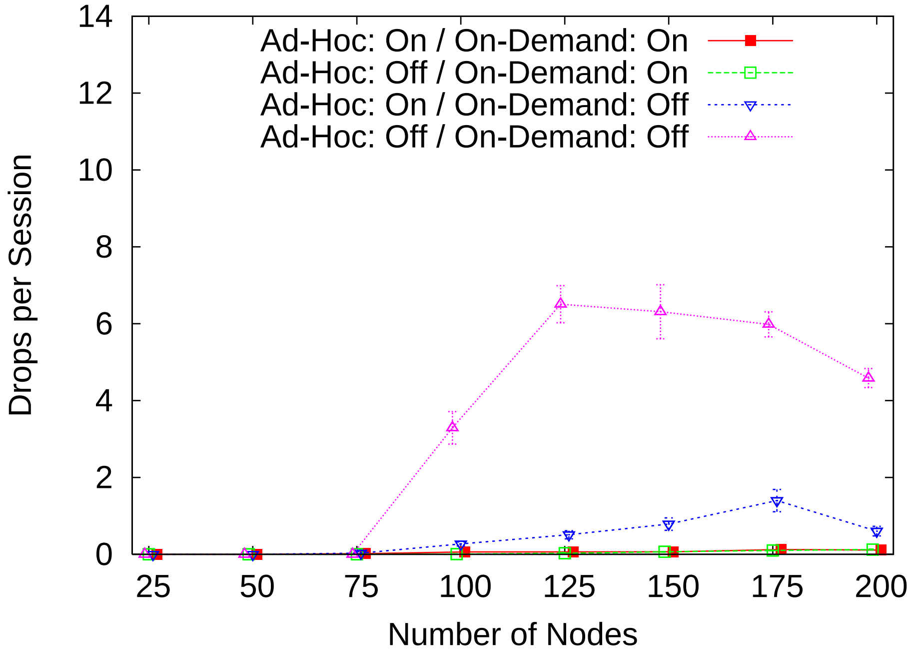

Figure 53: Session drop rate: RWP with 2 sessions per hour

in drop rates of about 100 drops per session. With 50 nodes, the drop rate reaches its maximum and decreases smoothly. For such a heavy loaded network the on-demand mode is not suited anymore. However, the ad-hoc mode further decreases the amount of drops. Figure 55 shows the drop rates with 8 sessions per hour. The outage ratio is depicted in Figure 56.
The RPGM mobility model increases the number of drops per session like it did for the session block probability, because the nodes moving in groups are competing for the same network resources. In Figure 57 and Figure 58 the drop rate and outage ratio are shown. In Figure 59 it is clearly visible that the on-demand feature is no more influencing the drop rate for heavily loaded networks. However, the ad-hoc feature still profits from the RPGM mobility model and reduces the drop rate up to 50%.

Session throughput

In the first simulation set the session throughput is evaluated assuming unlimited network resources. However, the values obtained from the second simulation set include the delays and outages resulting from the session blocking and dropping, which may occur if nodes share the network resources. These delays and outages increase the duration of a session and thus decrease the average throughput of the session. The throughput for each node is decreased further whenever the network has to perform load balancing to optimize the overall network capacity. If a node in a UMTS cell can no longer be served by UMTS, the network assigns GPRS to the specific node instead. This degradation did not occur in the first simulation set because all resources are assumed to have unlimited capacity. Figure 61 and Figure 62 show the simulation results for the two session densities of 2 and 8 session per hour, respectively, for the high coverage scenario, and RWP mobility. The ad-hoc mode

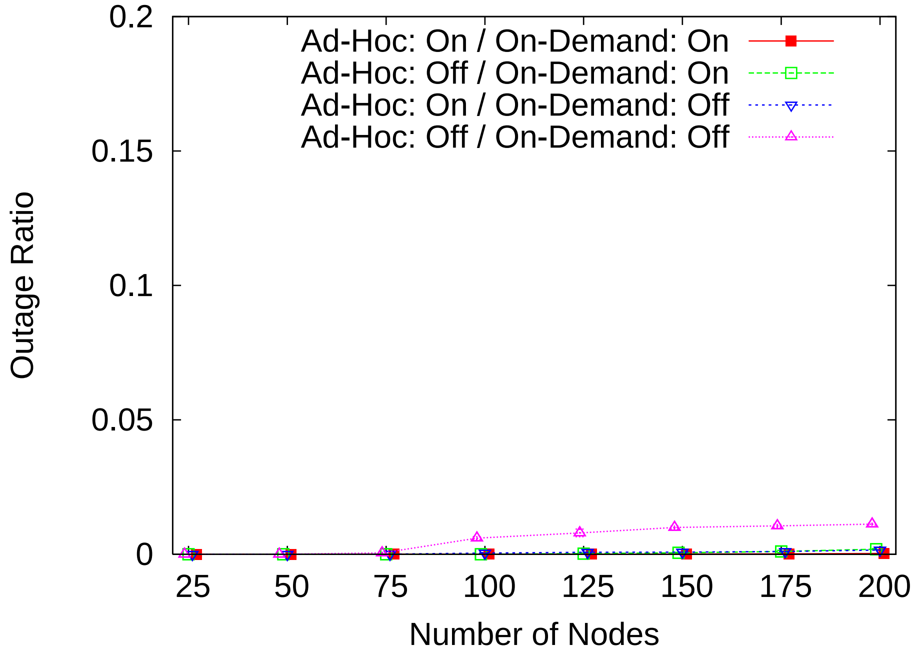

Figure 54: Session outage: RWP with 2 sessions per hour

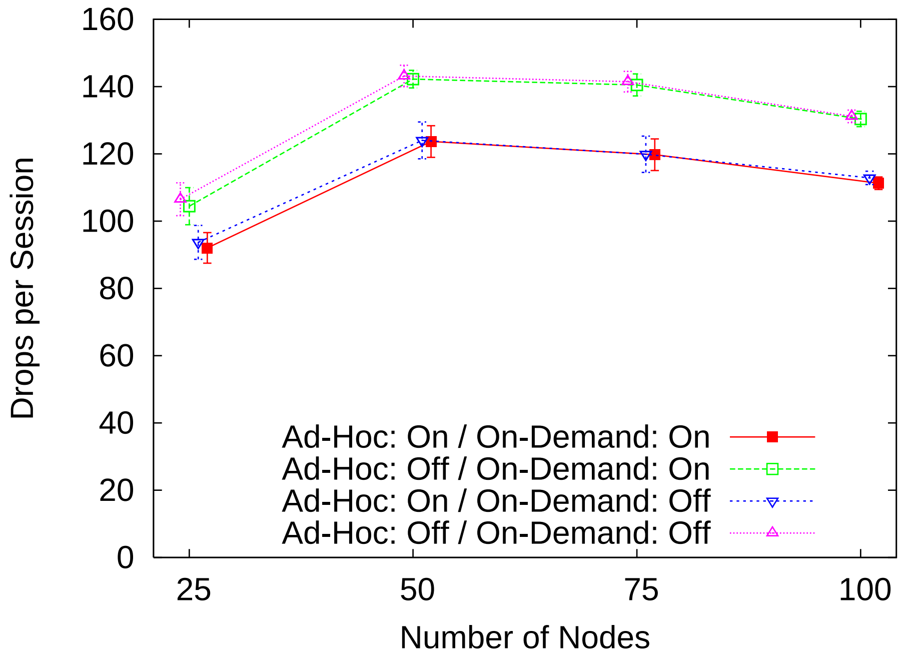

Figure 55: Session drop rate: RWP with 8 sessions per hour

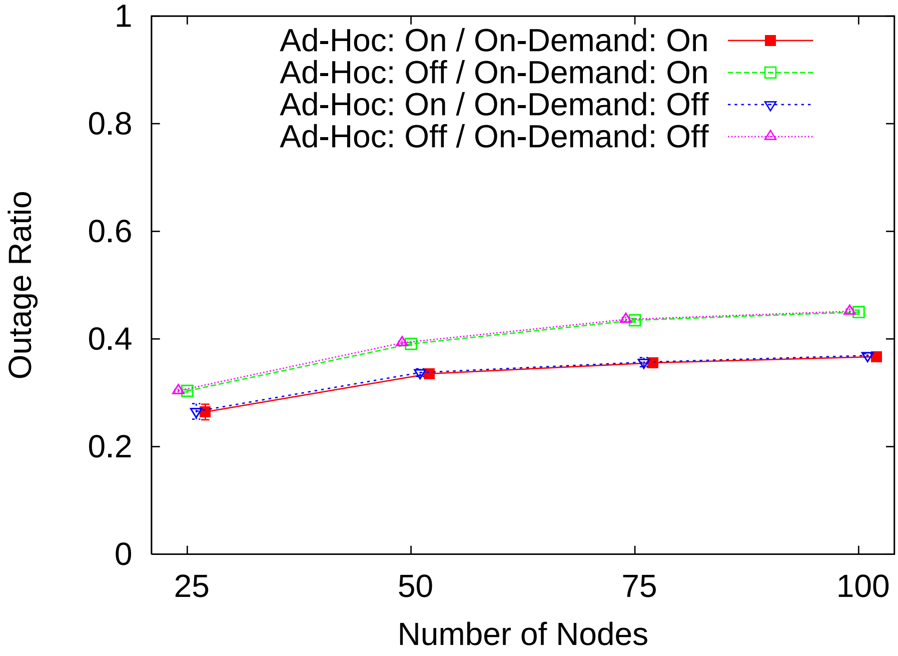

Figure 56: Session outage: RWP with 8 sessions per hour

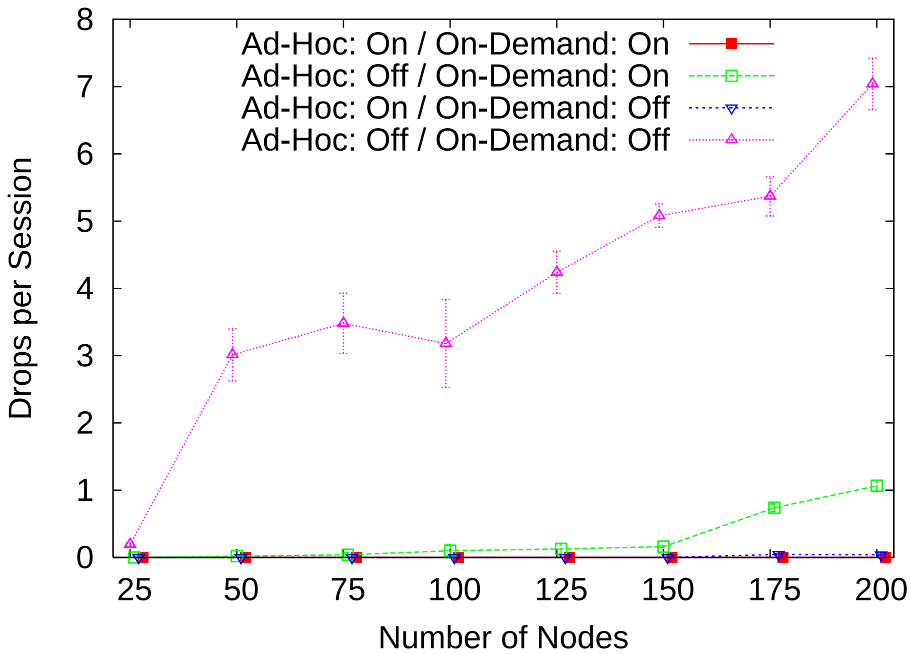

Figure 57: Session drop rate: RPGM with 2 sessions per hour

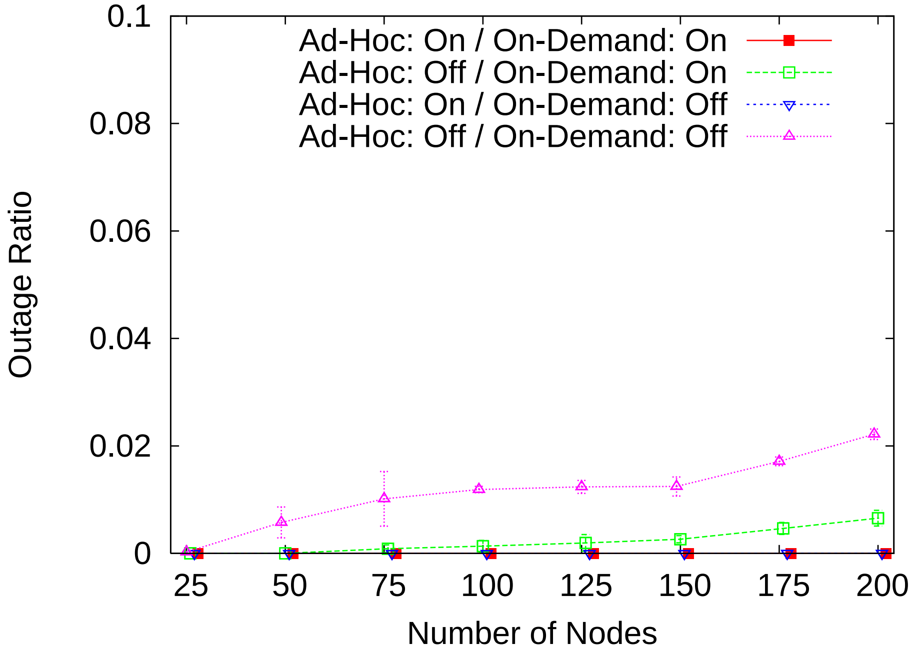

Figure 58: Session outage: RPGM with 2 sessions per hour

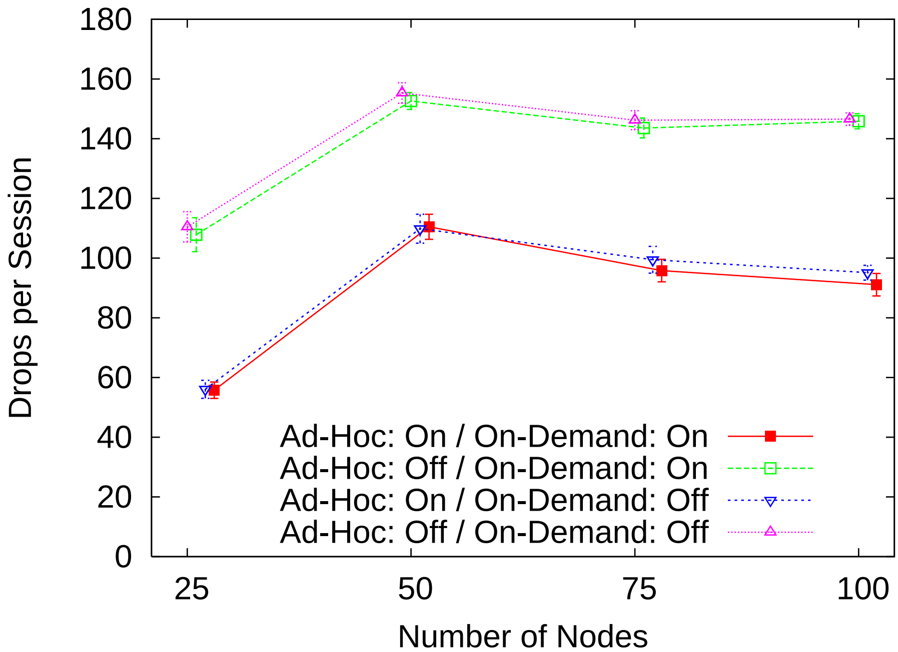

Figure 59: Session drop rate: RPGM with 8 sessions per hour

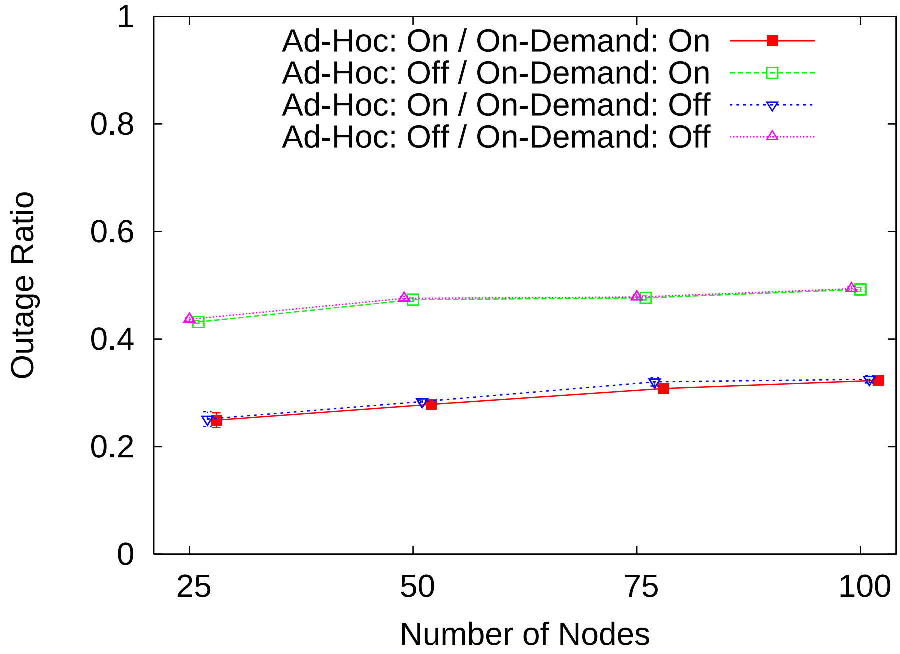

Figure 60: Session outage: RPGM with 8 sessions per hour

is much more beneficial for the measured throughput than the on-demand mode. This is mainly due to the high data rates offered by the infrastructure-less links. Nevertheless, the negative impact of the number of sessions per hour on the on-demand mode is clearly visible when comparing the two figures.

When choosing the RPGM mobility model, the average throughput is even further increased due to the ad-hoc mode. Considering the curve for the joint on-demand and ad-hoc features in Figure 63, there is a positive influence of the ad-hoc mode to the benefit introduced by the on-demand mode. Whenever communicating nodes are moving within the same group, the network is further discharged, which increases the throughput of the infrastructure-based sessions.

Figure 64 depicts the throughput for 8 sessions per hour. Both, the ad-hoc and the on-demand feature are highly beneficial for low node densities of 25 and 50 nodes.

5.4.2 Operator benefits

Network load

The overall network load is calculated as an average of the load of each technology. Each load is weighted according to the coverage provided by that specific access technology. This should take into account that the overall network load is more dependent on the load of networks serving a large area than a small one.

The network load is evaluated for varying node and session densities. Both, the ad-hoc mode as well as the on-demand feature are enabled or disabled independently to analyze the impact on the overall network load. Figure 65 shows the four resulting network loads in function of the number of nodes for 2 sessions per hour for the RWP mobility model.

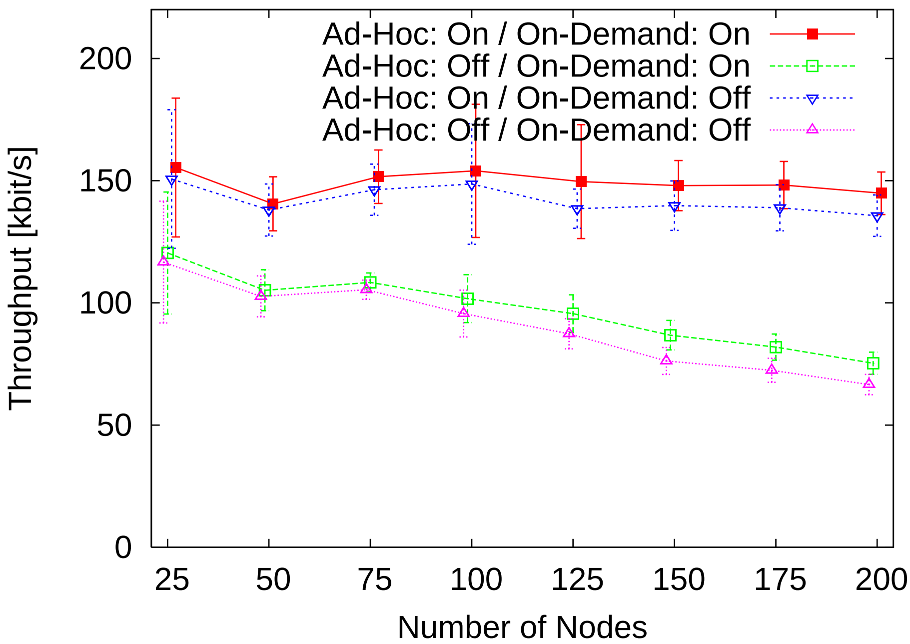

Figure 61: Session throughput: RWP with 2 sessions per hour

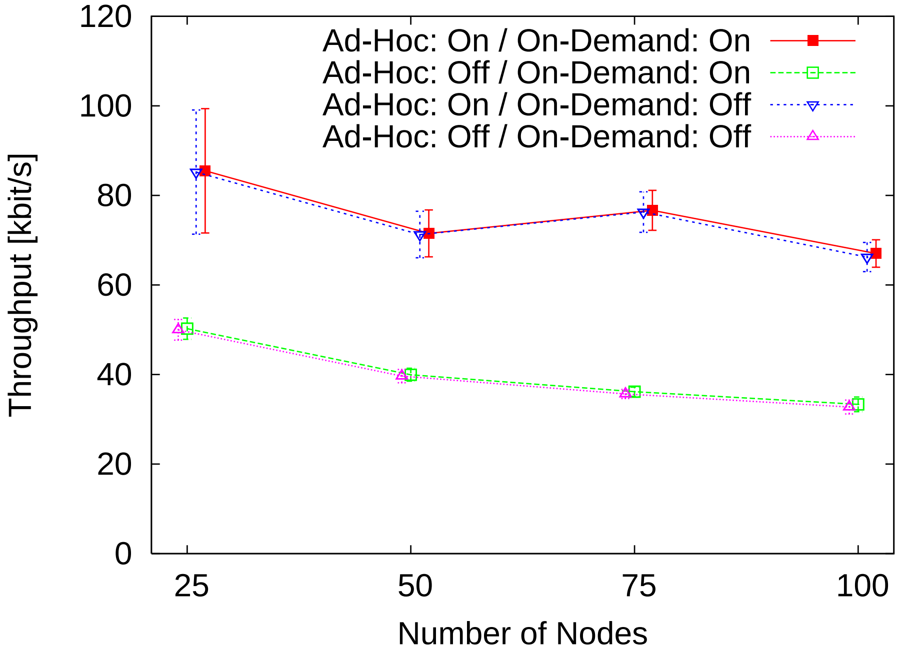

Figure 62: Session throughput: RWP with 8 sessions per hour

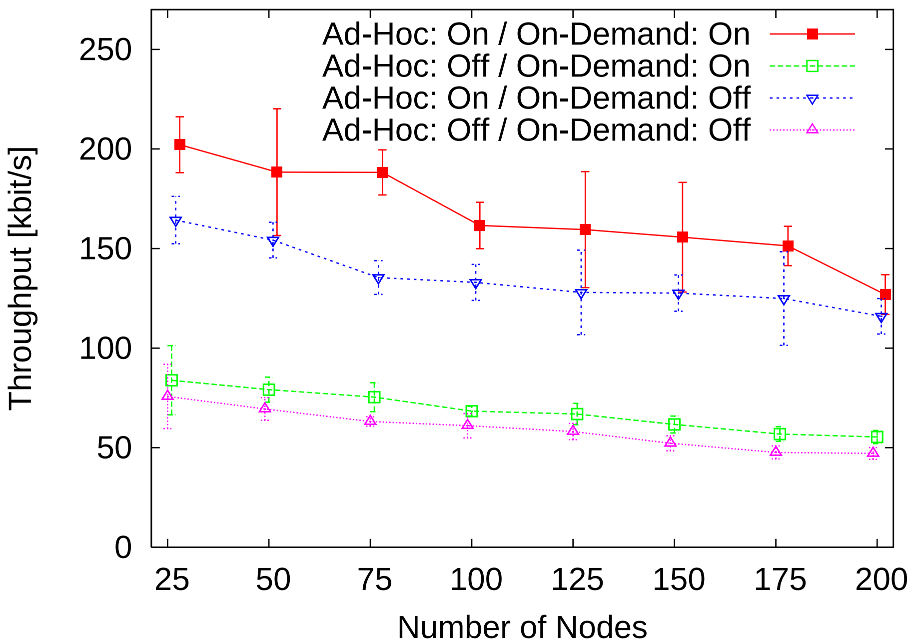

Figure 63: Session throughput: RPGM with 2 sessions per hour

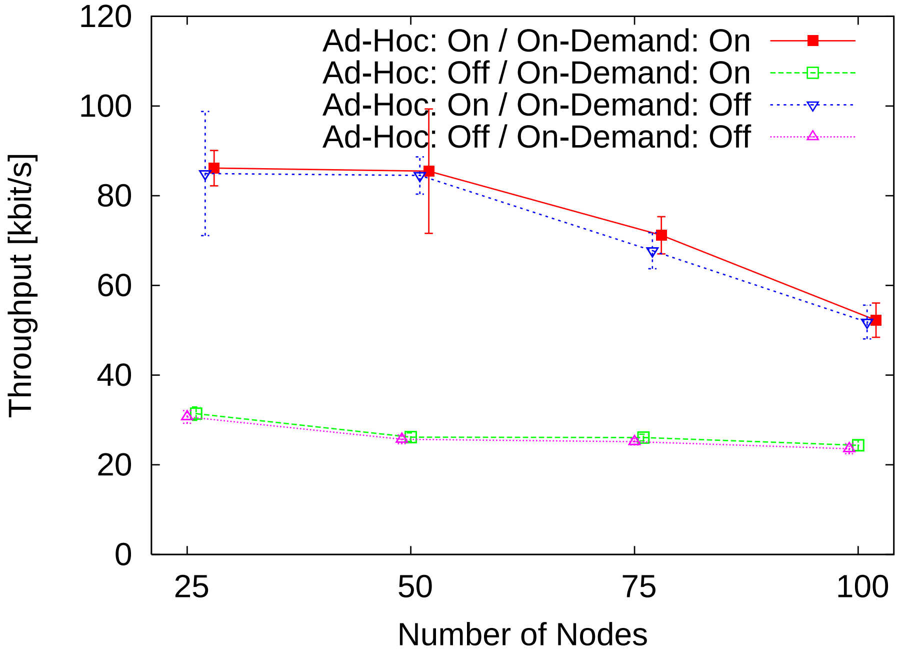

Figure 64: Session throughput: RPGM with 8 sessions per hour

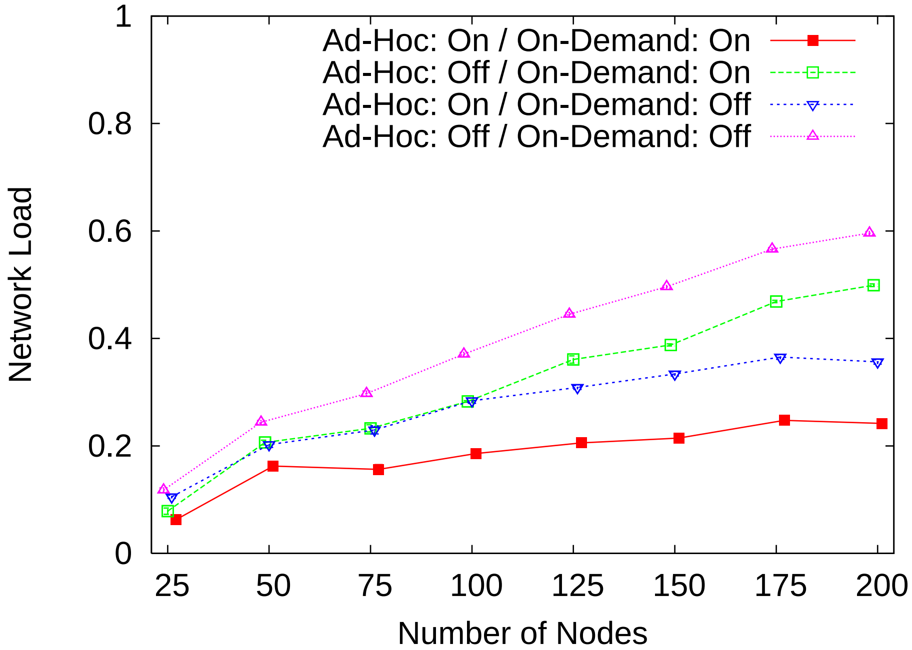

Figure 65: Network load: RWP with 2 sessions per hour

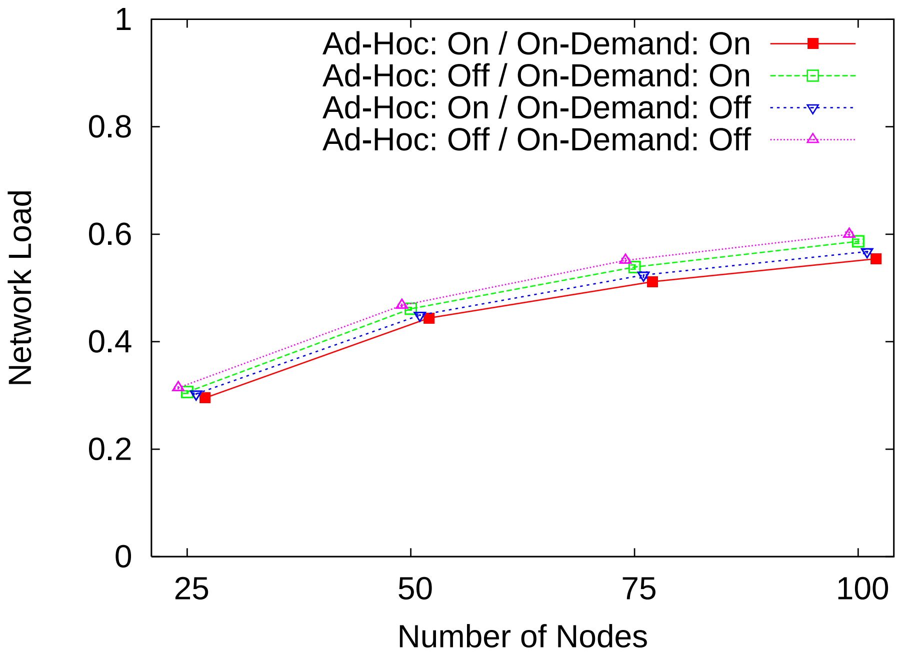

Figure 66: Network load: RWP with 8 sessions per hour

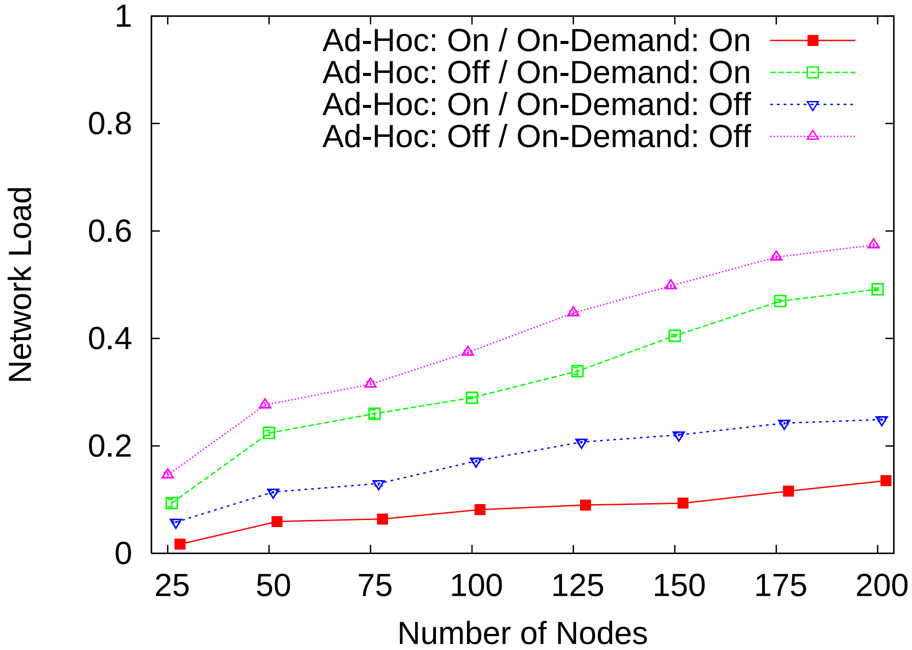

Figure 67: Network load: RPGM with 2 sessions per hour

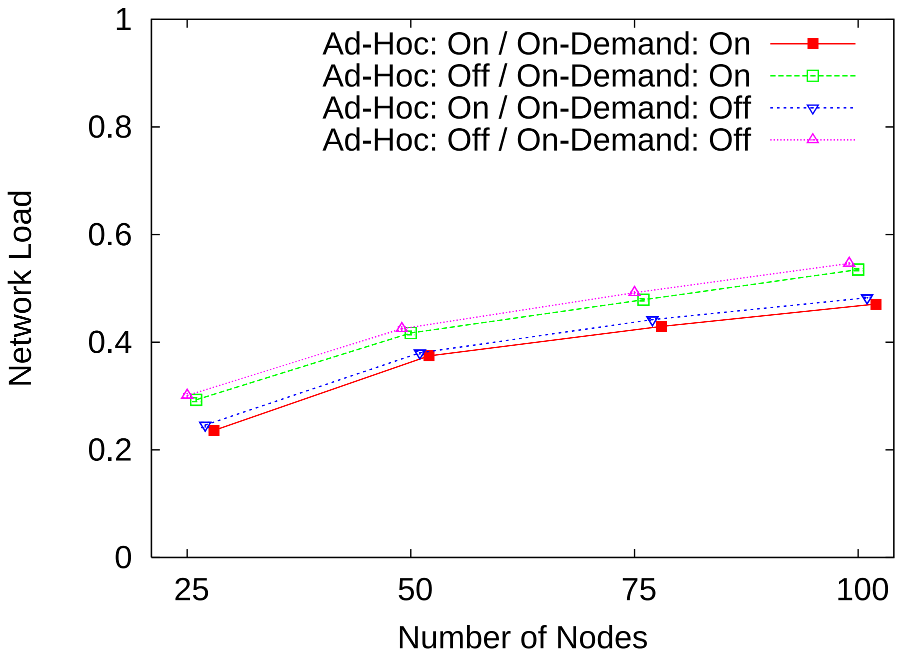

Figure 68: Network load: RPGM with 8 sessions per hour

The ability to free network resources by switching ongoing sessions to direct peer-to-peer communications whenever possible reduces the network load by up to 24% as shown in Figure 66. This implies that the network can serve more nodes at the same capacity due to the enabled ad-hoc mode. The on-demand feature is further increasing the number of nodes, which can be served. Thanks to the on-demand feature, nodes, which do not have an ongoing transmission, do not occupy network resources.
When comparing Figure 65 and Figure 66, it turns out that the lesser sessions the nodes have the higher is the resource saving potential of the on-demand feature. With 2 sessions per hour about 18% of the network resources can be freed thanks to the on-demand feature and eventually assigned to other nodes. With increasing number of sessions per hour, the benefit of the on-demand feature is noticeably decreasing. The shorter the period between the data sessions is, the less network resources can be freed. In Figure 66, the benefit from the on-demand feature is nearly negligible due to the increased session density. With 8 sessions per node, the chance to run into resource shortages is very high. Considering that some of these sessions are even overlapping, it turns out that they are competing for the same available resources. Waiting sessions are started immediately as soon as other sessions are offloaded to ad-hoc mode. Thus, the ad-hoc mode is not able anymore to discharge the network. The simulation results for the RPGM mobility model are shown in Figure 67 and Figure 68 for both session densities 2 and 8 sessions per hour.
The ad-hoc feature slightly benefits from the group mobility due to communications in the same group which result in direct peer-to-peer transmissions in most of the cases. The ad-hoc feature would be even more beneficial if it would be considered that in the real world, people traveling together more likely start connections among each others than to nodes outside of the group. Even for 8 sessions per hour the beneficial impact of the RPGM is visible in Figure 68. The on-demand capability is almost independent on the mobility pattern, which is also reflected in the figure.

Network efficiency

For a further investigation of the impact of the ad-hoc mode, a new metric is introduced. It is called "network efficiency" and is defined as the ratio between traffic sent using ad-hoc links and the overall traffic sent by the nodes. The network efficiency indicates how much traffic could be offloaded to direct peer-to-peer links. The bigger this ratio, the more operator resources can be freed and the overall capacity of the network increases. The network efficiency ratio is measured for both mobility models, 2 and 8 sessions per hour, and the two UMTS and WLAN coverages, respectively.
First, the efficiency is shown for RWP mobility. In Figure 69, the efficiency is constantly about 20%, which implies that the enabled ad-hoc feature saves one fifth of the network resources. If the number of sessions is increased as in Figure 70, the efficiency is even almost doubled.
Figure 71 and Figure 72 show the efficiency for RPGM mobility model with 2 and 8 sessions per hour and communicating pair, respectively. The group mobility model influences the impact on the efficiency compared to the random mobility in case of 2 sessions. The efficiency is more dependent of the on-demand feature, which is obvious because the traffic over infrastructure takes shorter to be sent without on-demand feature, while the peer-to-

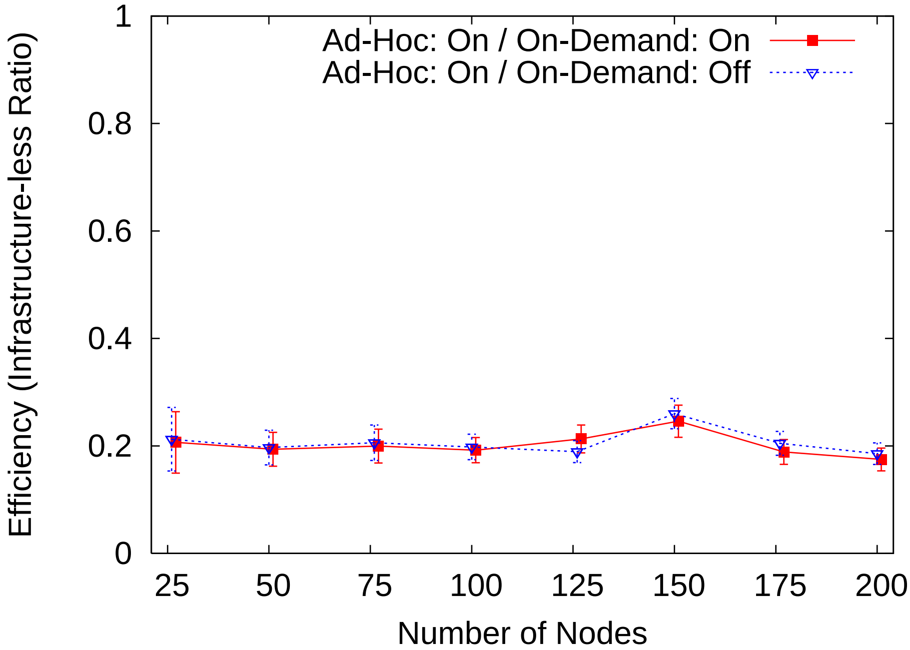

Figure 69: Network efficiency: RWP with 2 sessions per hour

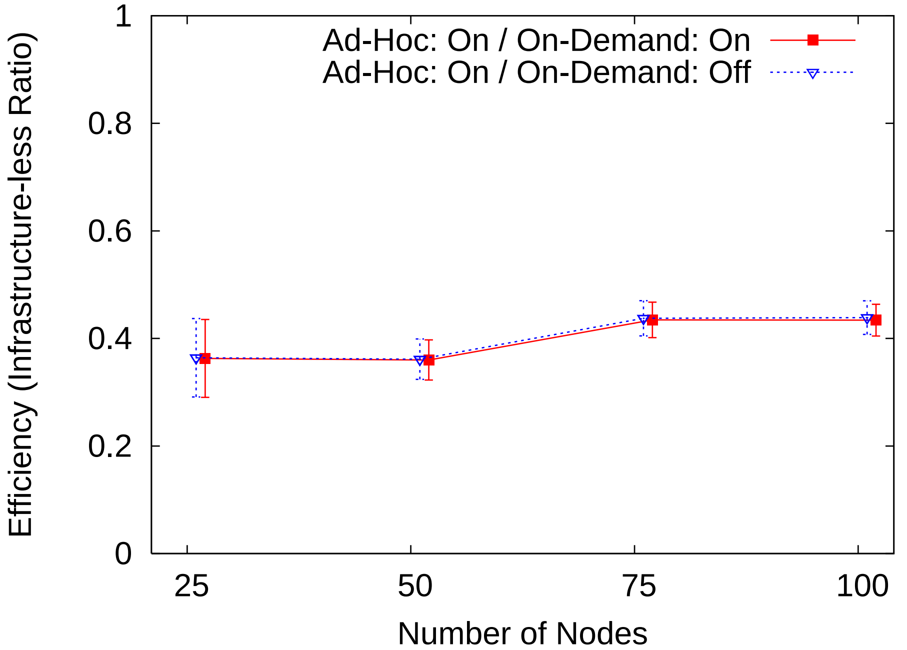

Figure 70: Network efficiency: RWP with 8 sessions per hour

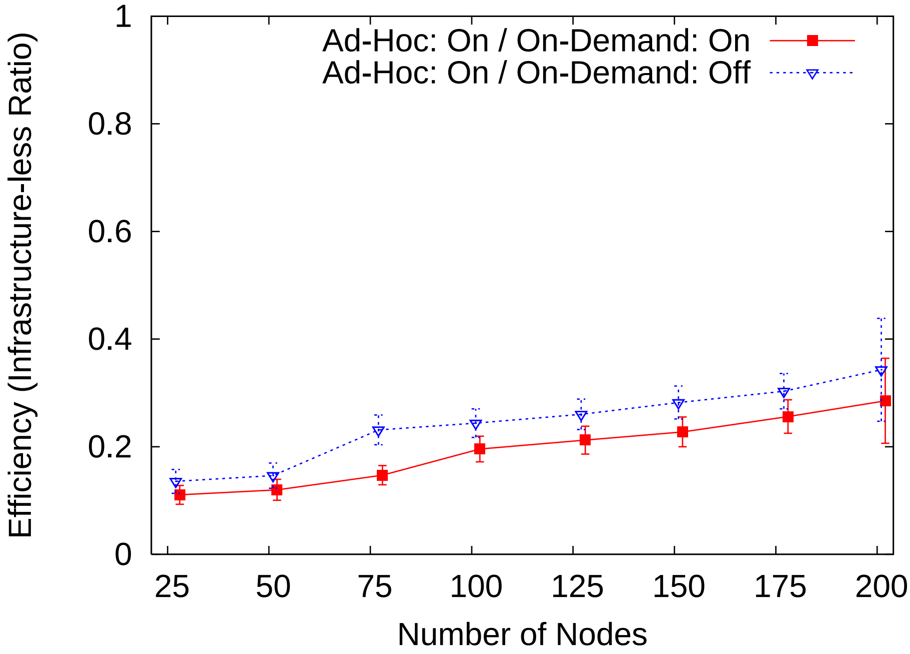

Figure 71: Network efficiency: RPGM with 2 sessions per hour

peer traffic amount is the same and the difference is therefore bigger. The same trend is also visible in the scenario with 8 sessions per hour; however the effect is not as strong because the overall traffic is higher.
The figures show that the ratio of infrastructure-less connections is not dependent on the number of nodes. On the average, the selection of the mobility model is only slightly influencing the ratio. Similarly, the influence of the on-demand feature is almost negligible.

5.5 Evaluation

The simulations show that the ad-hoc mode is clearly profiting from the RPGM mobility model, where nodes are likely to move within groups, sharing the same path. The characteristics of the RPGM are set on purpose to guarantee that the group members stay within the range of the infrastructure-less communication interface. Thus, whenever sender and receiver of a session happen to join the same group, a large part of the data is sent using the infrastructure-less technology providing high data rates. The nodes change the group with a pre-defined probability to assure a certain level of mobility of the individual node. Whereas in the simulations, the selection of the session end-points is done independently of the group membership, in reality there is most probably a higher probability that nodes moving in the same group also start sessions among them. This behavior would for sure be beneficial for the ad-hoc mode. On the other hand the on-demand feature is generally suffering from the group mobility offered by the RPGM. The members of the group traveling the same path have to compete for the available resources. The amount of sessions per hour also dramatically decreases the benefit of the on-demand mode. For higher session densities, where the period between the sessions is very short or even not present,

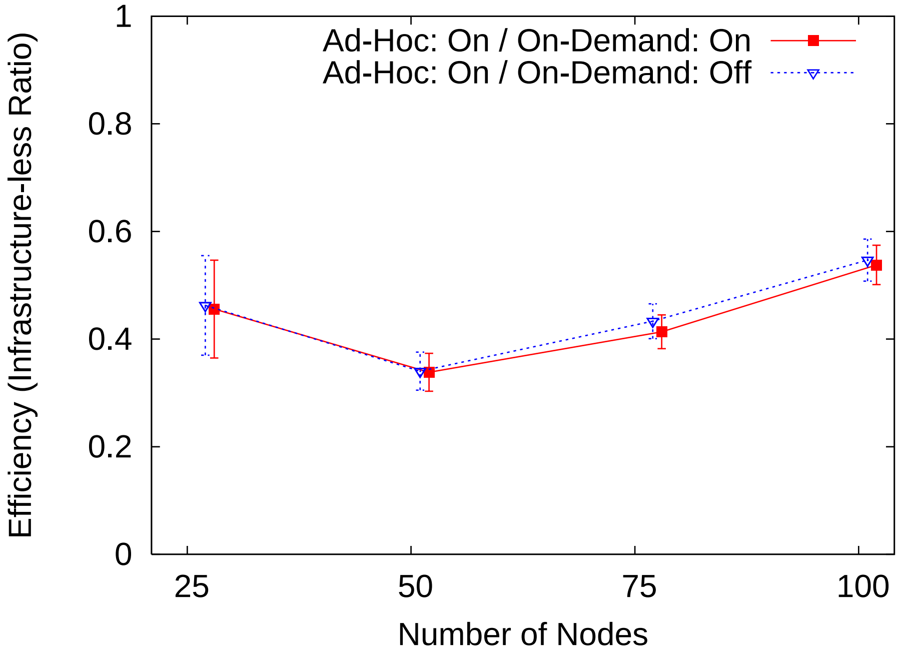

Figure 72: Network efficiency: RPGM with 8 sessions per hour

the on-demand mode fails to free network resources. However, the simulation results show that none of both newly introduced features ever resulted in a decrease of the network performance. Hence, there is no risk to enable both modes permanently. The simulations prove that the concepts of ad-hoc and on-demand mode can generally help to increase the throughput, the battery lifetime, the network efficiency, and decrease at the same time the session block probability, the session drop rate, and the session outage ratio of data sessions in heterogeneous networking environments.

6 Conclusion

6.1 Summary

As heterogeneous networking has become very popular in the last years and multimode devices have become available, the is interest in identifying the potential of such heterogeneous networking environment. To do so, simulations are a powerful mean. Today's network simulators do not support heterogeneity per se.
In this work a new network simulator is introduced which is designed to simulate heterogeneous networks with seamless handovers between WLAN, UMTS, GPRS, and peer-to-peer ad-hoc links. With this simulation tool, many scenarios with different seeds have been simulated to evaluate the benefit resulting from the on-demand and ad-hoc feature.
In the first simulation set, simulations with varying node density, mobility models, session density, bandwidth ratio, and coverage have been made and the results have shown that every scenario is profiting from the enabled ad-hoc and on-demand modes. The throughput was increased by a factor of four in average, whereas the energy consumption has been decreased by 20% for both features. The second evaluation set was made on varying number of nodes, mobility models and coverage ratios. The results have shown that the newly introduced modes improve the examined scenarios in terms of lower outage and delay times, lower block and drop probabilities, higher throughputs, higher efficiency, and lower network load. It can be stated that the introduction of the on-demand and ad-hoc features would have a positive impact for the operator as well as the user in terms of resource management and performance.
The introduction of ad-hoc link enabled network configurations should also be considered for homogeneous networks, e.g. public WLAN hotspots, UMTS cells, etc. Such ad-hoc links, which do not need any infrastructure, would relieve the probably overloaded WLAN access points and at the same moment reduce the costs on the user side. In the next few years it will most probably be possible to establish very fast peer-to-peer ad-hoc links over UWB, thus up to twenty times higher data rates can be expected and the added value of infrastructure-less connections will be even higher.

6.2 Future Work

A future task could include several extensions to the simulator such as lower layer complexity, additional access technologies, different handover strategies,and traffic patterns. Furthermore, a revision could be considered as the simulator contains already a lot of classes and a good idea would be to split it up in modules before adding new code.
As a big part of the time of this work was invested to build up the simulator not many simulations were run in the context of this work. Thus, the capabilities of the simulator were not exhausted so far. There are much more scenarios with different mobility models, traffic patterns, handover decision rules, etc. to simulate. Another interesting experiment would be the processing of real-world data, e.g. based on real traffic data of an operator.
From an architectural point of view, there are also some possible extensions. The introduction of peer-to-peer ad-hoc links, which should be built up spontaneously, requires the implementation of some mechanism to establish these connections securely. Additionally,

there should also be introduced an accounting strategy for operators in order to have a revenue out of the provided service.
A next step could be of course the consideration of a test implementation of the described ad-hoc and on-demand modes in a real environment.

7 Acknowledgments

I would like to thank Swisscom AG, who provided the material and the resources needed for my diploma work. Further, thanks go to all the people at Swisscom Innovations for their cooperation. I have been collecting many innovative experiences working on interesting tasks.
My special thanks go to Prof. Dr. T. Braun for making this cooperation with Swisscom possible and for his care during my diploma work. Sincere thanks also go to M. Danzeisen and Dr. D. Rodellar for the technical assistance and support during my diploma work.
Finally, I want to thank my family and my friends for their encouragement and support.

8 References

1. G.F. Riley, Ammar, Fujimoto et al., "Distributed Network Simulations using the dynamic simulation backplane", 2000, Georgia Institute of technology

2. Institute of Electrical and Electronics Engineers (IEEE) 802 LAN/MAN Standards Committee, http://www.ieee802.org, last visited 20.1.2006

3. P. Lescuyer, Bott, Frank, "UMTS",pp.236-242, 2002, ISBN 3-89864-141-4

4. C. Perkins, "RFC 2002 - IP Mobility Support", 1996, http://www.faqs.org/rfcs/rfc2002.html, last visited 20.1.2006

5. H. Soliman, "Mobile IPv6", Addison Wesley 2004, ISBN: 0-201-78897-7, Chapters 2.6 and 3.3

6. M. Danzeisen, D. Rodellar, S. Winiker, T. Braun, "Heterogeneous Networking facilitated by cellular Networks", Comtec journal, 2004

7. M. Danzeisen, D. Rodellar, T. Braun, I. Steiner, "Optimized heterogeneous spontaneous networking supplied by cellular mobile operators", March 2005, WSNET

8. M. Danzeisen, "Secured Mobile IP Communication", May 2001, Diploma work at the University of Bern

9. The Network Simulator - NS-2. University of California, Information Science Institute ISI, available on: http://www.isi.edu/nsnam/ns, last visited 20.1.2006

10. Qualnet Simulator - Scalable Networks Technology, available on: http://www.scalable-networks.com, last visited 20.1.2006

11. Opnet website, http://www.opnet.com, last visited 20.1.2006

12. Institute of Electrical and Electronics Engineers (IEEE) 802.11_21, available on: http://www.ieee802.org/21, last visited 20.1.2006

13. C. Bettstetter, H. Hartenstein, X. Pérez-Costa, "Stochasitic Properties of the Random Waypoint Model", ACM, 2004

14. BonnMotion - A mobility scenario generation and analysis tool, available on: http://www.cs.uni-bonn.de/IV/BonnMotion/, 2002-2005, last visited 20.1.2006

15. L.M. Feeney and M. Nilsson, "Investigating the energy consumption of a wireless network interface in an ad-hoc networking environment", Proceedings of the 20th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 01), Anchorage, AK, USA, Apr. 2001, pp. 1548 - 1557.

16. Merling U630 Wireless PC Card Modem, Lucent Technologies S.A., 2005, $http://www.lucent.com/livelink/090094038003841f_Brochure_datasheet.pdf$, last visited 1.3.2006

Appendix

A Integration of Swisscom Mobile Unlimited

At the beginning of the implementation process it was previewed to base the energy consumption simulation on an abstraction of an empirical energy consumption model from a real-world example. Therefore, some measurements have been made, which are presented in this paragraph.
The Swisscom Mobile Unlimited PCMCIA card has been developed to make it easier for the user to profit from the existing network technologies (UMTS, GPRS, WLAN) without being bothered with complicated configuration work. It makes use of the previously described protocols and features to detect available networks and process seamless handovers if necessary. EDGE has been integrated into the card lately.

Figure 73: Figure 81: Case study Unlimited, test setup

As the simulations should also show the benefit of the tested features in terms of the consumed energy, a test environment has been set up to measure the energy current in the PCMCIA Unlimited card for the use of the different network interfaces.
The test environment consists of a Dell Laptop Latitude D500, a Swisscom Mobile Unlimited card and an ampere-meter as shown in Figure 1. 12 measurements of 100 seconds at a sample rate of 0.5 seconds have been recorded and evaluated. The goal was to find out how much energy is used for the Unlimited card when using GPRS, UMTS, and WLAN for the different user activities: upload, download (both over http), browsing and being idle. The results are presented in Table 2.

Interface	idle	browse	upload	download
GPRS	3.081	6.4155	7.878	5.0115
UMTS	4.134	7.4295	7.02	7.8585
WLAN	6.006	5.9865	6.084	5.9865

Table 5: Average energy consumption in Watt for different interfaces and activities

The consumed power of the Laptop without running a PCMCIA card has also been measured as a reference value. Considering that a common Laptop consumes 17.55 to 33.15 Watt, the energy consumed by the Unlimited card is 17 to 28% of the totally consumed energy on average. The results show that especially WLAN consumes a lot of energy being idle whereas the other interfaces are already saving energy when no data is sent or received. The gained values provide a reference to evaluate simulation scenarios in terms of energy consumption.

B The HNS configuration file

```
###################################################
# PROTOSIM Config File     Overall Test scenario  #
###################################################
# Serie: nodes
# Simarea: 2000x2000
# Model: RWP
# Ad-hoc: on
# On-demand: always-on
# Seed: 8716
#

#Excel Path
path C:/Program/Microsoft Office/OFFICE11/EXCEL.EXE ;

starttime 1000
#the seed is for the random peer-to-peer sessions
seed 3468


###################################################################
#specify the mobility model
#mobility -f <filename> <Model name> <all the other params>
#for the load of manual files set
#mobility manual <filename> <simtime> <xdim> <ydim>
#see the file README_mobility for specific information
###################################################################
mobility -f rwp1 RandomWaypoint -d 4600 -i 1000 -n 20 -x 2000 -y 2000
-R 8716 -h 15 -l 1 -p 30


#mobility Statistics -f rpgm2 -r 50,100,150
#see stat file in util/<name>.stat

#mode [on-demand | always-on]
#default is always-on
mode always-on

#ad-hoc <off|on> means peer-to-peer
#default is on
ad-hoc on

#performance low_price | high_bw | best_signal
#default = best_signal
```

```
performance best_signal

#define P2P link capacity (per default it is 5200)
P2Pcapacity 11264

#Area x y rad (all n m), tech, max_users (-1 no matter) bwUL bwDL

#---------------------
#Simulation Area 2000x2000
#---------------------
Area 0 1000 1415 GPRS
Area 2000 1000 1415 GPRS
Area 750 350.48 288.6751346 UMTS
Area 1750 350.48 288.6751346 UMTS
Area 500 783.4927019 288.6751346 UMTS
Area 88 1456 57 WLAN
Area 88 202 57 WLAN
Area 1456 1456 57 WLAN
Area 1570 1114 57 WLAN
Area 1114 1000 57 WLAN
Area 886 202 57 WLAN
Area 1912 202 57 WLAN
Area 1114 1570 57 WLAN
Area 1000 88 57 WLAN
Area 1342 1684 57 WLAN
Area 1000 1456 57 WLAN

#pattern [random|manual]
#the pattern defines whether you have to set up traffic
#or just let decide the random machine for a pattern where
#nodes choose peers and traffic load randomly
#pattern must be defined
pattern manual
max_peers 1
#max_data is in KB (4000)
max_data 10000

#send event (always specify at the end of this config file)
#s <master node> ; <2nd> ; <third> ; .. <start> <data KB>
# traffic
s 10 ; 0 1225 43144
s 11 ; 1 3429 47830
s 12 ; 2 1936 4940
s 13 ; 3 1543 27577
s 14 ; 4 1698 97970
```

```
s 15 ; 5 2094 37314
s 16 ; 6 1867 32615
s 17 ; 7 1043 2856
s 18 ; 8 2819 54759
s 19 ; 9 3957 29565
```

C Implementation examples

C.1 Java implementation of the Wireless LAN model

```
public class WLAN extends Technology{

  /**
   * @param node
   * @param bs
 * @param futureBW will use resources in future but not until now
 * @param ongoing is an ongoing session in the same BS
   * @return the download capacity
   */
   public synchronized double getDL(Node node,BaseStation bs,
   boolean futureBW, boolean ongoing){
    double activeUsers = bs.getUserCountAll();
    activeUsers = Technology.getActiveUsers(true, node,activeUsers,
    futureBW, ongoing);
    double capacity = bs.getCapacity();
    if(activeUsers<=0)
    return capacity;
    double bwAvailable = capacity/(activeUsers+1);
    if(bwAvailable <= 0)
    return -1;
    return bwAvailable;
   }

   /**
    * @param node
    * @param bs
  * @param futureBW will use resources in future but not until now
  * @param ongoing is an ongoing session in the same BS
    * @return the upload capacity
    */
    public synchronized double getUL(Node node,BaseStation bs,
    boolean futureBW, boolean ongoing) {
     double activeUsers = bs.getUserCountAll();
     activeUsers = Technology.getActiveUsers(false, node,activeUsers,
     futureBW, ongoing);
```

```
    double capacity = bs.getCapacity();
    if(activeUsers<=0)
    return capacity;
    double bwAvailable = capacity/(activeUsers+1);
    if(bwAvailable <= 0)
    return -1;
    return bwAvailable;
   }
}
```

C.2 Java implementation of the UMTS model

```
public class UMTS extends Technology {

/**
 * @param node
 * @param bs
 * @param futureBW will use resources in future but not until now
 * @param ongoing is an ongoing session in the same BS
 * @return the download capacity
 */
public synchronized double getDL(Node node,BaseStation bs,
 boolean futureBW, boolean ongoing) {

double activeUsers = bs.getUserCountAll();
activeUsers = Technology.getActiveUsers(true,node, activeUsers,
  futureBW, ongoing);

double capacity = bs.getCapacity();
double maxUsers = bs.getMax_users();

if(activeUsers>maxUsers)
return -1;

if(384*(activeUsers)<=capacity){
return 384;
}else if(128*(activeUsers)<=capacity){
return 128;
}else if(64*(activeUsers)<=capacity){
return 64;
}else
return -1;
}
```

```
/**
 * @param node
 * @param bs
 * @param futureBW will use resources in future but not until now
 * @param ongoing is an ongoing session in the same BS
 * @return the upload capacity
 */
public synchronized double getUL(Node node,BaseStation bs,
 boolean futureBW, boolean ongoing) {
double activeUsers = bs.getUserCountAll();
activeUsers = Technology.getActiveUsers(false, node, activeUsers,
  futureBW, ongoing);
double capacity = bs.getCapacity();
if(activeUsers>bs.getMax_users())
return -1;

if(64*(activeUsers)<=capacity)
return 64;
else
return -1;
}
}
```

C.3 Java implementation of the GPRS model

```
public class GPRS extends Technology {

/**
 * @param node
 * @param bs
 * @param futureBW will use resources in future but not until now
 * @param ongoing is an ongoing session in the same BS
 * @return the upload capacity
 */
public synchronized double getUL(Node node,BaseStation bs,
 boolean futureBW, boolean ongoing) {

double activeUsers = bs.getUserCountAll();
activeUsers = getActiveUsers(false,node,activeUsers, futureBW, ongoing);

double maxUsers = bs.getMax_users();
double capacity = bs.getCapacity();

if(activeUsers>bs.getMax_users())
return -1;
```

```
if(42.8*(activeUsers)<=capacity)
return 42.8;
else if(21.4*(activeUsers)<=capacity)
return 21.4;

return -1;
}

/**
 * @param node
 * @param bs
 * @param futureBW will use resources in future but not until now
 * @param ongoing is an ongoing session in the same BS
 * @return the download capacity
 */
public synchronized double getDL(Node node,BaseStation bs,
 boolean futureBW, boolean ongoing) {

double activeUsers = bs.getUserCountAll();
activeUsers = getActiveUsers(true, node,activeUsers, futureBW, ongoing);

double maxUsers = bs.getMax_users();
double capacity = bs.getCapacity();
if(activeUsers>maxUsers)
return -1;

if(85.6*(activeUsers)<=capacity)
return 85.6;
else if(64.2*(activeUsers)<=capacity)
return 64.2;
else if(21.4*(activeUsers)<=capacity)
return 21.4;
else
return -1;
}
}
```

D Figures and Tables

List of Figures

1 Overview of the GSM . . . 5
2 CAHN connection establishment . . . 9
4 RWP . . . 11
3 Overview of mobility models . . . 12
5 RWP mobility node distribution [13] . . . 12
6 Gauss . . . 13
7 RPGM . . . 14
8 Manhattan . . . 15
9 Two differently scaled test scenarios . . . 19
10 Simulation procedure with pre- and post-processes . . . 22
11 Overview of packages in HNS . . . 22
12 Simplified class diagram of HNS . . . 23
13 The Graphical User Interface of HNS . . . 28
14 State Diagram of HNS . . . 30
15 Example grid for 4 by 4 nodes . . . 31
16 Measured connectivity of NS-2 and HNS compared . . . 32
17 Example of test scenarios . . . 32
18 Business scenario calculation . . . 38
19 Seamless handover to GPRS versus P2P only . . . 39
20 Seamless handover to UMTS versus P2P only . . . 40
21 Seamless handover to WLAN versus P2P only . . . 40
22 Scenario with five users . . . 42
23 Scenario with 5 users and a random number of sessions . . . 43
24 Scenario without handover to ad-hoc connections . . . 44
25 Scenario with handover to ad-hoc connections . . . 44
26 Test case 1: RPGM, no ad-hoc . . . 46
27 Test case 2: RPGM, ad-hoc . . . 46
28 Test case 3: RWP, no ad-hoc . . . 46
29 Test case 4: RWP, ad-hoc . . . 46
30 Test case 1: RPGM, no ad-hoc . . . 47
31 Test case 2: RPGM, ad-hoc . . . 47
32 Test case 3: RWP, no ad-hoc . . . 47
33 Test case 4: RWP, ad-hoc . . . 47
34 Test case 1: RPGM, no ad-hoc . . . 48
35 Test case 2: RPGM, ad-hoc . . . 48
36 Test case 3: RWP, no ad-hoc . . . 48
37 Test case 4: RWP, ad-hoc . . . 48
38 Test case 1: RPGM, no ad-hoc . . . 50
39 Test case 2: RPGM, ad-hoc . . . 50
40 Test case 3: RWP, no ad-hoc . . . 50
41 Test case 4: RWP, ad-hoc . . . 50

42 Throughput for varying node density . . . 53
43 Energy consumption for varying node density . . . 53
44 Throughput for varying session density . . . 54
45 Energy consumption for varying session density . . . 55
46 Throughput for varying bandwidth ratio . . . 56
47 Throughput for varying coverage . . . 57
48 Energy consumption for varying coverage . . . 57
49 Session block rate: RWP with 2 sessions per hour . . . 59
50 Session block rate: RWP with 8 sessions per hour . . . 60
51 Session block rate: RPGM with 2 sessions per hour . . . 60
52 Session block rate: RPGM with 8 sessions per hour . . . 61
53 Session drop rate: RWP with 2 sessions per hour . . . 62
54 Session outage: RWP with 2 sessions per hour . . . 63
55 Session drop rate: RWP with 8 sessions per hour . . . 63
56 Session outage: RWP with 8 sessions per hour . . . 64
57 Session drop rate: RPGM with 2 sessions per hour . . . 64
58 Session outage: RPGM with 2 sessions per hour . . . 65
59 Session drop rate: RPGM with 8 sessions per hour . . . 65
60 Session outage: RPGM with 8 sessions per hour . . . 66
61 Session throughput: RWP with 2 sessions per hour . . . 67
62 Session throughput: RWP with 8 sessions per hour . . . 67
63 Session throughput: RPGM with 2 sessions per hour . . . 68
64 Session throughput: RPGM with 8 sessions per hour . . . 68
65 Network load: RWP with 2 sessions per hour . . . 69
66 Network load: RWP with 8 sessions per hour . . . 69
67 Network load: RPGM with 2 sessions per hour . . . 70
68 Network load: RPGM with 8 sessions per hour . . . 70
69 Network efficiency: RWP with 2 sessions per hour . . . 72
70 Network efficiency: RWP with 8 sessions per hour . . . 72
71 Network efficiency: RPGM with 2 sessions per hour . . . 73
72 Network efficiency: RPGM with 8 sessions per hour . . . 74
73 Figure 81: Case study Unlimited, test setup . . . 79

List of Tables

1 Overview of the technologies (Values for Europe) . . . 7
2 Comparison of different handover technologies . . . 39
3 Bandwidths of technologies . . . 42
4 Used relative energy consumption values . . . 52
5 Average energy consumption in Watt for different interfaces and activities . . . 79