# On the Potential of Heterogeneous Networks

Marc Danzeisen, Torsten Braun, Isabel Steiner, and Marc Heissenbüttel

University of Bern, Bern, Switzerland,
{danzeis, braun, isteiner, heissen}@iam.unibe.ch

**Abstract.** Many research work is going on in the domain of pure ad-hoc networking and even more issues are raised due to the missing central infrastructure to properly manage resources, guarantee fairness, and provide security features. On the other hand lot of research effort is spent to increase the performance of infrastructure-based access networks to cope with the steadily increasing demand for broadband data. When making a step back, the most promising evolution of heterogeneous networking is the integration of both paradigms. Taking advantage of the well controlled cellular environment and the high capacity of ad-hoc and direct node-to-node communication. The resulting hybrid networks are incorporating the best of both worlds. The concept of *Cellular Assisted Heterogeneous Networking* (CAHN) provides a framework to offer convenient and secure management of heterogeneous end-to-end sessions between nodes. The introduced separation of the signaling and the data plane allows to switch on power demanding broadband interfaces like GPRS, UMTS, or even WLAN only, if actually required, i.e., data has to be sent or received. The proposed out-of-band signaling enables furthermore the integration of ad-hoc links to offer best performance whenever nodes are within vicinity. Extensive simulations show that both, the integration of ad-hoc links and the selective activation of high power broadband interfaces, can significantly increase the performance of heterogeneous networks.

## 1 Introduction

The layered structure of the ISO/OSI communication stack has enabled high flexibility by decoupling the underlying network technology from the applications to a certain extend. However, most of the applications are not independent of the communication characteristics provided by the underlying network. Applications often require a certain level of bandwidth, delay or security to work properly. Analyzing further these applications reveals that their requirements also depend on the user's situation. The level of mobility has an impact on the way applications (or services) are used. The capabilities of the devices used in the different situations is very much influencing the requirements on the service and therefore on the underlying network. A single network can not cope anymore with the different and often changing requirements. The development and deployment of the 3G cellular network clearly showed the complexity to build up one network, which fits all the requirements of nowadays and future mobile applications. To optimally meet the applications need, different communication technologies have to collaborate.

The motivation for such a collaboration is further discussed in [1], where the vision of being *always best connected* is introduced, connecting mobile nodes with the most appropriate communication technology. Dynamically assigning the most appropriate networking resources to each node, depending on its actual needs and capabilities, may increase the customer satisfaction and the overall performance of the heterogeneous network at the same time. The middleware presented in [2] and [3] is able to gather context information of the application layer to optimize the handover decision. However, network environment information is also permanently influencing the proper selection of the most appropriate communication technology. Mobility poses an additional challenge by constantly changing the network environment, e.g., used networks become unavailable or new ones get detected.

The concept of MIRAI [4–6] addresses this problem by defining dynamically one channel to be used for signaling information and negotiation of handover decisions. The authors proposed an agent based platform that provides location-based information on available access networks through a so-called basic access signaling, which is assumed to have a larger coverage than all other access networks. This concept is very beneficial, especially if the basic access signaling channel is a low power channel. In [7] it has been shown that such a signaling channel does not have to provide high data rates. MIRAI focuses on infrastructure-based access networks only. But communicating nodes can come close enough to establish direct links based on short range and infrastructure-less communication technologies. These links have the potential to deliver data rates which are orders of magnitude higher than infrastructure-based links will ever do. Especially in scenarios, where nodes are moving in groups and the probability of being within the range of direct communication is high, the average data rates can be considerably increased. This is the case for public transportation, battlefield scenarios, but also in smaller campus networks.

Extending the vision of being always-on and best connected to infrastructure-less communication technologies introduces a major conflict between the interests of a network operator and the ones of the end users. Offering better performance in terms of bandwidth and costs, it might become the first choice to exchange large amounts of data between nodes. At first glance, there is no motivation in keeping infrastructure-based connections offered by the operator in that case. However, simulations presented later in this paper are showing that the benefit of using direct links is considerably increased if infrastructure-based networks can take over the session whenever the direct link is lost. The usage of infrastructure-less links might also become economically interesting for the network operator. If the price of broadband connectivity is further falling, the operators will be forced to reduce the costs of their infrastructure. Introducing an intelligent resource management for heterogeneous network resources could help to reduce the cost of a data session. The possibility to offer connectivity through infrastructure-less technologies, which are already built in nowadays devices, may offer an interesting option.

The integration of both network paradigm is also beneficial for infrastructure-less communication technologies. Due to the missing authentication infrastructure, the provisioning of security features in pure ad-hoc networks is still challenging. In hybrid networks combining infrastructure-based and ad-hoc elements some of the problems can be solved. Therefore, efficient and secured signaling over infrastructure-based networks will probably be a key value proposition, even if the actual data is transported using ad-hoc networks free of charge.

The next section is addressing the possibilities of such a broadband on-demand mode and the integration of ad-hoc links enabled by the concept of *Cellular Assisted Heterogeneous Networking* in further detail. Detailed evaluations on the improvement potential of an intelligent end-to-end session management for the end user experience and for the operator are presented in the simulation section. The last section concludes the paper.

## 2 Cellular Assisted Heterogeneous Networking

When considering the interconnection of any two nodes, the optimal data path might be very heterogeneous. Depending on the available networks, the optimal end-to-end data path between the nodes can consist of infrastructure-based and ad-hoc links. Fig. 1 illustrates a scenario where two communicating nodes are moving abroad connecting to UMTS and Wireless LAN (e.g., Public Hotspot). Initially, node 1 and node 2 are communicating using the cellular network (step 1). Then the node 1 changes its point of attachment to WLAN and sends an update message to the node 2 (step 2). If now the nodes move towards the same WLAN access point, the session path can be optimized (step 3). The traffic is not further sent trough the Internet. If both nodes come close enough to each other, the session is switched to direct node-to-node communication (step 4).

The integration of infrastructure-less connections between communicating nodes to enhance the performance is very challenging. The missing infrastructure to properly and automatically negotiate required parameters to establish a secured and optimized direct links is imposing some major hurdles that have to be taken before users can benefit from higher performance without compromising security.

To enable such heterogenous networking enabling an always best connected experience throughout all available communication technologies including direct peer-to-peer links, a novel architecture and protocol was implemented and presented in [8, 9]. The system architecture proposed allows the independent routing of signaling and data related information over different communication technologies. Unlike the inband signaling of IP, where the data channel has to be established before the first signaling messages can be exchanged, the ability to first exchange signaling information on a dedicated signaling channel is highly beneficial in heterogeneous environments. The separated treatment of signaling and data related information allows optimal resource management. During the

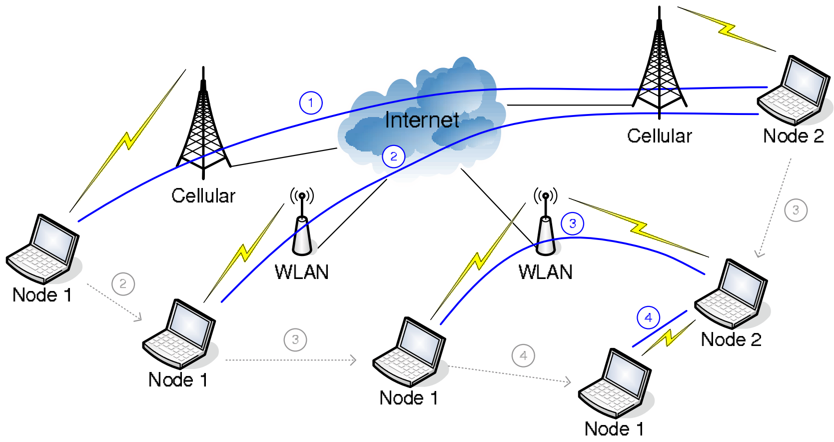


**Fig. 1.** Handover between Infrastructure-based and Infrastructure-less Communication

data session establishment the nodes can learn about the networking capabilities and actual environment, such as the currently available networks at the peer's location. Although the system architecture can be used to dynamically route the signaling messages over any type of communication technology, we argue that the use of the existing cellular mobile network as the primary signaling plane has several advantages. The well established and power optimized location, paging, and mobility management services can implicitly be shared for other communication technologies and networks, which are lacking such functionality. Especially the low power characteristics of the cellular mobile network is beneficial when addressing the power management of heterogeneous communication. With the help of the concept and architecture proposed in [10, 11], such a general signaling service could be offered to handle any type of transport network. To further increase the efficiency of the signaling of heterogeneous end-to-end sessions, we propose to do only the bootstrapping through the valuable cellular network, if no other secured IP connectivity is available. Furthermore, the signaling messages are sent inband after successful data channel setup, whenever possible. The end-to-end session management, together with the separation of signaling and data related information facilitates the concept of being reachable without wasting valuable networking resources. The ability to reach any node through the low power cellular network allows switching power demanding broadband data channels to sleep mode if no actual data session is going on. Reachability is no longer coupled with power demanding broadband IP connectivity. Waking up data channel interfaces only on-demand can considerably increase the power efficiency of the mobile devices. This *on-demand mode* also has the potential to increase the efficiency of allocated networking resources because the nodes are only attached to the network when data has to be transferred.

Furthermore, the security relation between the cellular subscriber and the network operator together with the roaming relations between the operators offer a secured communication channel which can be used to securely exchange the protocol messages. Especially if the establishment of infrastructure-less communication channels is considered, the secure exchange of configuration and security related parameters is absolutely mandatory to securely bootstrap the connections. Having established a secured initial communication between the communicating entities, all other parameters required to establish further communication channels can securely be negotiated. The concept of Cellular Assisted Heterogenous Networking (CAHN) provides the missing part to securely extend the scope of heterogeneous networking also to ad-hoc links. The ability of CAHN to securely bootstrap infrastructure-less communication between nodes enables the system to consider also direct node-to-node links when evaluating the most appropriate connection. Throughout the rest of this paper this feature is also referred to as *ad-hoc mode*.

## 3 Simulations

### 3.1 The Heterogeneous Network Simulator

We implemented a dedicated network simulator that allows the modeling of heterogeneous networks at a simplified level. The simulator does not account for any physical propagation medium properties or MAC layer functionality and simulates sessions between peer mobile nodes at the application level, i.e., the simulations are flow based. There are several reason why we did not use existing network simulators (ns2, Qualnet, OpNet, etc.). Among others they do not provide appropriate support for the simulation of heterogeneous networks with dynamic vertical handovers during runtime, end-to-end communication between nodes using different wireless technologies simultaneously, and switching between infrastructure and ad-hoc mode of operation. Furthermore, these simulators either do not yet implement certain wireless technologies, e.g., GPRS in Qualnet, or implement different technologies for different incompatible versions, e.g., UMTS for ns-2.26 and GPRS for ns-2b7a. The necessary modifications to the network simulators to enable the simulations of the intended scenarios would require a tremendous implementation work and is out of scope of this paper.

The transmission ranges for different wireless technologies are modeled as circles with varying radiuses with respect to their characteristics, e.g., small radius for broadband technologies such as WLAN and larger radius for 2.5 and 3 G technologies with narrower bandwidth such as GPRS and UMTS. Even though the simulator does not take into account the lower layers of the protocol stack, it allows the estimation of the possible benefits of the two main features of CAHN in heterogeneous networking environments, namely the ad-hoc and on-demand mode.

The different access technologies can be defined with limited capacity and adaptive data rates delivered to each node depending on the actual load of the cell or access point as described below. To model the bandwidth assignment for a UMTS node, we considered the up- and downlink channel separately. The uplink bandwidth is statically set to $64\,kbit/s$ per node, independently of the assigned downlink bandwidth. The overall capacity $m$ offered by the base station is assigned to the attached nodes $n$ until there is no capacity left to assign.

$$\text{bandwidth}_{up} = \begin{cases} 64\,kbit/s, & \frac{m}{n} \geq 64 \\ 0\,kbit/s, & \text{otherwise} \end{cases}$$

The downlink bandwidth is equally distributed to the attached nodes. The UMTS provides different bandwidth rates, namely $384\,kbit/s$, $128\,kbit/s$, and $64\,kbit/s$. If $n$ is the number of nodes and $m$ is the maximum capacity offered by the base station, the assigned bandwidth is modeled as follows:

$$\text{bandwidth}_{down} = \begin{cases} 384\,kbit/s, & \frac{m}{n} \geq 384 \\ 128\,kbit/s, & \frac{m}{n} \geq 128 \\ 64\,kbit/s, & \frac{m}{n} \geq 64 \\ 0, & \text{otherwise} \end{cases}$$

GPRS is modeled based on TDMA slots. The coding scheme (CS) is statically set to CS4, which is providing $21.4\,kbit/s$ per slot. In the model we assume that CS4 can be used for up- and downlink independent of the distance between the node and the base station. We further assume class 10 devices, allowing 4 downlink and 2 uplink slots maximum. The number of assigned slots is depending on the availability of slots. The network tries to assign the maximum number of slots supported but dynamically adapts the assignment to guarantee uniform distribution of the available resources. For the uplink this results in the following slot assignment model, where $n$ is again the number of nodes and $m$ the maximum capacity offered by the base station.

$$\text{bandwidth}_{up} = \begin{cases} 42.8\,kbit/s, & \frac{m}{n} \geq 42.8 \\ 21.4\,kbit/s, & \frac{m}{n} \geq 21.4 \\ 0\,kbit/s, & \text{otherwise} \end{cases}$$

The downlink slot assignment is modeled similarly, but allowing up to 4 slot per node.

$$\text{bandwidth}_{down} = \begin{cases} 85.6\,kbit/s, & \frac{m}{n} \geq 85.6 \\ 42.8\,kbit/s, & \frac{m}{n} \geq 42.8 \\ 21.4\,kbit/s, & \frac{m}{n} \geq 21.4 \\ 0, & \text{otherwise} \end{cases}$$

The medium access mechanism of WLAN is aiming at the provisioning of equal bandwidth for all attached nodes. Unlike in GPRS or UMTS, the up- and downlink are not treated separately. Sending and receiving nodes are competing for the same medium. To simplify the modeling of WLAN we assume that no collisions occur and the maximum available capacity can equally be assigned to the nodes without loss because of collision recovery mechanisms (e.g. backoff). The model used for resource assignment for WLAN nodes is consequently as follows, where $n$ is the number of nodes competing for the medium and $m$ the overall capacity of the medium:

$$\text{bandwidth}_{up/down} = \frac{m}{n}, kbit/s$$

### 3.2 Simulation Scenarios

The simulation area is set to $2000\,m$ x $2000\,m$ where the nodes move according to the random waypoint mobility model with a speed in the interval $[1, 15]m/s$ and a pause time of $30\,s$. The simulation time was set to $4600\,s$ including a $1000\,s$ warm-up phase for the mobility model to reach a stable state. Random sessions are established between pairs of nodes, where the session arrival rate is Poisson distributed with a rate of two sessions per hour and source-destination pair. The amount of data transferred per session is Pareto distributed between 10 KB - 100 MB.

Each node always uses the available wireless technology with the highest bandwidth, i.e., a vertical handover occurs whenever a nodes moves into the range of a technology with a higher bandwidth. Consequently, the effective session transfer rate is the minimum bandwidth of the technologies, currently used by the two communicating nodes.

Three different wireless infrastructure-based technologies, namely GPRS, UMTS, WLAN, are deployed over the simulation area. Furthermore, there is an infrastructure-less wireless technology that allows for peer-to-peer communication for which we also used WLAN. The overall capacity for GPRS nodes is set dynamically according to the number of nodes for each simulation. Assuming that network operators deploy enough bandwidth to serve all nodes, which equals $n$, with the minimal data rate of one TDMA slot for both, the up- and downlink, we define the overall capacity of GPRS as $n$ up- and $n$ downlink slots. These slots are uniformly distributed among the GPRS cells, resulting in blocked sessions if the nodes are not equally distributed among the cells. In all simulation scenarios GPRS covers the whole simulation area and the coverage of UMTS is set to 80 % and the WLAN coverage to 10 % of the overall simulation area. UMTS cells with coverage radius of $450\,m$ are supposed to offer a maximum capacity of $1024\,kbit/s$, which is uniformally distributed among the nodes according to the model previously defined. Analogously, the transmission range of WLAN is set to $150\,m$ and the maximum capacity ($m$) per access point was set to $11\,Mbit/s$ as provided by the 802.11b standard.

### 3.3 Simulation Results

We analyze the influence of CAHN on the overall network performance. The simulation results have been analyzed in terms of network load and efficiency, throughput, and session block and drop rates. The different terms are defined in the following sections. All simulation results are given with a 95 % confidence interval. We also conducted simulations with higher session arrival rate, lower coverage of UMTS and WLAN, and other mobility models. However due to lack of space, the results are not given in this paper.

**Network Load** The overall network load is calculated based on the load of each technology. Each load is weighted according to the coverage provided by that specific technology. This reflects the fact that the overall network load is mainly dependant on the load of technologies serving a large area. The network load is evaluated for different numbers of nodes and sessions. Both CAHN features, the ad-hoc and on-demand modes, are enabled and disabled to analyze the impact on the overall network load. Figure 2 shows the four resulting network loads as a function of the number of nodes. The ability to liberate network resources by

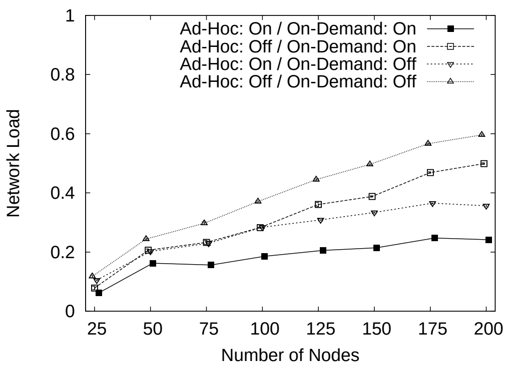


**Fig. 2.** Network Load

switching ongoing sessions to infrastructure-less technologies whenever possible reduces the network load by up to 24 %. Thus, the network is able to serve more nodes with the same capacity if the ad-hoc feature is enabled. The on-demand feature further increases the number of nodes that can be served. With the on-demand feature, inactive nodes do not occupy cellular network resources. The less sessions the nodes have, the higher the resource saving potential of this feature. With 2 sessions per hour about up to 12 % of the network resources can be liberated due to the on-demand feature and eventually assigned to other nodes.

**Network Efficiency** To further measure the influence when the ad-hoc mode is enabled, we introduce a new metric called network efficiency. With regards to the ad-hoc feature, we define the network efficiency as the ratio between traffic sent using ad-hoc links and the overall traffic sent by the nodes. This indicates how much traffic the network can offload to the direct ad-hoc links. The bigger this ratio, the less operator resources are used to transfer the session data. We additionally run simulations with an arrival rate of eight sessions per hour as depicted in Fig. 3. We can observe that the efficiency is not dependant

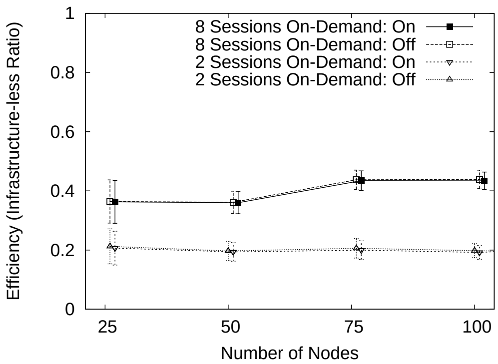


**Fig. 3.** Network Efficiency

on the number of nodes and is approximately 20 % and 40 % for the lower and higher session arrival rate, respectively. Similarly, the influence of the on-demand feature is almost negligible. However, for higher number of sessions the efficiency increases by more than a factor of two. This is mainly due to the high network load, which decreases the availability of the infrastructure-based links and thus their average throughput per source-destination pair, and consequently increases the impact of the data transferred using the high data rates offered by the infrastructure-less connections.

**Session Block Probability** Whenever a node is not able to start the session because of missing available network resources this is considered as a session block. The node continuously tries to start the data transmission until either another node releases resources in the congested cell or the node moves to another cell with available capacity. If the preferred technology is not available, the node connects to any other available networking technology, independent whether it is the most appropriate for that actual session. Thus, the network overrides the node's preference if the overall network performance can be increased. The network is able to serve approximately 75 nodes without blocking any resource request as depicted in Fig. . The nodes are therefore distributed to all available networks. The network forces the nodes to switch to another technology if this

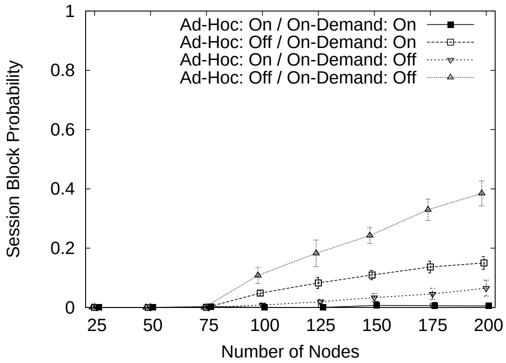


**Fig. 4.** Session Block Rate

increases the overall number of servable nodes. Depending on whether the ad-hoc and on-demand mode are enabled, the increase of the session block probability is steeper for a higher number of nodes. When both features are disabled, the probability for a block session is approximately 40 % for 200 nodes. The on-demand mode reduces that probability to less than the half. The ad-hoc mode has even a bigger impact on the session block probability. Together with the on-demand feature, the blocking probability can be kept close to 0 for up to 200 nodes.

**Session Drop Rate** Whenever a communicating node comes into a congested cell, the ongoing session is dropped. The cell is congested if according to the resource assignment model, the remaining capacity of the base station is smaller than the minimum assignable bandwidth.

Figure 5 shows the number of drops that occur during a session. If more than 75 nodes are using the network, the sessions get continuously dropped, frequently more than once, i.e., sessions are interrupted several times before completing. If there are more than 125 nodes, the average number of drops per session reaches its maximum.

During the time when a node is not connected because it suffered from a dropped call, no further drops can occur, which results in a decreasing number of drops per session again because the simulation period is limited and unfinished session are stopped at the end of the simulation. If the network load is high enough, the sessions do not terminate during the simulation period and tend thus to have less drops. In this stage, the network is not able to serve additional sessions, which further decreases the average number of drops per session. With the on-demand mode the drops can be avoided completely due to the released network resources between the sessions. The ad-hoc feature considerably reduces the drop rate but is not able to avoid all session drops.

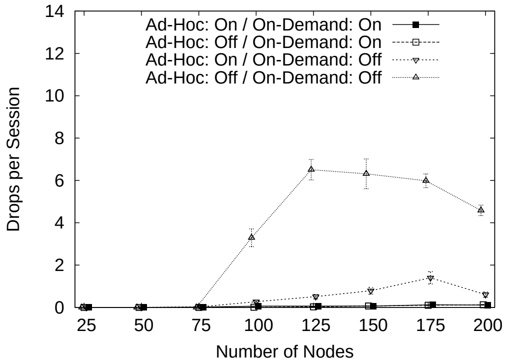


**Fig. 5.** Session Drop Rate

**Session Throughput** The amount of data transmitted is simply derived from the time attached to a certain technology and its bandwidth. The average session throughput includes the delays and outages imposed by the session blocking and dropping, which may occur if nodes have to share the network capacity. These delays increase the duration of a session and, thus, decrease the average throughput of the session. Furthermore, the actual throughput is decreased whenever the network has to perform load balancing. If a node moves into a congested UMTS cell, the network assigns GPRS instead. The ad-hoc mode is much more beneficial for the the throughput than the on-demand mode is as shown in Figure 6 and can increase the data rate from approximately 100 kbit/s to more than 150 kbit/s. This is mainly due to the high data rates offered by the ad-hoc links which provide much higher bandwidth links. The on-demand mode only marginally increases the average session throughput by freeing network resources from inactive nodes which increases the probability for transmitting nodes to use an available technology with a higher bandwidth.

## 4 Conclusion

In this paper we addressed the potential impact of the proposed CAHN concept on the performance of heterogeneous networks. We first introduced the two main features offered by CAHN, namely the ability to use power demanding broadband communication technologies only on-demand and bootstrap ad-hoc connection between nodes, whenever they are within the range of direct communication. We then presented the Heterogeneous Network Simulator, which allows the simulation of data sessions beyond the boundaries of communication technologies. The simulation results shows that the two features of CAHN enable the efficient use of available network resources. In certain scenarios the network load can be decreased by about 50 %, which results in lower session drop and blocking rates. Furthermore, the throughput can almost be doubled, depending on the number of nodes. These results motivate the tight collaboration of var-

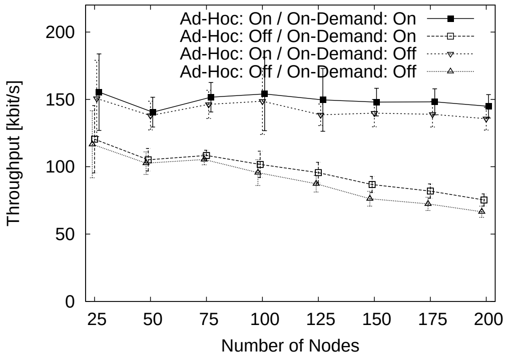


**Fig. 6.** Session Throughput

ious communication networks, including both, infrastructure-based and ad-hoc technologies.

## References


1. Gustafsson, E., Jonsson, A.: Always Best Connected. IEEE Wireless Communications Magazine **10** (2003) 49–55
2. Calisti, M., Lozza, T., Greenwood, D.: An agent-based middleware for adaptive roaming in wireless networks, AAMAS workshop on agents for ubiquitous computing, new york (2004)
3. Bhargava, B., Wu, X., Lu, Y., Wang, W.: Integrating heterogeneous wireless technologies: a cellular aided mobile ad hoc network (CAMA). Mob. Netw. Appl. **9**(4) (2004) 393–408
4. Wu, G., Mizuno, M., Havinga, P.J.M.: MIRAI architecture for heterogeneous network. IEEE Communications Magazine **40**(2) (2002) 126–134
5. Inoue, M., Mahmud, K., Murakami, H., Hasegawa, M.: MIRAI: A solution to seamless access in heterogeneous wireless networks. In: Proceedings of the IEEE ICC. (2003) 1033–37
6. Inoue, M., Mahmud, K., Murakami, H., Hasegawa, M., Morikawa, H.: Novel out-of-band signaling for seamless interworking between heterogeneous networks. IEEE Wireless Communications Magazine **11**(2) (2004) 56–63
7. et al., K.M.: On the required features and system capacity of basic access network in the MIRAI. In: WPMC. (2001) 1199–1204
8. Danzeisen, M., Braun, T., Rodellar, D., Winiker, S.: Heterogeneous communication enabled by cellular operators. Vehicular Technology Society (2006)
9. Danzeisen, M., Braun, T., Rodellar, D., Winiker, S.: Implementation of a cellular framework for spontaneous network. In: Proceedings of IEEE WCNC. (2005)
10. Danzeisen, M., Braun, T., Rodellar, D., Winiker, S.: Heterogeneous networking establishment assisted by cellular operators. In: Proceedings of MWCN. (2003)
11. Danzeisen, M., Rodellar, D., Winiker, S., Braun, T.: (Heterogeneous networking facilitated by cellular networks)