Dynamic Resource Management for Coverage and Capacity Enhancement in Cellular Systems with Infrastructure Based Multihop Relaying

Hang Zhang, Derek Yu, Shalini Periyalwar

Wireless Technology Labs, Nortel Networks, Ottawa

Abstract— Dynamic resource management strategies addressing active set management of BS and relays, dynamic resource partitioning and dynamic routing are examined for use in cellular infrastructure based multihop relaying systems. Simulation results demonstrate improved coverage with full queue user data rate distribution, along with a small increase in throughput. The average web page download rate is also shown to improve significantly, reflecting a reduction in end-to-end packet delay.

***Index Terms* — Cellular Systems, Multihop, Resource Management**

Introduction

Multihop relaying within the cellular infrastructure has been receiving considerable attention in WWRF [1] - [2] as well as in recent literature [3] - [5]. Multihop relays may be considered as a complement to the recent advances in air interface design and antenna processing technology which yield significantly higher spectral efficiencies than conventional 3G technologies. Multihop relays will enable the realization of these high spectral efficiencies over a wider area of the cell.

Relays used in cellular networks may be infrastructure based, involving the use of fixed equipment to perform relaying or may leverage UEs within the network which are equipped with relaying functionality. Infrastructure based relaying has some advantages over User Equipment (UE) based relaying. Infrastructure based relaying requires the deployment of relay nodes, which implies an increase of infrastructure cost. However, it will minimize the additional enhancements in the UE, which are particularly significant in an FDD system, thus containing the UE cost as well as conserving its power. If relaying is done through the UE, every UE needs to have relaying capability. Also, the mobility of UEs would weaken the stability of the relay-UE link, and the coverage performance would depend on the density of UEs within the cell. In infrastructure based relaying, the nodes can be carefully located to address the need for hot spots serving high speed data users or the existence of deep shadowing spots within the cell. Furthermore, infrastructure based relays can be designed such that they can be added to a cell to enhance coverage without impacting legacy UEs, and only requiring some software enhancements in the access network.

In [1], single hop conventional transmission and two-hop time-domain relaying are compared for different frequency reuse parameters, and it is shown that capacity and coverage gains are possible with simple two hop relaying, with increasing gains in larger reuse. A summary of multihop relaying research in WWRF projects is provided in [2]. Walke et al [3] propose a service architecture based on SIP for multihop relaying. Coverage enhancement in cellular networks with multihop relaying is addressed in [4]-[5]. Multihop relaying has also been examined for range extension in a number of other systems, e.g., [6] - [8].

Some important aspects of cellular infrastructure based multihop relaying have not been adequately covered in the literature. It is the objective of this paper to address dynamic resource management strategies in cellular infrastructure based multihop relaying for the following key areas, supported by simulation results for a system with nomadic users:

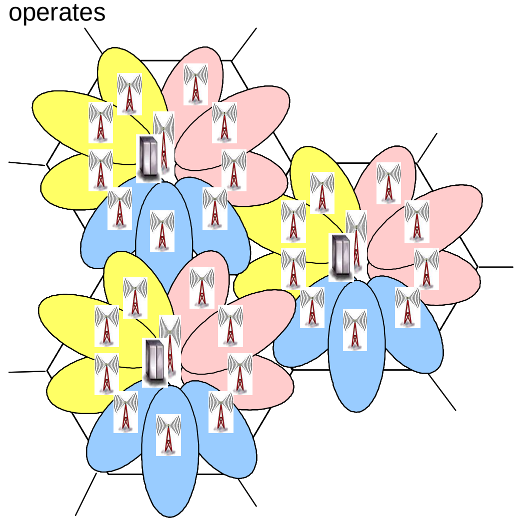
1. active set management
2. dynamic resource partitioning between BS-relay and relay-UE links
3. dynamic routing

In a multihop cellular system where the cellular spectrum is being reused for the relay to UE hop, it is important to ensure that the resources are assigned efficiently. These resource management aspects are critical in ensuring that resources are allocated not only to get the highest net link rate for BS-UE communication, but also the lowest delay so as to enable support of services with different QoS requirements.

The paper is structured as follows. Following a description of the system model used in the paper, dynamic resource management features are articulated. Active set management, dynamic resource partitioning and dynamic routing are addressed. Other resource management issues relating to relaying with mobility which are also discussed include efficient handling of MAC layer retransmissions and flow control during handoff. The performance improvements that the dynamic resource management schemes provide for nomadic users are demonstrated with end-to-end performance assessment for full-queue (FTP) and web browsing traffic.

System Model

An FDD based multi-beam[9] cellular system with universal (N=1) reuse is assumed, with relays deployed in all the beams in the system as shown in Figure 1. The multi-beam system may be considered equivalent to a multi-beam system, based on the antenna configuration. The focus of this work is on the downlink to the UE. The BS-relay and the relay-UE links employ the same frequency in a Time Division Multiplexed manner (Figure 1) where a fixed partitioning of resources between the BS-UE link, BS-Relay link and Relay-UE link is illustrated. The physical layer is assumed to be OFDMA, such that multiple users could be supported in one transmission unit. Thus, the relay operates



Slot	1 BS to UE	2 BS to Relay	3 Relay to UE
Base	Tx to UE	Tx to Relay	IDLE
Relay	IDLE	Rx from BS	Tx to UE
UE	Rx from BS	IDLE	Rx from Relay

Figure 1 System model and resource assignment map with fixed partitioning for the cellular system with relays

on the same frequency but in orthogonal TDM slots to communicate with the BS and UE respectively. Therefore, the UE is agnostic to whether the information was transmitted directly from the BS or through the relay. Proportionally fair scheduling [10] is assumed to be resident in the Base Station, where the relay transmissions slots are also defined. The relay incorporates a simple FIFO scheduler.

Dynamic Resource Management Features

With the objective of maximizing the performance of the cellular system with fixed relays, some dynamic resource management options are discussed below.

Active Set Management

Typically, the UE's active set comprises the BSs with the strongest pilot measurement. Similarly, for the relays, the relay to BS association is determined by the best BS pilot seen by the relay (termed the relay's parent BS). In order to find the best link to a UE however, the channel quality information on the BS-UE link and the relay-UE link is required. Measuring the channel quality for the relay to UE link would require separate pilot signals for the relays. The active set could then comprise of both BS pilots and relay pilots. Then the parent beam for the UE can be chosen in one of two ways:

- based solely on the best pilot among BSs in the active set, or
- based on the best of all the BS and relay pilot measurements.

In the latter case, it is possible that, for a given UE, if the relay pilot is the strongest of relay and BS pilots, the relay's parent BS's pilot may not necessarily be in the top of the UE's active set. Then the question is whether it is more beneficial to associate the UE with the BS-relay pair for the best relay pilot or the BS-relay pair for the best BS pilot, for that UE. If the signal strength from the relay's parent BS is adequate to support the control channel messaging to the UE, it makes perfect sense to assign the UE to the BS-relay pair for the best relay pilot. Otherwise, such a selection would require that the control signaling also be transmitted via the relay.

Dynamic Resource Partitioning

A simple approach to partitioning of resources in the multihop system is fixed partitioning, by allocating adjacent time slots to direct BS-UE transmission, BS-relay transmission and relay-UE transmissions, respectively, as shown previously in Figure 1. With dynamic resource partitioning, the resources are divided into two units and proportionally allocated to the BS-UE/relay transmissions and to relay-UE transmissions. The size of these units may be determined by the BS scheduler by taking into account the relay's transmission needs based on the relay-UE link conditions as well as the BS-UE/relay link conditions. Figure 2 illustrates an example of the dynamic resource partitioning scheme. It is possible to generalize this arrangement to have N BS-Relay/UE slots and M Relay-UE slots.

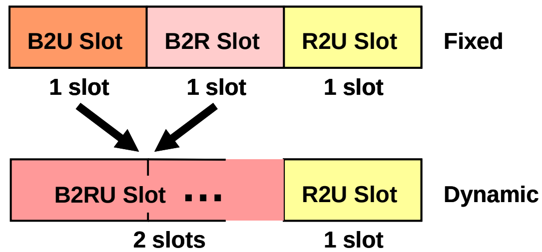


Figure 2 Dynamic Resource Partitioning Example

Dynamic Routing

A number of key criteria need to be considered to ensure the optimal routing of the UE's traffic either directly from the BS to the UE or through a relay.

<u>Resource Usage Efficiency:</u>

An important parameter for the routing decision is the net resource usage for the data rate required by the UE with the choice of the BS-UE link or the BS-relay-UE link. Although the available spectrum is divided amongst the multiple hops, the capacity is not necessarily compromised if the relaying is selective, in that the best path is dynamically chosen between the direct BS-UE link and the BS-relay-UE link to the user. A further optimization in routing is achieved if both routes (Direct, multihop) can be used to the UE when the UE is in a favorable position with respect to both of them. We classify this option as dual path dynamic routing.

<u>QoS Based Routing:</u>

The relaying operation with the proposed FDD/TDM mode introduces delay due to the store-and-forward nature of the relay. This would be problematic when the data packets are from an application that has a very tight delay requirement, such as interactive voice. In this case the path with least end-to-end delay may be chosen to be the direct BS-UE link. For delay tolerant services, the BS-Relay-UE could be selected over the BS-UE link. Thus routing in the cellular multihop system can be designed to meet the QoS requirement of different users/applications by incorporating the delay and throughput requirement into the routing decision.

Relay Buffer Overflow Prevention:

Each relay would buffer the data received from the BS until it is scheduled to be transmitted. The routing algorithm would have to consider the buffer occupancy of the relay node so that the traffic may be redirected to the second best path if the relay buffer in the best path is about to overflow. The impact of this factor on the uplink signaling load should be carefully examined as well.

User Mobility - Handover to Another Relay

When a UE moves in a cell, the best path to the UE would change dynamically. The routing algorithm needs to track the change of applicable metrics to update the decision for the best path to a specific UE. If necessary, the handover to neighbor relays should be handled by the routing algorithm. This may be done by keeping track of the channel quality measurement for each UE, both from relays and the BS.

It seems to be reasonable to locate the routing algorithm in the BS rather than in the relay when the number of relays or the number of hops is not large. With many relays and more than two hops, it may be worthwhile to consider letting the relays decide the forwarding path to reduce the computational burden at the base station. The details would require careful assumptions and performance simulations.

Other Design Considerations

While the major focus of this paper is on evaluating the performance of the system in a nomadic environment, there are some other important design considerations that need to be addressed for mobility.

Efficient Handling of MAC Layer Retransmissions:

In order to keep the design of the relays simple it is proposed that the relays only contain minimal layer 2 functionality. In this case, the radio link protocol extends from the BS to the UE and any MAC layer retransmissions will be required to be executed from the BS to the UE even if the errors occur only in the last hop. This is a costly use of resources and it is possible to consider a solution where the retransmissions occur only from the error-ed link, without replicating all the Layer 2 functions in the relay nodes.

Flow Control:

It is likely that a relay may not have completed the transmission of data to a UE, before the UE is redirected to a different relay or to the BS. Two events are likely to trigger a need for flow control:

- **Unnecessary packet retransmission with dynamic route selection:** With dynamic route selection, a UE may sometimes become connected directly to the BS, before the relay has completed its download of the buffered data to the UE. If there is no appropriate measure in place, the UE may issue a L2 NACK immediately, causing unnecessary retransmission.
- **Handoff:** If a UE performs a handoff to a neighbor beam or another relay while its data buffered in the old relay is still being downloaded, the residual data will have to be discarded by the UE's old relay. This lost data may need to be retransmitted by the new beam causing delay and resource wastage.

Performance Evaluation

Some of the proposed dynamic resource management strategies have been evaluated by simulation. An integrated Opnet based end-to-end simulation for the cellular infrastructure based multihop relaying system has been performed. The simulated system block diagram is shown in Figure 3 on the next page. The end-to-end full-queue (FTP) and web browsing application traffic generation is modeled with upper layer protocol stack (TCP/IP) and MAC functionality with Proportionally Fair (PF) scheduling at the BS and with FIFO scheduling at the relay. The simulations are performed in a multi-beam cellular system with frequency reuse of one and with two-hop communications between BS and UE. Up to two relays per beam are considered. The simulation results are generated for the center cell of a system with 18 surrounding cells (2 tiers) to model the interference. The relays employ a 60 degree antenna facing

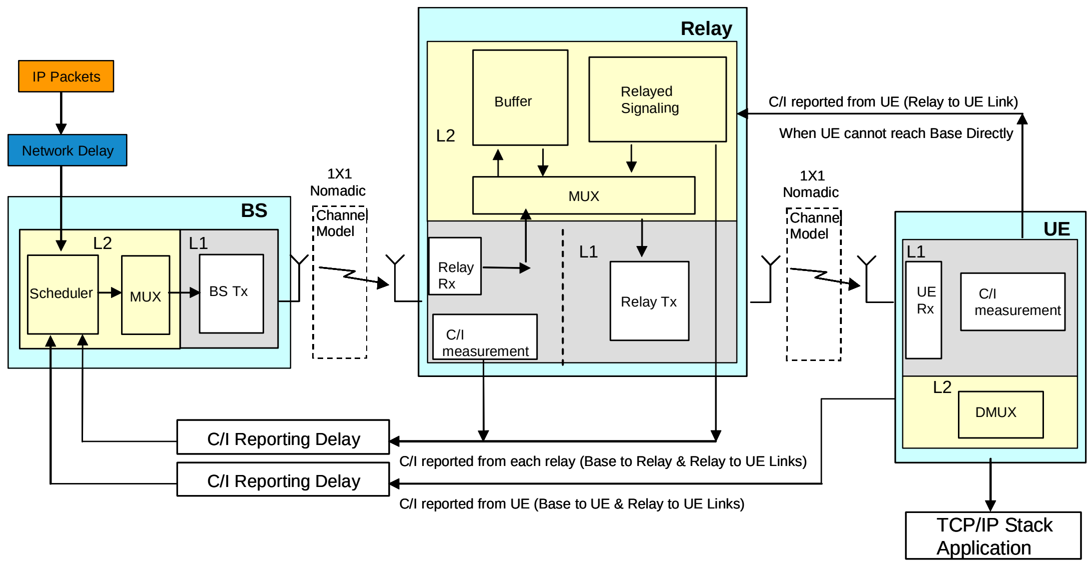


Figure 3 System Simulation Block Diagram

the BS and a 120 degree antenna facing away from the BS. Appropriate path loss and fading models are used at 1.9 GHz for the BS-relay (IEEE802.16 Type C Channel Model [11]), relay-UE (SCM channel model [12]) and BS-UE (Cost 231 W-I model) links.

A key assumption is that the dynamic resource partitioning boundary is a system wide parameter. Intra-cell & Inter-cell interference are modeled according to the resource partitioning assumptions:

- BS to UE transmission: Interference at UE only from BSs
- BS to Relay transmission: Interference at Relay only from BSs
- Relay to UE transmission: Interference at UE only from Relays.

Channel quality C/I measurement and reporting are modeled, and the physical layer behavior is modeled by mapping the measured C/I to a proprietary modulation and coding set.

The following metrics are used in the performance assessment:

- Data Rate Coverage Related
    - UE geometry C/I statistics
    - Full queue user rate distribution
    - Web Page download data rate distribution
- Throughput Capacity
    - Aggregate throughput

The Active Set Management is modeled to include both BS and relay pilots. In modeling the dynamic resource partitioning, a single slot is allocated for relay to UE transmission and two slots are allocated for BS - relay or UE transmission such that the resource is shared dynamically between relay and UE transmissions. The ratio of the number of BS - relay/UE slots to the Relay-UE slot(s) can be specified. The scheduling priority is for BS to relay relaying transmission to achieve a full relay to UE transmission slot. In single path dynamic routing, only those UEs which have been determined to have the BS-UE link better than the Relay-UE link are scheduled in the BS-UE scheduling interval. In dual path dynamic routing, all UEs are considered for scheduling in the BS-UE scheduling interval, such that the UE with multiple favorable paths to it may receive in both scheduling intervals.

Simulation Results

The reference results for performance comparison are the following:

- performance of the system with no relays
- performance of the system with relays, including dynamic resource partitioning, dynamic routing, and active set which includes relays.

The key reason for introduction of relaying is to improve the data rate distribution to the users. This benefit is shown in the average web page download data rate results for non full queue web browsing in Figure 4.

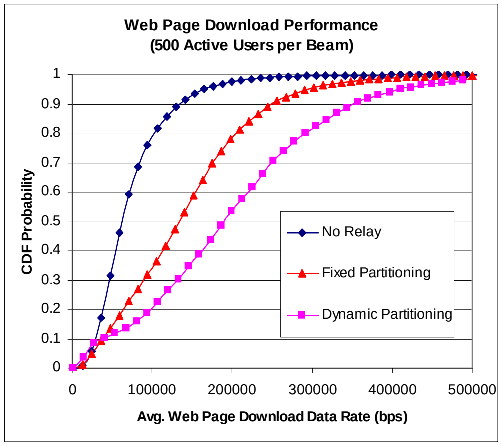


Figure 4 Web page download performance with dynamic resource management

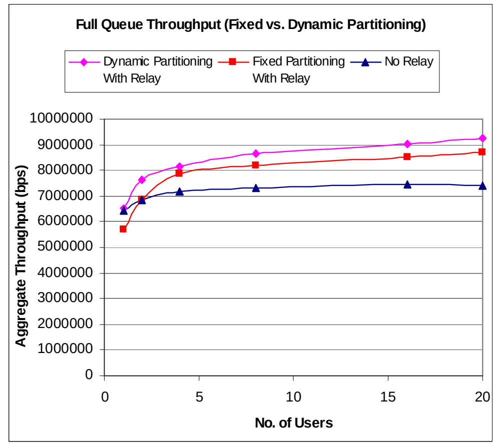


Figure 5 Aggregate throughput of system with dynamic resource management

The gain in the *average* user perceived download performance is more than 150% with respect to the no relay case. Figure 4 also shows the performance with fixed resource partitioning, where the gain in average web page download data rate decreases to ~90%. The benefit of dynamic resource partitioning is significant.

When the base to relay link is the limiting link for relay data transmission to a specific UE, dynamic resource partitioning can increase the effective data rate (up to two times in this setup) to that UE as it is possible to dedicate the full 2/3 resource for base to relay transmission.

Figure 5 shows the throughput performance of the proposed system. As seen in Figure 5, even after dedicating a fraction of the resources for transmission on the base to relay link, there is no loss in full queue capacity when there are sufficient number of simultaneous users. The corresponding capacity gain is approximately 21%.

Figure 5 also shows the performance of fixed resource partitioning, with a small decrease (6%) in full queue throughput over dynamic resource partitioning. Additionally, it was observed that some resource wastage occurs in the relay to UE slot when the transmission rate of the BS-relay link is less than that of the relay-UE link. Dynamic resource partitioning improves the resource usage with the unused resource at 0.6% versus fixed resource partitioning where the unused resource is at 8.2%.

Figure 6 and Figure 7 examine the impact of single path dynamic routing, where each UE is assigned to be served by either the direct link from the base or the relay link. The gain for non full queue web browsing in Figure 6 reduces to ~80%. One key reason for the gain reduction is the fact that for UEs which can operate on both direct and relay links, dual path routing allows simultaneous allocation of resources for transmission to the same UE using both links. The full queue throughput gain in Figure 7 reduces from ~21% to ~15% over the case with no relay.

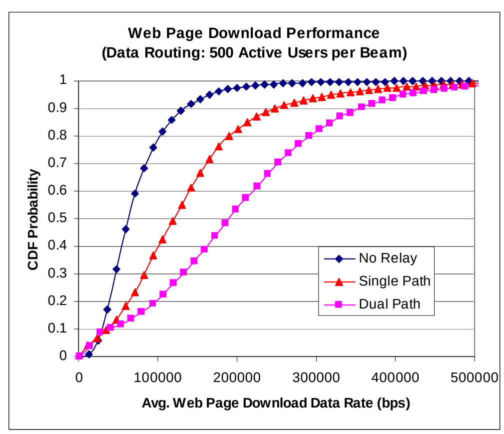


Figure 6 Web page download performance comparing single and dual path dynamic routing

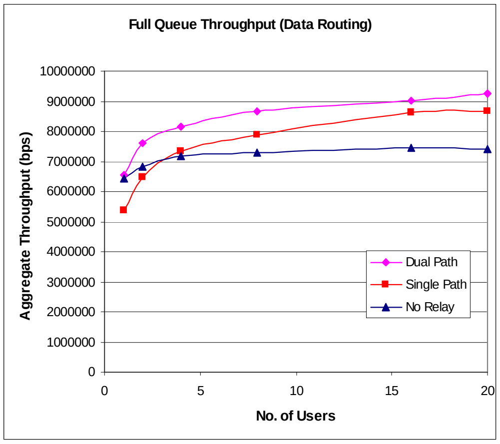


Figure 7 Full queue throughput performance comparing single and dual path dynamic routing

Next, the benefit of including relays in the active set is examined. For this simulation, fixed partitioning with dynamic routing is assumed. From the distribution plots of average C/I at the UE shown in Figure 8, the probability of the C/I exceeding 10dB increased from ~15% (no relay) to ~ 62% (with active set including relays). When the active set selection includes BSs only, the C/I improvement is not as significant in the lower C/I range. When a relay is chosen as the best link by a UE, its parent BS may sometimes not have an adequate link to the UE, thus requiring the relay to support both

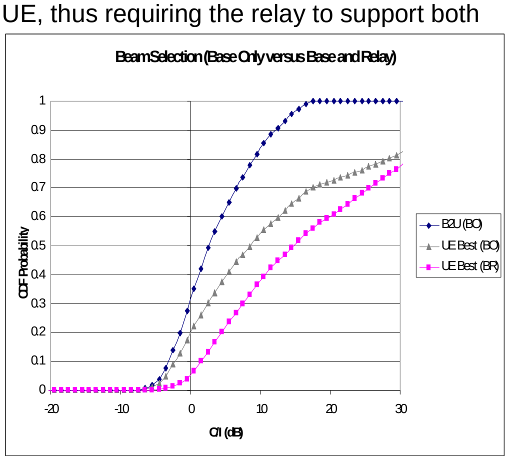


Figure 8 CIR improvement with active set comprising BSs and relays

signalling and data communications. The user data rate CDF for full queue traffic in Figure 9 indicates that almost half of the lower data rate users can be shifted to a

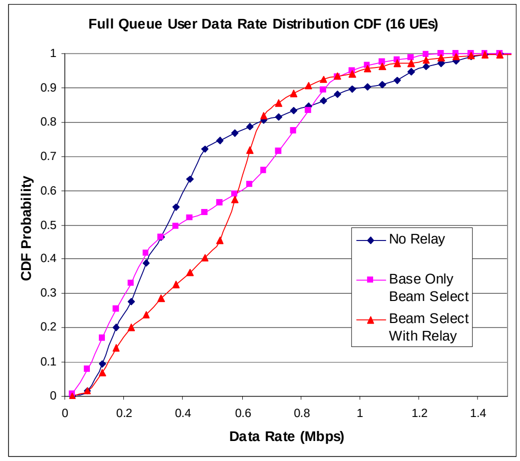


Figure 9 Full queue user data rate distribution for active set with relays

higher data rate.

With the introduction of a second relay into the beam, the location of the relays (staggered along the center path of the beam or distributed to the sides of the beam) impacts the observed throughput as shown in Figure 10. The full queue throughput capacity gain increases further to more than 25% of that without relay. With relays deployed offset from the beam center, the base to relay link now becomes more likely to be the limiting link on the BS transmission to the UE, creating a transmission bottleneck. This reduces the effectiveness of the PF scheduler and produces a lower full queue

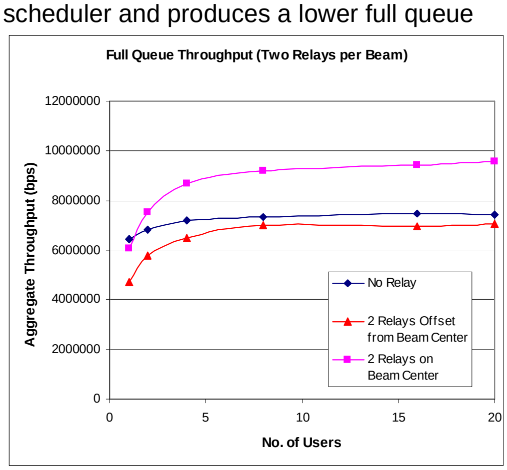


Figure 10 Full queue throughput with two relays per beam

aggregate throughput than the case without relays. This indicates that a good BS-relay link is an important design parameter. There is minimal impact on the average web page download performance with two relays as seen in Figure 11, with some improvement at the low data rate region. The weak BS-relay

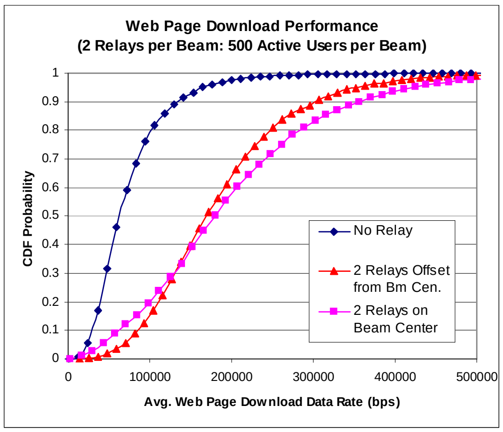


Figure 11 Web page download performance with two relays

link with the offset relays does not significantly affect the coverage for the bursty web browsing traffic, because the dynamic partitioning algorithm prioritizes the BS-relay/UE resource partition to carry the data for the relay-UE transmission of the data burst to the UE.

For the simulation scenario examined in this paper, the majority of the gain from relaying is obtained from a single relay per beam.

Conclusion

Dynamic resource management strategies for cellular infrastructure based multihop relaying have been addressed in this paper. The key strategies of active set management of BS and relays, dynamic resource partitioning and dynamic routing are addressed and their performance assessed via simulation. The performance improvements that the dynamic resource management schemes provide have been shown with end-to-end performance assessment for FTP and web browsing. It is seen that the maximum benefit from relaying is obtained with a single relay per beam.

The focus of this work has been on the downlink to the UE. The applicability to the uplink requires further study. Other important aspects pertinent to relaying include studying the other aspects of dynamic routing (QoS, relay buffer overflow prevention), and mobility related considerations such as the management of layer 2 retransmissions when the errors occur midway through the BS-relay(s)-UE hops, and the flow control necessary for ensuring that the packets are not unnecessarily retransmitted either in the case of dynamic routing or in the case of handoff. Further work is needed to develop and evaluate algorithms to address these issues.

References

[1] N. Esseling, B. H. Walke, R. Pabst, Performance of MAC-Frame-Based Protocols for Mobile Broadband Systems using Layer 2 Relays, 11th Wireless World Research Forum, 10.-11. June 2004, Oslo, Norway.

[2] R. Pabst, *et al*, Relay-based Deployment Concepts for Wireless and Mobile Broadband Radio, IEEE Communications Magazine, Sep 2004.

[3] B..H. Walke, I. Herwono, R. Pabst, Service Architecture for Infrastructure Based Multi-hop Networks Base don SIP, ICCT 2003, Beijing, China

[4] Sreng, V.; Yanikomeroglu, H.; Falconer, D, Coverage Enhancement Through Two-hop Relaying in Cellular Radio Systems, IEEE WCNC2002, vol. 2, 17-21 March 2002, pp. 881 - 885

[5] A. Negi, S. Singh, Improvement in Reliability of Coverage Using 2-hop Relaying in Cellular Networks, Proceedings of 14" IEEE 2003 International Symposium on Personal Indoor and Mobile Radio Communication, pp.2348-2352.

[6] B. Walke and R. Briechle, A Local Cellular Radio Network for Digital Voice and Data Transmission at 60GHz, Proc. Cellular & Mobile Comms. INTL, pp. 215-225, London, 11/1985, available from www.comnets.rwth-aachen.de/~publications

[7] N. Esseling, H.S. Vandra and B. Walke, A Forwarding Concept for HiperLAN/2, Proc. of European Wireless 2000, pp. 13-18, 09/2000, available from www.comnets.rwth-aachen.de/~publications

[8] B. Walke, N. Esseling, J. Habetha, A. Hettich, A. Kadelka, S. Mangold, J. Peetz and U. Vornefeld, IP over Wireless Mobile ATM - Guaranteed Wireless QoS by HiperLAN/2, Proc. of the IEEE, Vol.89, pp. 21-40, 01/2001

[9] M.S. Smith, M. Newton, J. E. Dalley, Multiple Beam Antenna, US Pat. No. 6480524, Nov.12, 2002

[10] Jalali A, Padovani R, Pankaj, R. Data Throughput of CDMA-HDR a High Efficiency-High Data Rate Personal Communication Wireless System, Proc. 51st IEEE Vehicular Technology Conference 2000; 3:1854-1858.

[11] Channel Models for Fixed Wireless Applications, IEEE 802.16.3c-01/29r4, July, 2001

[12] Spatial Channel Model Text Description, SCM-130, April 10, 2003.