Performance Analysis of Access Selection and Transmit Diversity in Multi-Access Networks

Fredrik Berggren
Huawei Technologies
Kista, Sweden
fredrik.b@huawei.com

Remco Litjens
TNO ICT
Delft, The Netherlands
remco.litjens@tno.nl

ABSTRACT

Motivated by the 'beyond 3G' vision of radio access network integration and coordinated radio resource management, a purely analytical performance assessment is presented for a single access point integrating multiple radio accesses. Principal focus is placed on the evaluation of multi-user diversity, multi-access diversity and trunking gains. Scenarios with persistent and non-persistent data flows are investigated, concentrating on throughput and transfer time performance, respectively. A number of numerical experiments are included in order to quantify the relative contribution of the distinguished aspects to the performance gain. These experiments indicate that the exploitation of multi-user diversity with a channel-aware access selection scheme attains the most significant gains, while also the trunking gain that is due to an above-proportional performance enhancement when aggregating system-specific capacities, is noted to be significant. The assignment of multiple accesses to a given flow is demonstrated to have limited potential.

Categories and Subject Descriptors

C.2 [**Computer communication networks**]: Network operations.; G.3 [**Probability and statistics**]: Stochastic processes.

General Terms

Algorithms, performance.

Keywords

Multi-access networks, radio access selection, transmit diversity, scheduling, proportional fair, multi-user diversity, stochastic analysis, performance evaluation, trunking gains.

1. INTRODUCTION

Radio access networks in personal communication systems today generally either operate a single radio interface or in-tegrate multiple radio interfaces at flow (or call) level, e.g. 3GPP standardised inter-system handovers between 2G/3G networks. The large number of RADIO ACCESSES (RAs) that are currently available or in development, could in principle be closely integrated to provide *multi-radio access*, not only enabling flow level cooperation but potentially also allowing access coordination at packet or even frame level. Such networking solutions, integrating multiple, possibly heterogeneous RAs, i.e. operating distinct RADIO ACCESS TECHNOLOGIES, have recently started to draw interest. For example, the AMBIENT NETWORKS project [2, 19, 20] investigates solutions to combine and compose networks, potentially even associated with different administrative domains, to enhance capacity, resource efficiency and service quality. Therefore, RADIO ACCESS selection becomes an important radio resource management mechanism, that should be able to coordinate multiple RAs, possibly with different capacities and coverage areas. Radio access selection for certain specific RAs (e.g. WCDMA/GSM) capturing e.g. trunking gains, can be found in [11, 23]. In this work we will address the RA selection problem from a more general perspective and highlight the diversity effects.

With regard to RA selection and scheduling in an advanced multi-radio access network, two distinct forms of diversity are envisaged. On the one hand, *multi-user diversity* stems from the fact that a channel-aware scheduler time-multiplexes multiple data flows over a single shared transport channel, based on the flows' uncorrelated fast fading processes. By scheduling at each time instant the data flow with the most favorable current (relative) channel conditions, the expected throughput is enhanced. Schedulers exploiting multi-user diversity have been developed previously for e.g. HSDPA or 1xEV-DO systems (see e.g. [1, 4, 6, 14] and references therein). The availability of several RAs, however, extends the scheduling problem with yet another dimension. The proposed opportunistic scheduling schemes take advantage of the varying radio channel and the elasticity of the data flow. A common fairness criterion is to grant flows access to the channel an equal amount of time. Both the ROUND ROBIN (RR) and the PROPORTIONAL FAIR (PF) schedulers are designed to achieve such fairness. The principle distinction between the RR and PF schedulers is that the channel-oblivious RR scheduler cyclically grants channel access to the different flows, regardless of the flows' actual channel conditions, while the channel-aware PF scheduler carefully selects when to serve which flow, while still establishing fairness, albeit possibly on a slightly larger time scale.

On the other hand, *multi-radio transmit diversity* (MRTD; or multi-access diversity) stems from the fact that a channel-aware RA selection scheme can choose between multiple RAs when serving a given data flow, based on the flows' uncorrelated fast fading processes on the different RAs. By assigning at each time instant the RA with the most favorable conditions, the expected throughput is enhanced. Two distinct modes of MRTD can be identified. The case where the RA selection scheme is able to assign multiple RAs to any given flow simultaneously is denoted *parallel* MRTD. Parallel MRTD can be exploited to enhance the robustness of the transmissions, by transmitting identical information via multiple RAs, or to increase throughputs, by transmitting different information via multiple RAs. In this paper, parallel MRTD is applied to enhance throughputs. The alternative case, where a flow may only be assigned a single RA at a time is denoted *switched* MRTD. In both cases, the R&D ambition is to enable RA selection at the frame level in order to exploit the SIR fluctuations due to multipath fading.

Aside from the identified diversity gains, the integration of multiple RAs yields a *trunking gain*. The trunking gain stems from the generic observation that an L times increase in the capacity, yields an above-L times increase in the supportable traffic load, given some performance target, e.g. a maximum average transfer time or blocking probability. The significance of the trunking gain, which is due to flow level multiplexing, can be readily observed from traditional queueing models for single RA systems, e.g. the Erlang loss model for circuit-switched systems [10] or the processor sharing model for packet-switched systems [17, 18].

In this paper we devise integrated scheduling and RA selection schemes and present a purely analytical assessment of the performance of a multi-radio access system deploying these access selection schemes. A *persistent* flow analysis is presented to capture the channel-oblivious or channel-aware access selection operations at a small time scale, while a *non-persistent* flow analysis captures the flow level dynamics that is due to the initiation and completion of finite flows. As it is in principle non-trivial which source of performance gain contributes the most to the overall system performance, the obtained closed-form expressions are used in a numerical section to demonstrate the magnitudes of the multi-user diversity, multi-access diversity and trunking gains. As our focus is on the assessment of the relative *performance potential* of the different aspects of access selection, intrinsic implementational complexities, e.g. signalling overhead, synchronisation issues, feedback delay, ..., are ignored, in order to allow analytical tractability.

The remainder of this paper is organised as follows. Section 2 defines the assumed system, propagation and traffic model, and introduces the associated notation. Some remarks are made regarding the impact of some key approximations. In Section 3 distinct integrated scheduling/RA selection schemes are proposed and a throughput analysis is presented for the case of a given number of persistent flows. Section 4 subsequently presents a transfer time analysis for the case of non-persistent flows, where finite flows originate at random time instances, depart upon completion of the transfer and thus establish a dynamic process where the number of flows simultaneously in progress varies over time. The presented flow level analysis utilises the results derived in the previous section. Section 5 presents and discusses the results from a set of numerical experiments,

including a numerical validation of some key modelling assumptions. Section 6 ends this paper with some concluding remarks.

2. MODEL

We consider a single access point which integrates a total of L Radio Accesses (see Figure 1). These RAs may e.g. consist of different channels, Radio Access networks of distinct operators or Radio Access Technologies. In order to provide general insights, the considered Radio Accesses are of a generic nature, characterised by a shared transport channel whose capacity is given by the well-known Shannon-Hartley formula [21]. Denote with $W^{(l)}$ the bandwidth available to RA l, $l = 1, \cdots, L$, expressed in MHz.
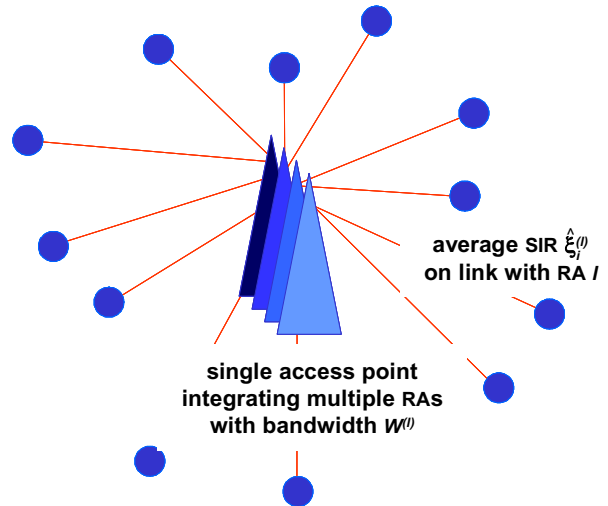

Figure 1: The considered system model comprises multiple radio accesses in a single access point, serving multiple flows.

Flows are assumed to originate in the coverage area of the access point according to a Poisson process with rate λ (in flows/s). A flow corresponds to the transfer of a file with a generally distributed size with mean A (in Mbits/flow). Once a flow has transferred all associated bits, it departs from the system. The offered traffic load ρ is defined as λA (in Mbits/s). An admission control scheme limits the total number of admissible flows to $N_{\max}$.

For a given flow i, the radio link quality of RA l at time frame t is denoted $\xi_i^{(l)}(t)$. In our approach $\xi_i^{(l)}(t)$ describes the Signal-to-Interference Ratio (SIR), although other quality measures, such as channel quality indicators or error rates could in principle also be used. The stochastic process $\left(\xi_i^{(l)}(t), t \geq 0\right)$ describing the evolution of flow i's SIR on RA l, is assumed stationary and ergodic, while the different SIR processes are assumed independent among both RAs and flows. The SIR variations are specified by an average SIR, denoted $\hat{\xi}_i^{(l)}$, and a small-scale Rayleigh fading process, so that the random variable $\xi_i^{(l)}$, expressing the SIR of flow i on RA l at an arbitrary time instant, is exponentially distributed with mean $\hat{\xi}_i^{(l)}$. Or, equivalently, $\xi_i^{(l)}/\hat{\xi}_i^{(l)} \sim Exp(1)$. For each channel state, a link adaptation scheme is assumed

to provide a suitable transmission rate, viz.

$$r_{\text{SH}}^{(l)}(\xi^{(l)}) \equiv W^{(l)} \log_2(1 + \xi^{(l)}),$$

(in Mbits/s), applying the well-known Shannon-Hartley channel capacity formula [21]. It is assumed that channel estimation and link adaptation are error-free and instantaneous.

An integrated packet scheduling and RA selection scheme operates at the access point to assign the available RAs to the present flows. Different alternatives are considered, characterised by either ROUND ROBIN or PROPORTIONAL FAIR access, and allowing either switched or parallel MRTD. A key common feature of all schemes is the inherent fairness, in the sense that each flow obtains equal long-term access to the resources, regardless of its average SIR. Furthermore, under any scheme, a given RA may be assigned to a single flow only at any given time, while in case of parallel MRTD, a given flow may utilise several RAs simultaneously.

A note on the key modelling assumptions

In line with our focus on the analytical assessment of the performance *potential* of different access selection schemes in multi-access networks, a number of approximations are made. For instance, in practical implementations the gains of channel-aware scheduling come at the expense of a required feedback channel (in the reverse direction) and thus an additional load that is imposed on the system. It was shown in [12] that a 'thin' feedback channel may suffice to achieve the principal performance gains. Furthermore, the inherent delay in the feedback loop may influence the accuracy of the channel quality estimation and thus affect the adequacy of both rate adaptation and channel-aware access selection. We will address this issue in the validation experiments of Section 5. Other integration complexities, e.g. the real-time processing of RA- and UE-specific link quality measurements in a multi-radio environment place requirements on the processing capacity of the access points, but not necessarily influence the radio network performance.

3. THROUGHPUT ANALYSIS FOR PERSISTENT FLOWS

In this section, different fair access selection schemes are defined and the throughput performance is analysed. This analysis is carried out for a fixed number N of persistent flows contending for the resources associated with the multiple RAs. In Section 4 the derived analytical expressions are used in an analysis with flow level dynamics. Numerical experiments are presented in Section 5 to illustrate the performance impact of the different schemes. The organisation of this section is according to the distinct access selection schemes considered, i.e. ROUND ROBIN versus PROPORTIONAL FAIR access selection.

3.1 Round robin radio access selection

The purpose of the access selection is to assign flows to RAs. This problem could consider different objective functions, e.g. to maximise throughput or to establish some form of fairness. Our focus is on establishing the same long-term access to any RA among flows, which can be regarded as a form of resource fairness. One access selection scheme that achieves such fairness is the channel-oblivious ROUND ROBIN scheme. Ignore for a moment the potential restrictions imposed by (switched or parallel) MRTD and consider RA l in

isolation. Define the auxiliary function

$$
\begin{aligned}
H(k, l, \hat{\xi}^{(l)}) &\equiv \int_0^\infty r_{\text{SH}}^{(l)}(\xi) \frac{1}{\hat{\xi}^{(l)}} e^{-\frac{\xi}{\hat{\xi}^{(l)}}k} \, d\xi \\
&= W^{(l)} \frac{e^{k/\hat{\xi}^{(l)}}}{k \ln 2} \Gamma\left(\frac{k}{\hat{\xi}^{(l)}}\right),
\end{aligned}
\tag{1}
$$

where $\Gamma(x) = \int_x^\infty t^{-1} e^{-t} \, dt$ is an incomplete Gamma function, applying substitution and partial integration techniques to obtain the final expression. For $k = 1$, $H(1, l, \hat{\xi}^{(l)})$ is readily seen to express the long-term average throughput experienced by an isolated flow on RA l with average SIR $\hat{\xi}$, which is obtained by conditioning on the exponentially distributed instantaneous SIR[1]. Consider now a scenario with N persistent flows. Denote with the random variable i_l^* the flow that is assigned RA l in an arbitrary time frame, so that flow i is selected with probability $\Pr\{i_l^* = i\} = 1/N$. The long-term average throughput experienced by flow i on RA l is then given by

$$R_i^{(l)} = \frac{1}{N} H(1, l, \hat{\xi}_i^{(l)}). \tag{2}$$

Consider now the scenario with multiple RAs and *switched* MRTD. Suppose that the number of flows exceeds the number of RAs, i.e. $L \leq N$, and we assign the flows to the RAs sequentially, starting with, say, RA $l = 1$. If flow i is assigned to RA 1 in the considered arbitrary time frame, i.e. if $i_1^* = i$, flow i is not allowed to take part in the allocation of RA 2. More generally, the probability that RA l is assigned to flow i is equal to the probability that flow i has not already been endowed with any of the previously assigned RAs times the probability that flow i is assigned RA l among the remaining unendowed flows, i.e.

$$
\Pr\{i_l^* = i\} =
\begin{cases}
\frac{1}{N}, & l = 1, \\
\prod_{l'=2}^{l} \left(1 - \frac{1}{N+2-l'}\right) \frac{1}{N+1-l} = \frac{1}{N}, & l > 1,
\end{cases}
$$

$l = 1, \cdots, L$. Hence, as is intuitively clear, a flow is indeed assigned with equal probability to any RA. The expected throughput experienced by flow i can then be expressed as

$$R_i = \frac{1}{N} \sum_{l=1}^{L} H(1, l, \hat{\xi}_i^{(l)}). \tag{3}$$

If the same access selection scheme were adopted for the alternate case of $L > N$, only the first N RAs would be assigned. Thus, the access probability to these RAs would be $1/N$, but the remaining $L - N$ RAs are not utilised and have access probability zero. To avoid this, we may alternatively perform the access selection by assigning RAs to flows sequentially, i.e. flow 1 selects arbitrarily from L RAs, flow 2 selects from $L - 1$ RAs, and so on. Then, it is easy to see that

$$\Pr\{i_l^* = i\} = \frac{1}{L},$$

$l = 1, \cdots, L$. The expected throughput of flow i can then

---

[1] Although for general k, $H(k, l, \hat{\xi}^{(l)})$ may not allow such an easy intuitive interpretation, it is nonetheless a convenient auxiliary function that enables simplified notation throughout the paper.

be expressed as

$$R_i = \frac{1}{L} \sum_{l=1}^{L} H(1, l, \hat{\xi}_i^{(l)}). \qquad (4)$$

Combining the obtained expression for both cases, we get

$$R_i = \frac{1}{\max\{N, L\}} \sum_{l=1}^{L} H(1, l, \hat{\xi}_i^{(l)}), \qquad (5)$$

for any $N, L \geq 1$.

Consider for a moment the case where $\hat{\xi}_i^{(l)} = \hat{\xi}$, for all $i = 1, \cdots, N$, $l = 1, \cdots, L$. Expression (5) then simplifies to $R_i = (L/\max\{N, L\}) H(1, l, \hat{\xi})$. Hence for small N (compared to L), the expected throughput does not increase with L due to the single RA restriction inherent to switched MRTD, while for large N the expected throughput is proportional in L. Above-proportional throughput gains from an increase in the number of RAs can be achieved by a channel-aware access selection scheme, as will be shown in Section 3.2.

In a scenario with *parallel* MRTD, flows can be assigned multiple RAs simultaneously and hence access selection can be done independently per RA. As is intuitively readily understood, the possibility to assign multiple RAs to a given flow potentially raises the flow's throughput. Note that the probability that RA l is assigned to flow i is now trivially equal to $\Pr\{i_l^* = i\} = 1/N$, for all $l = 1, \cdots, L$. The probability that a given flow is assigned k RAs is given by the binomial probability

$$\Pr\{a\text{ flow gets }k\text{ RAs}\} = \binom{L}{k}\left(\frac{1}{N}\right)^k \left(\frac{N-1}{N}\right)^{L-k}.$$

The expected throughput of flow i can then be derived as

$$
\begin{aligned}
R_i &= \sum_{l=1}^{L} \left(\frac{1}{N}\right)^1 \left(\frac{N-1}{N}\right)^{L-1} H(1, l, \hat{\xi}_i^{(l)}) + \\
&\quad \sum_{l \neq k} \left(\frac{1}{N}\right)^2 \left(\frac{N-1}{N}\right)^{L-2} \left(\begin{array}{c} H(1, l, \hat{\xi}_i^{(l)}) \\ + H(1, k, \hat{\xi}_i^{(k)}) \end{array}\right) + \cdots \\
&= \sum_{l=1}^{L} \sum_{k=1}^{L} \frac{k}{L} \binom{L}{k}\left(\frac{1}{N}\right)^k \left(\frac{N-1}{N}\right)^{L-k} H(1, l, \hat{\xi}_i^{(l)}) \\
&= \frac{1}{N} \sum_{l=1}^{L} H(1, l, \hat{\xi}_i^{(l)}), \qquad (6)
\end{aligned}
$$

which is also immediately clear from summing expression (2) over $l = 1, \cdots, L$, as it is implied by the application of parallel MRTD that all RAs are utilised. For $L \leq N$, expression (6) is the same as expression (3), and hence the flow (or aggregate) throughput is not changed by applying parallel rather than switched MRTD if the number of flows is relatively large. Alternatively, for $L > N$, comparing expressions (6) and (4), note that the summation index is the same but the throughput terms are weighted with N and L, respectively. Hence, if the number of flows is relatively small, parallel MRTD provides better throughput than switched MRTD, which is intuitively clear since in the case of switched MRTD some RAs remain unassigned.

3.2 Proportional fair radio access selection

A channel-aware access selection scheme exploits the channel variations to increase the throughput. Since flows may

have different mean radio link qualities $\hat{\xi}_i^{(l)}$, an adequate access selection scheme avoids starving flows with relatively poor quality. A Proportional Fair (PF) access selection scheme that manages this operates as given by

$$i_l^* = \arg\max_{1 \leq i \leq N} \frac{\xi_i^{(l)}}{\hat{\xi}_i^{(l)}}.$$

For the PF scheme, the probability that RA l is assigned to flow i, given instantaneously experienced SIR $\xi_i^{(l)}$, is equal to

$$
\begin{aligned}
\Pr\left\{i_l^* = i \,\Big|\, \xi_i^{(l)}\right\} &= \Pr\left\{\xi_i^{(l)} \geq \frac{\hat{\xi}_i^{(l)}}{\hat{\xi}_j^{(l)}} \xi_j^{(l)}, \forall j \neq i\right\} \\
&= \left(1 - e^{-\frac{\xi_i^{(l)}}{\hat{\xi}_i^{(l)}}}\right)^{N-1}, \qquad (7)
\end{aligned}
$$

which is noted to be increasing in $\xi_i^{(l)}$. Deconditioning expression (7) with respect to the experienced SIR $\xi_i^{(l)}$, it is readily derived that

$$
\begin{aligned}
\Pr\{i_l^* = i\} &= \int_0^\infty \left(1 - e^{-\frac{\xi_i^{(l)}}{\hat{\xi}_i^{(l)}}}\right)^{N-1} \frac{1}{\hat{\xi}_i^{(l)}} e^{-\frac{\xi_i^{(l)}}{\hat{\xi}_i^{(l)}}} d\xi_i^{(l)} \\
&= \left[\frac{1}{N}\left(1 - e^{-\frac{\xi_i^{(l)}}{\hat{\xi}_i^{(l)}}}\right)^N\right]_0^\infty = \frac{1}{N},
\end{aligned}
$$

i.e. the PF access selection scheme indeed establishes fair access to the resources, regardless of a flow's average SIR, which is also immediately clear from the fact that access selection is based on the *ratio* $\xi_i^{(l)}/\hat{\xi}_i^{(l)}$, which is probabilistically identical for all flows. Since both the RR and PF schemes establish this fairness property, a throughput comparison can be made between the RR and the PF access selection schemes. Ignore for a moment the potential restrictions imposed by MRTD and consider RA l in isolation. Under the PF scheme, the long-term average throughput of flow i on RA l is given by

$$
\begin{aligned}
R_i^{(l)} &= \int_0^\infty r_{\text{SH}}^{(l)}(\xi) \left(1 - e^{-\frac{\xi}{\hat{\xi}_i^{(l)}}}\right)^{N-1} \frac{1}{\hat{\xi}_i^{(l)}} e^{-\frac{\xi}{\hat{\xi}_i^{(l)}}} d\xi \\
&= \sum_{k=0}^{N-1} \binom{N-1}{k}(-1)^k \int_0^\infty r_{\text{SH}}^{(l)}(\xi) \frac{1}{\hat{\xi}_i^{(l)}} e^{-\frac{\xi(k+1)}{\hat{\xi}_i^{(l)}}} d\xi \\
&= \frac{1}{N} \sum_{k=1}^{N} k \binom{N}{k}(-1)^{k+1} H(k, l, \hat{\xi}_i^{(l)}), \qquad (8)
\end{aligned}
$$

applying Newton's binomium, with $H(\cdot)$ again as defined in expression (1). It can be verified that expression (8) is larger than expression (2), basically since $\Pr\left\{i_l^* = i \,\Big|\, \xi_i^{(l)}\right\}$ is increasing in $\xi_i^{(l)}$ while overall access is fair, i.e. a flow is more likely to be granted its fair share of access time at instances where it experiences a more favourable radio link quality and thus a higher potential bit rate. This establishes the multi-user diversity gain that is achieved by the PF scheme.

Consider now the scenario with multiple RAs, PF access selection and *switched* MRTD. Suppose that the number of

flows exceeds the number of RAs, i.e. $L \leq N$, and the access selection scheme sequentially considers the available RAs and assigns it to the flow with the best relative channel quality, starting with, say, RA $l = 1$. Since there are ever fewer flows contending for the RAs as more and more RAs have been processed, the potential multi-user diversity gain will decrease for each subsequent RA assignment. It is stressed, however, that the resulting performance cannot become worse than under RR access selection. A flow may contend for RA l if it has not been assigned any RA with index lower than l. Hence, when selecting a flow for RA l, the probability that flow i has not yet been assigned to any of the previously processed RAs $1, \cdots, l-1$, and is thus still available for RA assignment is given by

$$\Pr \{ \text{flow } i \text{ available} | \, l - 1 \text{ RAs assigned} \}$$

$$= \begin{cases} 1, & l = 1, \\ \prod\limits_{l'=2}^{l} \left(1 - \frac{1}{N+2-l'} \right), & l > 1. \end{cases} = \frac{N+1-l}{N},$$

for $l \geq 1$. Performing the access selection starting from the first RA, for the l-th RA we have

$$
\begin{aligned}
R_i^{(l)} & = \Pr \{ \text{flow } i \text{ available} | \, l - 1 \text{ RAs assigned} \} \times \\
& \int_0^\infty r_{\mathrm{SH}}^{(l)}(\xi) \left(1 - e^{-\frac{\xi}{\hat{\xi}_i^{(l)}}} \right)^{N-l} \frac{e^{-\frac{\xi}{\hat{\xi}_i^{(l)}}}}{\hat{\xi}_i^{(l)}} \, d\xi \\
& = \frac{N+1-l}{N} \times \\
& \int_0^\infty r_{\mathrm{SH}}^{(l)}(\xi) \left\{ \sum_{k=0}^{N-l} \binom{N-l}{k} \left(-e^{-\frac{\xi}{\hat{\xi}_i^{(l)}}} \right)^k \right\} \frac{e^{-\frac{\xi}{\hat{\xi}_i^{(l)}}}}{\hat{\xi}_i^{(l)}} \, d\xi \\
& = \frac{N+1-l}{N} \times \\
& \sum_{k=1}^{N-l+1} \binom{N-l}{k-1} (-1)^{k-1} \int_0^\infty r_{\mathrm{SH}}^{(l)}(\xi) \frac{e^{-\frac{\xi}{\hat{\xi}_i^{(l)}}k}}{\hat{\xi}_i^{(l)}} \, d\xi \\
& = \sum_{k=1}^{N+1-l} \frac{k}{N} \binom{N+1-l}{k} (-1)^{k+1} H(k, l, \hat{\xi}_i^{(l)}),
\end{aligned}
$$

applying Newton's binomium, and hence

$$R_i = \sum_{l=1}^{L} \sum_{k=1}^{N+1-l} \frac{k}{N} \binom{N+1-l}{k} (-1)^{k+1} H(k, l, \hat{\xi}_i^{(l)}). \quad (9)$$

It is important to realise that the order in which the different RAs are addressed will affect the overall performance. The applied *heuristic* is to address the RAs in the order dictated by the associated indices. An alternative implementation could be to try each possible order, register the overall performance according to e.g. the aggregate value $\sum_{l=1}^{N} \left(\xi_{i_l^*}^{(l)} / \hat{\xi}_{i_l^*}^{(l)} \right)$ and effectuate the found 'best order'. It is noted, however, that not only is such a scheme analytically hard to evaluate, an implementation in a live network is also computationally very intensive. Particularly for large L, the required computation time is likely to be too large to be carried out in real-time.

Once again, if we would adopt the same access selection scheme for the alternate case of $L > N$, only the first N

RAs would be assigned and a lower degree of multi-RA diversity gain would be achieved. If the access selection is instead carried out by addressing the different flows in some order to allocate the available RAs according to the PF principle, a larger set of samples per access selection step will enhance the multi-access diversity gain. Following an equivalent analysis as that presented above, we find

$$R_i = \sum_{l=1}^{L} \sum_{k=1}^{L+1-i} \frac{k}{L} \binom{L+1-i}{k} (-1)^{k+1} H(k, l, \hat{\xi}_i^{(l)}). \quad (10)$$

It is readily derived that the probability that RA l is assigned to flow i is given by $\Pr \{ i_l^* = i \} = 1/L$, for all $l = 1, \cdots, L$. Comparing the alternative schemes applied for the cases $L \leq N$ and $L > N$, it is noted that in the former case, the throughput is higher, the sooner the channel is scheduled, while in the latter case the throughput is higher, the sooner the flow is scheduled. To let all flows experience the same degree of multi-RA diversity, the order in which they are scheduled can be varied. If we alternate the flow ordering for (10) in a random (or RR) manner, the expected flow throughput becomes

$$\hat{R}_i = \frac{1}{N} \sum_{j=1}^{N} \sum_{l=1}^{L} \sum_{k=1}^{L+1-j} \frac{k}{L} \binom{L+1-j}{k} (-1)^{k+1} H(k, l, \hat{\xi}_i^{(l)}). \quad (11)$$

Observe that if all flows have the same average SIR $\hat{\xi}$, the long-term average throughput given in expression (11) is equal for all flows.

In the scenario with *parallel* MRTD, the assignment of flows to RAs is done independently per RA and flows can be assigned multiple RAs simultaneously. Since there is no restriction on the number of RAs assigned to any given flow, the flow throughput is readily given by

$$R_i = \sum_{l=1}^{L} R_i^{(l)} = \sum_{l=1}^{L} \sum_{k=1}^{N} \frac{k}{N} \binom{N}{k} (-1)^{k+1} H(k, l, \hat{\xi}_i^{(l)}), \quad (12)$$

using $R_i^{(l)}$ as given in expression (8). It can be noted also here that $\Pr \{ i_l^* = i \} = 1/N$ for all $l = 1, \cdots, L$.

4. TRANSFER TIME ANALYSIS FOR NON-PERSISTENT FLOWS

In the previous section, analytical expressions have been derived for the long-term average flow throughput as a function of the number of available RAs (L), the number of flows in the system (N) and the flow/RA-specific average SIRs ($\hat{\xi}_i^{(l)}$). Applying a separation of time scales technique, we will use these throughput expressions in a flow level analysis, where *finite* (or non-persistent) flows originate, undergo transfer and subsequently depart from the system.

4.1 Round robin radio access selection

Before presenting the flow level analysis for the RR RA selection, recall from expressions (3) and (4) that the flow-specific throughput of flow i under RR RA selection and with *switched* MRTD can be written in the form

$$R_i \left(N, \hat{\xi}_i \right) = \frac{1}{\max \{ N, L \}} \sum_{l=1}^{L} H(1, l, \hat{\xi}_i^{(l)}) \equiv \psi_L (N) \, \varphi_L \left(\hat{\xi}_i \right), \quad (13)$$

where $\hat{\xi}_i \equiv \left(\hat{\xi}_i^{(l)}, l = 1, \cdots, L \right)$ denotes the average SIR vector associated with flow i, function $\psi_L(N)$ is noted to depend on (a priori given L and) N only, and function $\varphi_L\left(\hat{\xi}_i\right)$ is noted to depend on (a priori given L and) $\hat{\xi}_i$ only. Note that $\varphi_L\left(\hat{\xi}_i\right)$ is the long-term time-average throughput of flow i when continuously utilising all L RAs. Function $\psi_L(N)$ implicitly captures both the fair sharing factor $1/N$ and the multi-RA diversity gain $N\psi_L(N) = \mathbf{1}_{\{L \leq N\}} + (N/L)\,\mathbf{1}_{\{L > N\}}$, with $\mathbf{1}_{\{\cdot\}}$ the indicator function, which is equal to 1 (no gain) if $L \leq N$ and less than 1 (loss) if $L > N$, as observed above.

For the case with *parallel* MRTD, we have from expression (6) that

$$R_i\left(N, \hat{\xi}_i\right) = \frac{1}{N}\sum_{l=1}^{L} H(1, l, \hat{\xi}_i^{(l)}) \equiv \vartheta_L(N)\,\varphi_L\left(\hat{\xi}_i\right),$$

where function $\vartheta_L(N)$ is noted to depend on (L and) N only. Observe now that $\vartheta_L(N)$ is precisely equal to the fair sharing factor, expressing an absence of multi-RA diversity gain. In other words, the gain from having multiple RAs merely stems from summing capacity, *not* from intelligent access selections.

As for both MRTD cases, the effects of the SIR averages $\hat{\xi}_i$ and of the number of flows N on the aggregate throughput can be split as indicated, along with the fair resource sharing that is inherent to the RR scheme, we can model the flow level system behaviour as a Generalised Multi-Class Processor Sharing (GMCPS) model with state-dependent aggregate service rates [5, 6]. Note that while L and $\hat{\xi}$ are a priori fixed, N is now dynamically varying over time due to the flows' arrival and departure process. The classification of flows is done by discretisation of the continuous multi-dimensional space of $\hat{\xi}_i$ into K distinct flow classes, characterised by average SIR vector $\hat{\xi}_k$, $k = 1, \cdots, K$. Class-k flows are thus characterised by average SIR vector $\hat{\xi}_k$ and are assumed to originate according to a Poisson process with arrival rate λ_k. Denote with random variable N_k the number of class-k flows present in equilibrium, $k = 1, \cdots, K$.

The GMCPS model belongs to the class of product-form 'networks' and is analytically tractable. In particular, the joint distribution of the number N_k of flows of class k in the system, $k = 1, \cdots, K$, is given by [8, 16]

$$\Pr\{N_1 = n_1, \cdots, N_K = n_K\} = G\frac{n!}{\phi(n)}\prod_{k=1}^{K}\frac{\hat{\rho}_k^{n_k}}{n_k!},$$

$0 \leq n \equiv \sum_{k=1}^{K} n_k \leq N_{\max}$. In this expression, $\hat{\rho}_k \equiv \lambda_k A / \varphi_L\left(\hat{\xi}_k\right)$ denotes the normalised traffic load of class k, $k = 1, \cdots, K$, the function $\phi(n)$ captures the multi-RA selection gain:

$$\phi(n) \equiv \begin{cases} \displaystyle\prod_{n'=1}^{n} n'\psi_L(n'), & \text{in case of switched MRTD}, \\ \displaystyle\prod_{n'=1}^{n} n'\vartheta_L(n') = 1, & \text{in case of parallel MRTD}, \end{cases}$$

with $\phi(0) \equiv 1$ by convention, and G denotes the appropriate normalisation constant, i.e.

$$G \equiv \left(\sum_{n=0}^{N_{\max}} \frac{\hat{\rho}^n}{\phi(n)}\right)^{-1}$$

with $\hat{\rho} \equiv \sum_{k=1}^{K} \hat{\rho}_k$ the aggregate normalised traffic load.

Applying Little's formula (e.g. [22]) the expected transfer time $\mathbf{T}_k$ of a flow of class k can be derived as

$$\begin{aligned} \mathbf{T}_k &= \frac{1}{\lambda_k \Pr\{N < N_{\max}\}}\mathbf{N}_k \\ &= \frac{1}{\lambda_k \Pr\{N < N_{\max}\}}\frac{\hat{\rho}_k}{\hat{\rho}}G\sum_{n=0}^{N_{\max}}\frac{n\hat{\rho}^n}{\phi(n)}, \end{aligned}$$

where $N \equiv \sum_{k=1}^{K} N_k$ and $\mathbf{N}_k$ denotes the expected number of class-k flows present. Observe for the specific scenario with a single class, no admission control ($N_{\max} = \infty$) and parallel MRTD, that the expected transfer time is given by $\mathbf{T} = (1/\lambda)\,\hat{\rho}/(1 - \hat{\rho})$, i.e. the perhaps more familiar expression associated with the egalitarian Processor Sharing model (see e.g. [22]).

Some additional interesting results for the applied GMCPS model have been derived (see [8]). In particular, the conditional expected transfer time $\mathbf{T}_k(x)$ of a class k flow of given size $x \geq 0$ can be computed explicitly and grows linearly in x, i.e. $\mathbf{T}_k(x) = (x/A)\,\mathbf{T}_k$, a result which further expresses the fair allocation of capacity to the served flows, regardless of the flow size. An important additional feature of the GMCPS model is that all these performance measures are *insensitive* with respect to the specific form of the flow size distribution, depending on its mean A only.

4.2 Proportional fair radio access selection

As is immediately clear from expressions (9), (11) and (12), the Proportional Fair RA selection schemes with either switched or parallel MRTD, the flow-specific throughput function is a more complicated function of (L,) N and the $\hat{\xi}_i$'s. In particular, it cannot be split into distinct components comprising the effects of N and the $\hat{\xi}_i$'s, respectively.

Moreover, as the PF RA selection scheme does not necessarily establish fair resource sharing on smaller time scales, in contrast to the RR RA selection scheme, the application of a Processor Sharing type queueing model is not trivially sane here. In light of the fairness property derived in the persistent flow analysis, however, indicating that in the long-term each flow is assigned an equal fraction of the resources, we do propose to use the GMCPS model as an approximation to evaluate the flow level behavior under PF scheduling (see also [5, 6]). The validity of this approximation largely depends on the distinctness of the time scales at which PF scheduling achieves its fairness property, i.e. the time scale of multipath fading, and that at which flows originate and depart from the system. In particular, if the fading process is relatively fast, fairness is likely to be achieved well within a flow's lifetime and hence a Processor Sharing discipline may serve as a good approximation for PF scheduling. At the other extreme, if the fading process is relatively slow, the PF scheme behaves more like a Random Order of Service discipline, where upon termination of a flow, the next flow is selected randomly depending on its (virtually constant) actual SIR realisation (see [7, 9] for some equivalence properties between PS and ROS).

In order to apply the GMCPS model for the PF scheme, given the unfortunate form of the throughput functions, one must enforce a separation as in expression (13) that is robust for the most likely range of the number of flows N and the SIR averages. A possible approach (see also e.g. [5, 6]) is to

approximate the Shannon-Hartley formula by

$$r_{\text{SH}}^{(l)}(\xi^{(l)}) \approx \frac{W^{(l)}\xi^{(l)}}{\ln 2},$$

which is noted to be fairly accurate for low $\xi^{(l)}$. Given this approximation, the applied auxiliary function $H(k,l,\hat{\xi}^{(l)})$ reduces to

$$H(k,l,\hat{\xi}^{(l)}) \approx \int\limits_0^\infty \frac{W^{(l)}\xi}{\ln 2} \frac{1}{\hat{\xi}^{(l)}} e^{-\frac{\xi}{\hat{\xi}^{(l)}}k}\, d\xi = \frac{W^{(l)}\hat{\xi}^{(l)}}{k^2 \ln 2},$$

using partial integration, which is readily verified to allow expressions (9), (11) and (12) to be split into disjunct parts that depend only on (a priori given L and) N and $\widehat{\xi}_i$, respectively, as desired.

For our current purposes of demonstrating the multi-user and multi-RA gains in multi-access networks, however, it suffices to limit the investigation to a single flow class, leaving a more extensive consideration of the multiple class case for the near future. The reduced queueing model is the Generalised Processor Sharing (GPS) model, where $R_i\left(N,\hat{\xi}\right)$ may have an arbitrary form. The expression for the equilibrium distribution then reduces to

$$\Pr\{N = n\} = G\frac{\rho^n}{\phi(n)} \quad \text{with} \quad \phi(n) \equiv \prod_{n'=1}^n n' R_i\left(n',\hat{\xi}\right),$$

$0 \leq n \leq N_{\max}$, with $\rho \equiv \lambda A$ as defined previously, G the appropriate normalisation constant, i.e.

$$G \equiv \left(\sum_{n=0}^{N_{\max}} \frac{\rho^n}{\phi(n)}\right)^{-1},$$

and $R_i\left(n',\hat{\xi}\right)$ corresponding to the applicable MRTD scenario. The expected transfer time $\mathbf{T}$ is follows from applying Little's formula:

$$\mathbf{T} = \frac{1}{\lambda \Pr\{N < N_{\max}\}} G \sum_{n=0}^{N_{\max}} \frac{n\rho^n}{\phi(n)}.$$

In Section 5, the presented numerical results will be accompanied by validation results obtained via dynamic simulations.

5. NUMERICAL RESULTS

In this section the results of some indicative numerical experiments are presented. In the experiments, both the Round Robin and Proportional Fair access selection schemes are considered, applying switched or parallel MRTD. The number of available RAs (L) is varied from 1 to 4. A uniform bandwidth of $W^{(l)} = W = 1$ MHz is assumed for all RAs $l = 1, \cdots, L$, while also the flows' average SIRs are assumed uniform over all RAs, i.e. $\hat{\xi}_i^{(l)} = \hat{\xi} = 10$, $i = 1, \cdots, N$, $l = 1, \cdots, L$. Applicable only in the case with non-persistent flows, an average flow size of $A = 1$ Mbit was assumed, while the admission control threshold has been set at $N_{\max} = 100$.

5.1 Throughput results for persistent flows

Figure 2 presents the expected flow throughput versus the N/L, i.e. the number of *persistent* flows N normalised with respect to L, where the case of switched (parallel) MRTD is shown in the left (right) chart. The reason for normalising

the persistent 'traffic load' is to allow an adequate assessment of the performance impact of L. It is obvious that an increase in L will not degrade the performance, but the interesting question is whether or not a *proportional* performance improvement exists. Observe that cases of $L > N$ are also depicted for $L > 1$, which obviously correspond to a fractional number of flows per RA. For the charts with switched MRTD, expressions (3) and (9) are applied if $L \leq N$, while expressions (4) and (11) are applied if $L > N$, for the RR and PF scheme, respectively.

A number of observations can be made from these charts. Observe from both charts that the PF scheme always performs strictly better than the RR scheme due to multi-user and/or multi-access diversity, except for the case of $L = N = 1$, which obviously offers neither form of diversity. For $L > N$, the PF scheduler gives better performance than the RR scheme due to multi-access diversity, while for $N \geq L$, the PF scheduler additionally and increasingly exploits multi-user diversity and consequently outperforms the RR scheme to a larger degree, comparing the obtained throughput values in a relative sense. The relative contribution of multi-access and multi-user diversity to the expected throughput is changing from a predominant gain from multi-access diversity for low N/L to a predominant gain from multi-user diversity for higher (and more realistic) values of N/L.

The gain from parallel over switched MRTD is most significant for very low N/L, i.e. those scenarios where the available number of RAs per flow is relatively large. The rationale for this is that in the case of switched MRTD a number of RAs are left idle, which *are* utilised under parallel MRTD. In particular, note that the gain from parallel MRTD quickly vanishes in the practically rather likely case where N exceeds L.

Comparing the curves for different L, we note that an increase in the available capacity, yields a *proportional* enhancement of the flow throughput for the case of RR access selection, for both switched and parallel MRTD (observe the overlapping curves and recall the applied 'traffic load' normalisation). This is also immediately clear from the derived expressions: $R_i = H(1,l,\hat{\xi})$, for $N/L < 1$, and $R_i = (L/N) H(1,l,\hat{\xi})$, for $N/L \geq 1$. For the PF scheduler however, it is readily observed that the curves for higher L lie above those for lower L, which indicates an *above-proportional* performance improvement for increasing L. Since the access selection scheme is designed to intelligently choose which flow to assign to which RA, an increase in L (which for a given N/L also implies an increase in N) provides a larger degree of freedom in the access selection process and thus enhances the diversity gain.

5.2 Transfer time results for non-persistent flows

Figure 3 presents the expected transfer time $\mathbf{T}$ versus the offered traffic load ρ, normalised with respect to L, for the scenario with *non-persistent* flows, i.e. including the flow level dynamics. Again, the case of switched (parallel) MRTD is depicted in the left (right) chart.

As expected from standard queueing results, the transfer time curves show an exponential increase as the offered traffic load approaches some critical value. Actually, for traffic loads beyond this 'critical load' value, the admission threshold imposed by $N_{\max}$ causes the expected transfer time curve to converge to an upper bound. Although this effect natural-
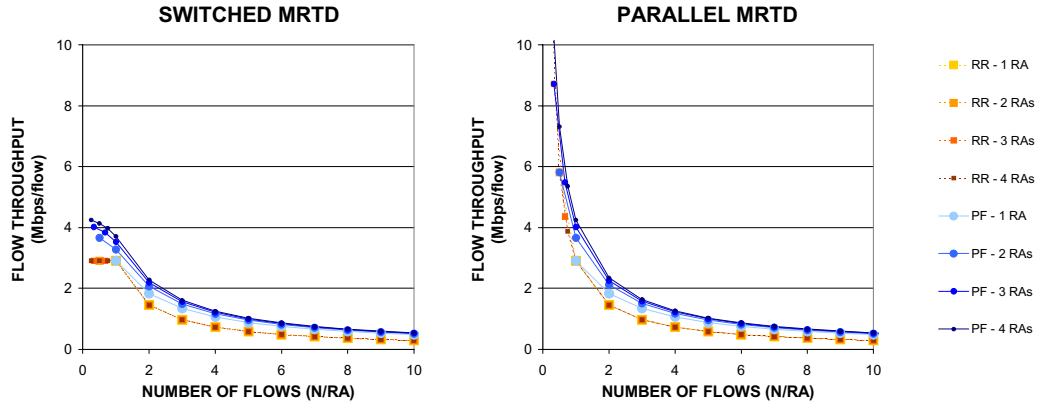
SWITCHED MRTD

PARALLEL MRTD

Figure 2. Persistent flow analysis - The figure shows the expected flow throughput versus the number of persistent flows, normalised by the number of RAs, for the case of switched (left) and parallel (right) MRTD.
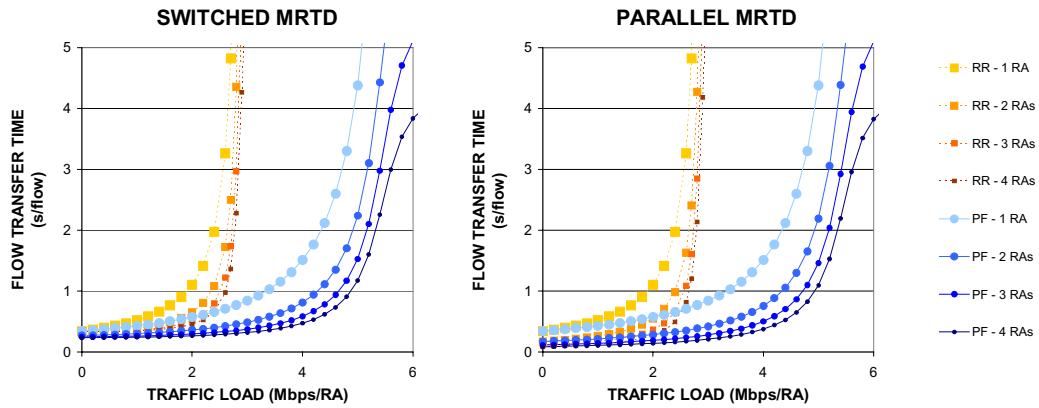
SWITCHED MRTD

PARALLEL MRTD

Figure 3. Non-persistent flow analysis - The figure shows the expected flow transfer time versus the offered traffic load, normalised by the number of RAs, for the case of switched (left) and parallel (right) MRTD.
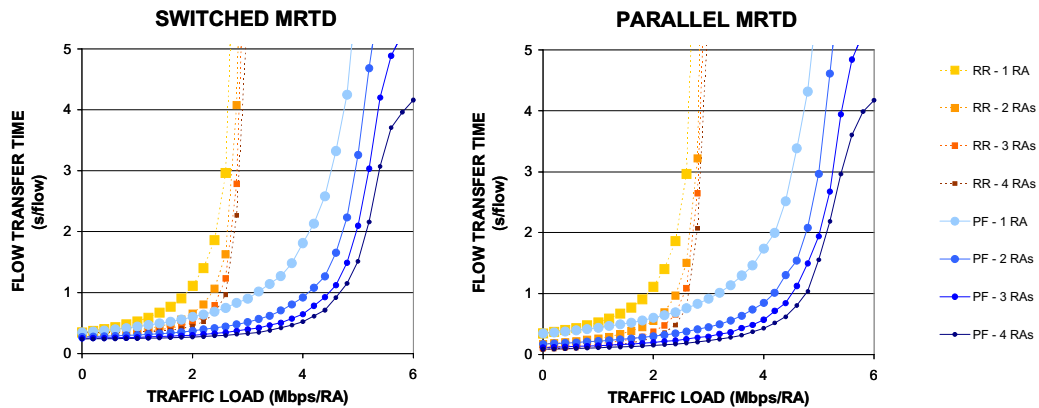
SWITCHED MRTD

PARALLEL MRTD

Figure 4. Validation results - The figure shows the expected flow transfer time versus the offered traffic load, normalised by the number of RAs, for the case of switched (left) and parallel (right) MRTD. The results are obtained via dynamic simulations, integrating the effects that were separated in the persistent and non-persistent flow analyses, as well as including a realistic multipath fading process and channel quality feedback delays.

ly appears for any access selection scheme and any choice of L, in the presented chart it is only (slightly) visible for the PF scheme ($L = 3, 4$). The 'critical load' value loosely indicates the supportable traffic load, which for the case of RR scheduling corresponds with $\lim_{N \to \infty} R_i\left(N, \hat{\xi}\right)$ (see Figure 2). Due to the ever increasing multi-user diversity gain, for the PF case, $R_i\left(N, \hat{\xi}\right)$ diverges to ∞ for $N \longrightarrow \infty$, i.e. the resource efficiency and effective service rates increase in the traffic load. As is clear from Figure 2, this capacity increase is less than linear in N, and hence the expected transfer time curve is still exponential in ρ. Consistent with the results for the case of persistent flows, we observe that the PF schemes can support much higher traffic loads for a given transfer time target than the RR schemes or, conversely, achieve significantly lower transfer times for a given traffic load. As noted above, this performance gain is due to multi-user and multi-RA diversity, and grows as the system is more heavily loaded.

An important distinction between the throughput results presented for the scenario with persistent flows, and the transfer time results presented for the scenario with flow level dynamics, is the achieved trunking gain. In the persistent flow scenario, it was noted that an increase in system capacity leads to a proportional (RR scheme) or slightly above-proportional (PF scheme) increase in the achieved performance. In the dynamic scenario, the performance gains, whether expressed in the achieved transfer times or the supportable traffic load for a given transfer time target, are significantly *above*-proportional for *both* access selection schemes (recall again the applied traffic load normalisation). This gain is generally referred to as a trunking gain.

Considering the different types of gains that are attainable, the presented numerical experiments indicate that the most significant gains come from exploiting multi-user/multi-access diversity, then from aggregating system-specific capacities (trunking gain), while the benefits of parallel over switched MRTD appear to be rather limited. Still, we note that in lightly loaded scenarios parallel MRTD does enhance the peak rates and may thus enable support of highly demanding services that may not be admissible on a single RA.

5.3 Validation experiments

Dynamic simulations of the model have been carried out to validate the separation of time scales and the applied two-stage stochastic analyses, as well as to assess the effect of the ignored feedback delay on the performance of the RA selection schemes. As the effect of the feedback delay on the applied bit rate is independent of the applied RA selection scheme, we concentrate solely on its effect on the access selection decisions, for reasons of transparency.

A scheduling heartbeat of 2 ms and a feedback delay of 8 ms were applied, in accordance with the HSDPA technology (see e.g. [13]). The Rayleigh fading process was modelled according to Jakes' model [15] with a fading velocity of 0.8, 4.0 or 20.0 m/s, where a higher fading velocity is expected to yield a more significant outdatedness of the channel quality feedback information and hence lead to worse (more random) RA selection decisions in case a PF scheme is applied. The flow sizes were sampled from a hyperexponential distribution with a mean of $A = 1$ Mbit (in accordance with the analytical results) and a coefficient of variation of 3. All

other parameters were matched to those used for Figures 2 and 3. For a fading velocity of 0.8 m/s, which is noted to reflect a realistic scenario for data users, Figure 4 depicts the obtained results for the scenario with flow level dynamics, which is noted to implicitly capture the scenario with persistent flows as well.

Comparing Figures 3 and 4, it is clear that the results obtained via analytical methods are quite accurate, and all analytically obtained insights are perfectly valid. In order to give an indication of the performance enhancement of PF over RR access selection, for both MRTD flavours, and to see how this gain is affected by the fading velocity in case of a realistic feedback delay, Table 1 presents the percentual gain of the supportable traffic load when using PF rather than RR access selection, given a 1 second target for the expected transfer time. It is noted that higher fading velocities reduce the PF performance but leave the RR performance unaffected. Observe that, as intuitively expected, the gain from PF over RR access selection diminishes to zero as an increased outdatedness of the channel quality feedback makes the access selections more and more random.

Table 1: Increase of the supportable traffic load when using PF rather than RR access selection.

fading velocity	feedback delay	SWITCHED MRTD			
		1 RA	2 RAs	3 RAs	4 RAs
analysis		76%	81%	86%	88%
0.8 m/s	0 ms	68%	75%	80%	81%
0.8 m/s	8 ms	68%	71%	80%	79%
4.0 m/s	8 ms	21%	28%	28%	31%
20.0 m/s	8 ms	0%	0%	0%	0%
fading velocity	feedback delay	PARALLEL MRTD			
		1 RA	2 RAs	3 RAs	4 RAs
analysis		76%	81%	84%	88%
0.8 m/s	0 ms	68%	75%	78%	81%
0.8 m/s	8 ms	68%	75%	76%	81%
4.0 m/s	8 ms	21%	25%	29%	30%
20.0 m/s	8 ms	0%	0%	0%	0%

5.4 A different perspective: isolated versus integrated access networks

As a final remark in this section, we wish to place a different perspective on the presented results. Throughout the paper, our focus was on a single access point integrating multiple Radio Accesses and the performance impact of a.o. the number of integrated RAs was assessed. The presented analysis and numerical evaluation can also be applied to assess the difference in performance between a scenario with L *isolated* Radio Accesses and a scenario with L *integrated* Radio Accesses.

Consider the scenario with L isolated RAs. Assume for reasons of transparency that all L RAs are equally loaded, either in the persistent sense, i.e. equal number of persistent flows per RA, or in the non-persistent sense, i.e. an equal offered traffic load λA per RA. As the average throughput (persistent scenario) and the average transfer time (non-persistent scenario) are identical for each individual RA, the overall performance is equal to the performance for an individual RA, which has been evaluated above under the label '$L = 1$'. Since traffic loads have been normalised with respect to the number of RAs in the presented charts, a comparison of scenarios with L' equally loaded *isolated* versus

integrated RAs, is readily done by comparing the curves for $L = 1$ and $L = L'$. Observing Figures 2 and 3 once again from this perspective, the performance gain from network integration is immediately clear, most significantly for the scenario with non-persistent flows, as a consequence of flow level multiplexing.

6. CONCLUSIONS

Motivated by the 'beyond 3G' vision of radio access network integration/composition and coordinated radio resource management, as formulated e.g. within the Ambient Networks integrated project, we have presented a purely analytical performance assessment for a single access point integrating multiple Radio Accesses. Principal focus was placed on the evaluation of multi-user diversity, switched/ parallel Multi-Radio Transmit Diversity and trunking gains. Scenarios with persistent and non-persistent flows have been investigated, concentrating on throughput and transfer time performance, respectively. A number of numerical experiments have been included to illustrate the relative contribution of the distinguished aspects to the performance gain under a long-term fair Radio Access selection policy. From these experiments, it was concluded that the exploitation of multi-user/multi-access diversity with a channel-aware access selection scheme attains the most significant gains, while also the trunking gain that is due to an above-proportional performance enhancement when aggregating system-specific capacities, is noted to be significant. Although parallel MRTD was demonstrated to enhance peak rates in light traffic scenarios, in general the achieved benefits appeared to be rather limited when compared to the implementationally less complex alternative of switched MRTD. Applying a slightly different perspective, the derived results have been argued to also allow an assessment of the performance difference between scenarios with *isolated* versus *integrated* Radio Access Networks, which has been demonstrated to be particularly significant when considering flow level dynamics.

Although the derived results indicate that channel-aware access selection provides significant potential gains, further research is necessary considering a more practical setting, particularly in a scenario with heterogeneous Radio Accesses, where measurement errors, outdated feedback information and protocol delays may reduce the observed multi-user/ access diversity gains. Furthermore, while the presented analysis and numerical experiments concentrated on network capacity and service quality enhancement as a benefit from integrating collocated access points at a single site, additional capacity, quality and also coverage gains may be achieved when integrating non-collocated access points. As another suggestion for continued research, we note that the presented two-dimensional access selection approach, i.e. scheduling over users and Radio Accesses, may also be applicable to the scheduling problem OFDMA systems [24], where the scheduling policy operates on users and frequency sub-carriers.

Acknowledgment

7. REFERENCES

[1] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana and A. Viterbi, 'CDMA/HDR: a bandwidth-efficient high-speed wireless data service', *Communications magazine*, vol. 38, no. 7, pp. 70-77, 2000.

[2] F. Berggren, A. Bria, L. Badia, I. Karla, R. Litjens, P. Magnusson, F. Meago, H. Tang and R. Veronesi, 'Multi-Radio Resource Management for Ambient Networks', *Proceedings of PIMRC '05*, Berlin, Germany, 2005.

[3] F. Berggren and R. Jäntti, 'Asymptotically fair transmission scheduling over fading channels', *Transactions on wireless communications*, vol. 3, no. 1, pp. 326-336, 2004.

[4] T. Bonald, 'Flow-level performance analysis of some opportunistic scheduling algorithms', *European transactions on telecommunications*, vol. 16, no. 1, pp. 65-75, 2005.

[5] T. Bonald and A. Proutière, 'Wireless downlink data channels: user performance and cell dimensioning', *Proceedings of Mobicom '03*, San Diego, USA, 2003.

[6] S.C. Borst, 'User-level performance of channel-aware scheduling algorithms in wireless data networks', *Proceedings of INFOCOM '03*, San Francisco, USA, 2003.

[7] S.C. Borst, O.J. Boxma, J.A. Morrison and R. Núñez Queija, 'The equivalence between processor sharing and service in random order', Operations. Research Letters, vol. 31, pp. 254-262, 2003.

[8] J.W. Cohen, 'The multiple phase service network with generalized processor sharing', *Acta informatica*, vol. 12, pp. 245-284, 1979.

[9] J.W. Cohen, 'On processor sharing and random service' (letter to the editor), *Journal of applied probability*, vol. 21, pp. 937–937, 1984.

[10] A.K. Erlang, 'Sandsynlighedsregning og Telefonsamtaler' (in Danish; translated: 'The theory of probabilities and telephone conversations'), *Nyt tidsskrift for matematik B*, vol. 20, pp. 33-39, 1909.

[11] A. Furuskär and J. Zander, 'Multiservice allocation for multiaccess wireless systems', Transactions on Wireless Communications, vol. 4, no. 1, pp. 174-184, 2005.

[12] D. Gesbert and M.-S. Alouini, 'How much feedback is multi-user diversity really worth?', *Proceedings of ICC '04*, Paris, France, 2004.

[13] H. Holma and A. Toskala (editors), '*WCDMA for UMTS: radio access for third generation mobile communications*', John Wiley & Sons, Chichester, England, 2002.

[14] J.M. Holtzman, 'CDMA forward link waterfilling power control', *Proceedings of VTC '00*, Tokyo, Japan, 2000.

[15] W.C. Jakes, '*Microwave mobile communications*', John Wiley & Sons, New York, USA, 1974.

[16] F.P. Kelly, *'Reversibility and stochastic networks'*, John Wiley & Sons, Chichester, United Kingdom, 1979.

[17] L. Kleinrock, 'Analysis of a time-shared processor', *Naval research logistics quarterly*, vol. 11, pp. 59-73, 1964.

[18] L. Kleinrock, 'Time-shared systems: a theoretical treatment', *Journal of the Association for Computing Machinery*, vol. 14, no. 2, pp. 242-261, 1967.5.

[19] P. Magnusson, F. Berggren, I. Karla, R. Litjens, F. Meago, H. Tang and R. Veronesi, 'Multi-radio resource management for communication networks beyond 3G', *Proceedings of VTC '05*, Stockholm, Sweden, 2005.

[20] N. Niebert et al., 'Ambient networks: an architecture for communication networks beyond 3G', *Wireless Communications,* vol. 11, no. 2, pp. 14-22, 2004.

[21] C. E. Shannon, 'Communication in the presence of noise', *Proceedings of the Institute of Radio Engineers*, vol. 37, no.1, pp. 10-21, 1949.

[22] H.C. Tijms, *'Stochastic modelling and analysis: a computational approach'*, John Wiley & Sons, Chichester, United Kingdom, 1986.

[23] A. Tölli, P. Hakalin, and H. Holma, 'Performance of common radio resource management (CRRM)', *Proceedings of ICC '02*, New York, USA, 2002.

[24] C. Wengerter, J. Ohlhorst, and A. Golitschek, 'Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA', *Proceedings of VTC '05 (Spring)* , Stockholm, Sweden, 2005.