Multi-RAT Traffic Steering – Why, when, and how could it be beneficial?

M. Danish Nisar, Volker Pauli, Eiko Seidel, Nomor Research GmbH, Munich, Germany December, 2011

Summary
Simultaneous deployment of multiple Radio Access Technologies (RATs) and multiple bands of the same RATs in a given coverage area, call for an efficient utilization of the overall available resources to carry the complete offered user traffic, and to provide a superior user experience. Traffic steering aims at distributing the offered load in an optimal manner across different RATs and bands in order to fulfil the desires of both the network operators and the end-users. This concise white paper presents the fundamental concepts involved in the design of generic traffic steering algorithms, and provides a brief description of some primitive traffic steering algorithms explaining the intuition behind them. Finally, we present the current status of our system level multi-RAT / multi-band simulation capabilities, and provide an overview of how more sophisticated traffic steering algorithm can be designed, and how could their performance be compared in terms of different performance metrics.

Introduction & Motivation
The emergence of new cellular communication standards offering higher data rates and better spectral efficiencies, in combination with the desire/necessity of retaining the legacy technologies for as long as possible (in order not only to satisfy the old customers who do not switch to the new technologies, but also to reap the maximum benefits from the already installed infrastructure), is leading us to a cellular communication world where multiple technologies simultaneously co-exist in a given area. Moreover, each operator typically gets a fragmented spectrum (multiple frequency bands of operations) for a given RAT owing to licensing issues. This eventually leads to a multi-RAT and multi-band cellular communication environment1 as shown in the diagram below.

A question of key interest that arises here is how to assign and distribute the given traffic between multiple RATs and bands. The question is of fundamental importance to the cellular operators as well as vendors, because
• it allows to control the load on various RATs and bands, and ensures that all installed infrastructure is being optimally utilized, and

1 In the sequel, we use often the term multi-RAT to refer to both multi-RAT as well as multi-band scenarios.
• it allows to enhance the user experience by assigning a user equipment (UE) to a better RAT/band in comparison with its currently assigned RAT/band.

Incorporating the effects of user mobility, time-dependent network loads, and evolving user demands and quality of experience, this splitting of users between RATs and bands needs to be continually adapted and requires the design of sophisticated traffic steering (TS) algorithms.

In this white paper, we discuss the motivation for multi-RAT traffic steering and highlight some of its underlying fundamental principles. This is followed by brief descriptions of some simple intuitive algorithms for traffic steering, and finally, we present some details of our capabilities with regard to multi-RAT traffic steering simulations. The conclusions and directions for future work are then summarized in the last section.

Fundamental Traffic Steering Principles
Traffic steering in live networks can be realized in two distinct ways:

• Static steering (achieved via RRC redirection): where at the time of connection establishment, a decision-logic is pursued to select the RAT/band to be used by the user throughout the connection.

• Dynamic steering (achieved via RRC handover): where the the network load as well as the user experience is monitored throughout the connection duration, and if needed an inter-RAT handover procedure is invoked to steer the UE to a different RAT.

In the following, we briefly outline some fundamental aspects that need to be incorporated while designing the traffic steering algorithms.

A. Theoretical Considerations
As far as selection between multiple RATs is concerned, since the newer RATs offer higher data rates and better spectral efficiencies, as a matter of fact, it is always best to keep all capable user equipments to the newest RAT. This holds as long as the RAT can handle the offered load. There exist exceptions however; especially with regard to bursty (throughput hungry) traffic such as browsing and FTP/P2P. For bursty traffic, it might be beneficial to utilize the inferior but empty RAT in comparison with the superior but heavily loaded RAT.

Selection between multiple bands of the same RAT admits an even simpler theoretical analysis. A higher frequency band typically offers higher bandwidths, so supported data rates are higher, and can be exploited especially by bursty traffic users. The higher frequencies, however, encounter heavier path losses so are often coverage limited. Thus, as a general rule of thumb, a higher frequency band is reserved for the cell-center, while the lower frequency band is reserved for the cell-edge users.

B. Practical Considerations
When considering traffic steering for practical networks, some other aspects, mentioned below, need to be additionally considered.

• The real-life networks might not be as homogeneous as it may seem. The newer RATs are likely to be installed in phases meaning that they are not available at all places, and additionally the differences in path losses, antenna gains makes the practical network deployments too complicated to derive traffic steering hints based purely on theoretical considerations above.

• The offered traffic consists of a variety of traffic mix which includes some continuous bit rate traffic (e.g. voice or streaming) as well as some bursty traffic (e.g. browsing, FTP/P2P) which all have their different characteristics and requirements. A practical traffic steering algorithm needs to take these into considerations.

• There might be constraints imposed on traffic steering based on traffic types or user
priority classes. An example could be to keep the voice traffic on the most stable and mature RAT instead of a superior but possibly less mature RAT. Traffic steering algorithms in such scenarios need to act in accordance with the imposed constraints and policies.

Traffic Steering Algorithms

In this section, we present some primitive traffic steering algorithms in the order of their increasing complexity (sophistication) with a view to highlight the possible strategies in achieving a superior multi-RAT performance.

A. Pure Coverage based TS

A pure coverage based TS algorithm pursues simple non-blind reference signal received power (RSRP) comparisons at periodic predefined intervals for all UEs. Taking into account some hysteresis (5 dB) and handover biases, each UE can then steered to its "strongest" RAT. The handover biases can be tuned to reflect our relative preference of different RATs or based on RSRP cumulative distribution functions (CDFs) for each RAT to artificially maintain a balance of loads between RATs.

B. Pure Load based TS

A pure load based TS, as described above, pursues RSRP checks without explicitly looking at the RATs’ loads (although handover biases do provide a measure to artificially adjust the loads). A load based TS, on the other hand focuses explicitly on achieving a good load balance between RATs. Thus, based on some sliding window load measurements from all RATs, a pair of over-loaded and under-loaded RAT is identified periodically. A randomly selected active and eligible UE can then be moved from the over-loaded RAT to the under-loaded RAT in order to gradually arrive at the preferred load balance.

C. Load + Coverage based TS

A load + coverage based TS algorithm not only tries to balance the loads between various RATs, it does so keeping in view the coverage perspective. For instance, while offloading to a higher frequency RAT, a UE with highest RSRP level in current RAT can be selected to ensure a good coverage on the target RAT, and while offloading to a lower frequency RAT, a UE with lowest RSRP level in the current RAT can be selected to ensure that a poor coverage on current RAT is avoided.

D. Load + Coverage + Policy based TS

On the top of any TS algorithm, a steering policy can be separately invoked. For instance, the load + coverage based TS can be combined with the policy of restricting the voice traffic on a more stable RAT and avoid its handover to other RATs even in cases of deviations from preferred load balance between RATs. In this setting, the TS algorithm will try to maintain the load balance only as much as possible under the given constraints.

Nomor’s Multi-RAT Simulation Capabilities

In this section, we describe Nomor’s capabilities with regard to

- carrying out multi-RAT system level simulations with different RATs and deployment scenarios,
- development and testing of different traffic steering algorithms, and
- conducting offline analysis for comparing the performance of various traffic steering algorithms and policies.

A. Multi-RAT Simulator

The Nomor multi-RAT simulator is basically composed of individual system level simulators for different RATs (HSPA and LTE) in conjunction with a central control unit (CCU). The individual
RAT simulators are able to carry out system level simulations with multiple cells/UEs deployments almost in real-time, and their results are well calibrated at 3GPP standardization evaluations. The CCU of the multi-RAT simulator is responsible for synchronization between the various active RATs (and bands) in terms of:

- simulation time
- UE traffic activity/inactivity behaviour
- UE positions and their mobility.

Thus, with this central control unit in place, we are able to model a multi-RAT (and multi-band) network deployment and test various traffic steering algorithms that are implemented. This multi-RAT simulator can therefore be efficiently used to analyze and compare the benefits of various traffic steering algorithms.

B. Deployment Scenario

In our preliminary multi-RAT simulations, we consider a simple deployment scenario where we have three RATs deployed via identical (co-sited) towers. The three RATs considered are:

- HSPA with a carrier frequency of 2.1 GHz and a bandwidth of 5 MHz, labelled U21,
- LTE with a carrier frequency of 800 MHz and a bandwidth of 10 MHz, labelled L8,
- LTE with a carrier frequency of 2.6 GHz and a bandwidth of 20 MHz, labelled L26.

An inter-site distance (ISD) of 500 meters is assumed for all three RATs. Most other simulation parameters are adopted from 3GPP recommended simulation scenarios [1]. Further details can be provided on request.

C. Traffic Model

A traffic mix is simulated consisting of browsing, P2P/FTP, streaming, and VoIP users. A birth-death model with different active and inactive intervals for each traffic type is also implemented. Thus, during the course of a simulation run, some users transition from active to inactive state and vice versa. With the adopted settings of active and inactive times, we obtain the following equilibrium probabilities$^3$ of various traffic types:

- VoIP 19%, configured with ca. 64 kbps
- Streaming 27%, configured with ca. 1 Mbps
- P2P/FTP 6%, configured according to the NGMN FTP traffic model with a mean file size of 100 Mbits, and a mean reading time of 60 seconds.
- Browsing 48%, configured according to the NGMN http traffic model with a mean reading time of 10 seconds.

We simulate our multi-RAT simulator at three load levels corresponding to 50, 100, and 150 total UEs. The observed aggregate network load in downlink under these settings has been 40%, 66%, and 80% respectively.

D. Performance Analysis

The multi-RAT simulator besides running a live illustration of network dynamics, records traces of various quantities and parameters of interest into some binary files stored in Nomor’s proprietary format. The so-called trace files can then be later read and parsed by offline evaluation scripts to produce CDFs of various key performance indicators (KPIs). CDF plots of these KPIs along with metrics for the mean and the worst-case user performance can then be used for comparison between various steering algorithms.

As an example, we provide below, in Fig. 2, the throughput CDFs for the P2P/FTP users under a medium load without and with traffic steering in place. The curves (and text) in blue represents U21 RAT, red represents L8, green represents L26, while black represents the aggregate performance. We observe that traffic steering not only changes the fraction of assigned UEs to each RAT, but also steers the UEs to their better

---

$^3$ An equilibrium probability of 19% for VoIP users means that at any instant approximately 19% of the active users are VoIP users.
RATs based on their individual coverage. An increase in mean as well as worst-case user performance is therefore evident.

(a) Without traffic steering; Users randomly assigned to each RAT in proportion to their bandwidth.

(b) With Cov + Load traffic steering algorithm

Figure 2. Throughput CDFs for P2P/FTP users without & with traffic steering in place. Both mean and worst-case performance show improvements.

Table 1: Typical performance improvement by traffic steering as compared to no traffic steering.

<table>
<thead>
<tr>
<th>User KPI</th>
<th>Performance Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Browsing throughput</td>
<td>Up to 15%</td>
</tr>
<tr>
<td>P2P/FTP throughput</td>
<td>Up to 90%</td>
</tr>
<tr>
<td>Streaming latency</td>
<td>0.5x – 0.9x</td>
</tr>
<tr>
<td>VoIP latency</td>
<td>0.5x – 5x</td>
</tr>
</tbody>
</table>

As an alternative to evaluations based on CDFs, different steering algorithms can be compared by defining certain acceptance criterion for the KPIs of different traffic types. These acceptance criteria can then be used along with the CDF plots to obtain the fraction of satisfied UEs for each traffic type, and then weighting with the equilibrium probabilities we can compute the fraction of overall satisfied UEs for each TS algorithm. This allows us to directly compare the overall performance of various steering techniques and policies.

Conclusions

In this white paper, we explained the origin of multi-RAT and multi-band cellular deployments and highlighted the motivation carrying out dynamic traffic steering. We explained why, when and how the properly designed traffic steering algorithms can lead to benefits for both the network operators as well as the end-users. We discussed some primitive steering algorithms and observed that their design is based on certain theoretical and practical considerations as well as designer (operator) preferences.

Finally, we described Nomor’s capabilities with regard to conducting multi-RAT simulations and designing efficient traffic steering algorithms with different objective functions, and presented some sample results showing the potential gains from efficient traffic steering strategies. Thus, we are looking forward for collaborative projects on multi-RAT traffic steering simulations. These could involve simulations with more realistic
traffic mix, and heterogeneous deployment patterns (such 1:3 instead of co-sited towers), with the focus on developing algorithms that level out the user experience by improving the worst-case metrics.

References


Note: This white paper is provided to you by Nomor Research GmbH. Similar documents can be obtained from www.nomor.de. Feel free to forward this issue in electronic format. Please contact us in case you are interested in collaboration on related subjects.
System Level Simulation Services

Nomor Research has developed a comprehensive simulation environment supporting various standards such as LTE, LTE Advanced and HSPA and offers related services to support research, development and standardisation.

Features of the dynamic multi-cell, multi-user system level simulator include:
- macro-cell and HetNet deployments (pico-, femto-cell, relay nodes)
- flexible base station and user configurations and drop models
- different transmitter and receiver chains incl. MIMO, ZF, MMSE
- channel modeling with slow/fast fading, pathloss, full user mobility
- intra- and intercell interference modeling for OFDMA, SC-FDMA and WCDMA
- 2D and 3D antenna pattern and multi-antenna beam forming
- Extensive metrics and KPIs: capacity, throughput, spectral efficiency, user QoS etc

The simulator can be used on project basis or in customized simulation campaigns. The performance of the system level simulator has been calibrated to simulation results obtained in standardisation.

Research on advanced algorithms include, but are not limited to:
- advanced features as link adaptation, HARQ, power control, measurements
- scheduling and resource allocation algorithms considering channel and buffer status, QoS etc.
- inter-cell interference coordination, avoidance and cancellation
- Single user-, multi-user MIMO with open and closed loop feedback
- Cooperative multi-point transmission and reception
- functions for self-organising and self-optimizing networks (e.g. load balancing, mobility optimization, tilt optimisation, range extension, power saving etc.)

If you are interested in our services please contact us at info@nomor.de or visit us at http://www.nomor-research.com/simulation