

Laiho J., Kylväjä M. and Höglund A., 2002, Utilization of Advanced Analysis Methods in UMTS Networks, Proceedings of the 55th IEEE Vehicular Technology Conference (VTC 2002 Spring), vol. 2, pp. 726-730.

© 2002 IEEE. Reprinted with permission.

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of Helsinki University of Technology's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org.

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

Utilization of Advanced Analysis Methods in UMTS Networks

Jaana Laiho, Mikko Kylväjä, Albert Höglund

Nokia Networks

E-mail: {jaana.laiho, mikko.kylvaja, albert.hoglund}@nokia.com

Phone +358718023773, Fax +358718023876

Abstract – The scope of this paper is to introduce new analysis and visualization methods for WCDMA cellular networks. The proposed examples are mainly based on the Self-Organizing Map (SOM) method, but also other neural and statistical methods are equally applicable. The main motivation for advanced methods is to increase the abstraction level from the raw network measurements, i.e. radio access network language, to network functional areas or a language closer to the business of network operator. Furthermore, the vast amount of Quality of Service (QoS) and service combinations 3G will enable, require effective data handling procedures.

I. INTRODUCTION

The main driver of 3G mobile networks is the availability of wide range of multimedia applications and services including those requiring up to 2 Mb/s user bit rates. This new multi-service aspect brings totally new requirements into network optimization process and radio resource management algorithms. One of the modifications is related to the quality of service (QoS) requirements and control. So far it has been adequate to provide the speech services with continuous coverage and with acceptable blocking probability. In the case of UMTS the problem is more multidimensional. For each provided service and service profile the QoS targets have to be set and naturally also met.

An additional complication to the optimization process arises from the fact that the network is optimized based on network measurements. Currently over a thousand possible measurements in the WCDMA radio access network (RAN) have been introduced. Considering networks with thousands of cells, there is a possibility for several offered QoS profiles, it is clear that for optimum handling of the radio access network (RAN), effective cell Key Performance Indicator (KPI) visualization and analysis methods are required. Furthermore, the network management system should not only identify a lack of capacity in the current network but also identify where there is potential to introduce data services where they currently do not exist. Therefore tools aiding the trend analysis can be utilized. Furthermore there is need for noticing

abnormal incidents, analyzing them and providing possible solution.

This paper is organized as follows: Section II describes the performance spectrum, how to build it and its' utilization for RAN visualization and optimization. In section III trend analysis using the performance spectrum is presented. Section IV shows how anomaly detection can be used to ease the anomaly spotting during network monitoring. Section V concludes this paper.

II. PERFORMANCE SPECTRUM

Mobile networks produce a huge amount of spatio-temporal data. The data consists of counters and performance indicators of base stations and quality information of calls. Furthermore, the KPIs can be combined into a cost function, providing a high level view of the network performance. The scope of this section is to demonstrate possible ways to perform the analysis and performance visualization of a network. This section demonstrates how a Performance Spectrum (PS) can be utilized in visualizing the network status, finding a "performance point" for a cell and furthermore in the capacity and quality of service analysis and optimization process of a network. The core of PS is to further process the information provided by a clustering method to such a format that one could fast and effectively conclude the performance and status of the RAN QoS.

A. Building the Performance Spectrum

The Performance Spectrum is built by teaching neural network with data that is obtained from the operational RAN. Typical data used for teaching is performance data in terms of counters, KPIs or cost function results derived from those. The requirement for the teaching algorithm is that it has the ability to find similarities between input vectors and based on that, the network is taught so that similarly behaving neurons are close to each other. The Self-Organizing Map (SOM) is a method suitable for building the PS and it is particularly suitable due to its' excellent visualization capabilities. For details and applications of the SOM

see [2], [3] or [4]. A suitable method for creating the SOM was described in [5].

B. Performance Spectrum for RAN Visualization

A network operator can find use for PS in several phases of a network's lifecycle. PS typically requires collection of performance data from the network and it is well suited for monitoring, troubleshooting and optimization of the network system. As it can be seen in later chapters PS can be used for analyzing the network as an entity or analyzing the behavior of just one cell.

B1. RGB utilization

An enhanced version of PS visualisation is utilising R(red)G(green)B(blue) mapping in case of visualising three performance measures. With this implementation a user can detect the dominating colour in the network, thus the condition of network can be understood fast. Even if KPIs are most commonly used to analyse the network condition and quality, also counters and cost-functions derived from those can be used, as well as any combination of counters, KPIs and cost function results.

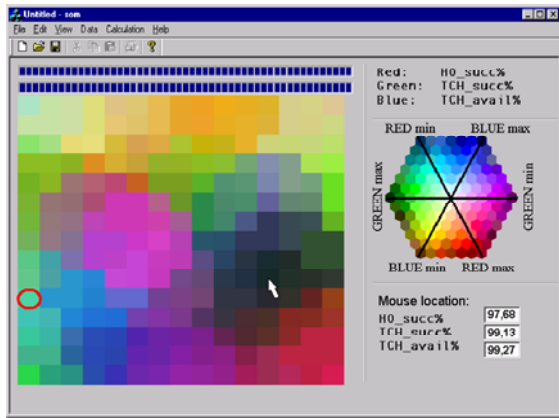


Figure 1. An example of RGB utilization. RGB is used to visualize three different KPIs on the same PS. A simple legend map is drawn with three non-orthogonal axes.

RGB mapping can be understood as a function from the three-dimensional space to a color

$$f_{RGB}(x, y, z) \rightarrow \text{“displayed” color} \quad (1)$$

where $x, y, z \in [0, 1]$ and they represent the intensity of colors red, green and blue correspondingly. With these colors three KPIs can be now presented. The normal variation ranges of the KPIs are scaled to $[0, 1]$. In Table 1 there is simple example with one imaginary cell and three KPIs.

Table 1. Example of RGB utilization.

KPI	Range	Normal variation range	Sample value	[0,1] Scaled sample
HO success%	0 – 100 %	85 – 100 %	96 %	0.73
Dropped Call %	0 – 100 %	0 – 20 %	4 %	0.2
DL quality	0 - 7	0 - 7	3	0.38

The cell in example would now be represented by the color returned by $f_{RGB}(0.73, 0.2, 0.38)$, which is close to violet. The color can be used to analyze the cell status using a picture. As another example, if a cell is located in the place marked with a circle in Figure 1 it can easily be seen from the legend map that green and blue colors (TCH success and availability) are close to maximum, but on the red axis the value is close to the minimum. This means that in this cell the HO (hand over) success is not performing as well as in the other cells and some action should be taken.

B2. KPI Prioritization (Weighting) Vector

The Performance Spectrum itself does not have any upper or lower limits on how many KPIs or counters are needed for teaching the neural system. However, selecting correct and sensible performance indicators as input data for the PS, is required in order to get a useful output from the system. This might be difficult in some cases, because in complex systems it is not transparent, which factors affect each other. Also the selection of performance indicators is dependant on what the PS is used for. As an example, if the goal was to analyze and increase the network's handover performance, the performance indicators that are selected should be related to handovers. Because all of the KPIs and counters are not equally important it has to be taken into account when teaching the neural system. A prioritization-weighting vector (PWV) is one solution, which can be easily used for example when using the SOM. Let m denote the dimension of input data (number of KPIs) used to teach the neural network. Let \mathbf{x} be any of the input vectors used for teaching the system.

$$\mathbf{x} = [x_1, x_2, \dots, x_m]^T \quad (2)$$

User defines a PWV \mathbf{p} of dimension m .

$$\mathbf{p} := [p_1, p_2, \dots, p_m]^T \quad (3)$$

In competitive process the synaptic distance \mathbf{d} , between \mathbf{x} and the j :th neuron, with synaptic weight \mathbf{w}_j is calculated [3]:

$$\mathbf{d} = \arg \|\mathbf{x} - \mathbf{w}_j\|, \quad (4)$$

where $\| \cdot \|$ usually denotes $\sum_{i=1}^m (x_i - w_{ij})^2$ (the squared Euclidean distance). If we take the PWV \mathbf{p} into account $\| \cdot \|$ is $\sum_{i=1}^m (p_i(x_i - w_{ij}))^2$.

It is easier to define \mathbf{p} if all KPIs and counters are scaled so that the range is the same. Otherwise it should be taken into account in \mathbf{p} , which might be difficult.

C. Utilization of Cluster Information in RAN Optimization

An important property of the SOM is that it is able to form clusters of similarly behaving network elements and objects (for example cells) that can be analyzed visually. The scope of this section is to demonstrate the utilization of the cell cluster information in the optimization process. The optimization process utilizing the SOM-generated clusters is depicted in Figure 2. It is assumed that the lower left corner of the performance spectrum indicates cells that have a certain defect. These cells are automatically selected as optimization targets.

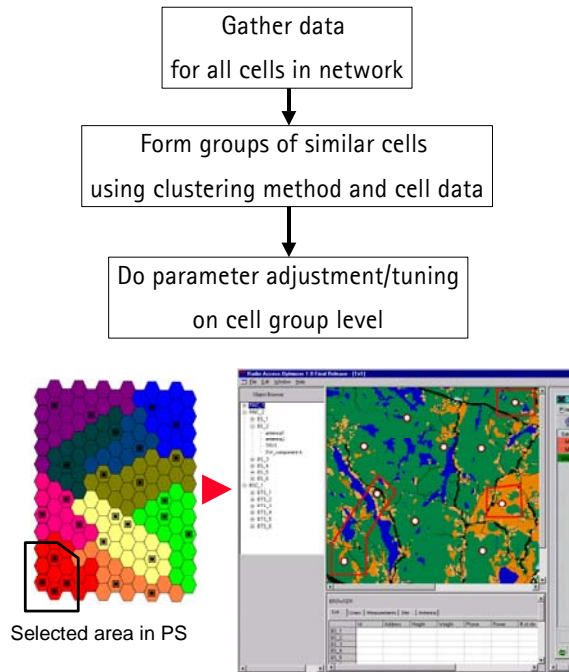


Figure 2. Example of utilization of cluster information in the optimization process. Selection of cells to be optimized or auto-tuned. PS stands for performance spectrum.

In a network with hundreds of cells manual correction of parameter settings is a very time-consuming and close to impossible task. When utilizing cell clustering

it is much easier for the operator to optimize the cell-specific parameters. With the help of performance spectrum the cells can be clustered (grouped) based on traffic profile and density, propagation conditions, cell types, RRM functionality performance etc. Grouping based on multiple criteria instead of just one (like cell type) is more accurate and the optimization of the network will benefit from this. First the network is started with default parameter settings. During the initial sub-optimal operation of the network measurements from the cells are collected. With the help of a clustering method each cell is automatically assigned to a cluster, number of clusters being well under the number of cells in the network. Each cell in a cell-group will use the same configuration parameter values and the optimization process is greatly simplified, improved and made more error prone and efficient. The optimization phase will concentrate on the optimization/automation of cell groups owning a parameter set, rather than optimizing each individual cell with its own selection of configuration parameter settings. In the troubleshooting case the problematic cells can be found rapidly using a clustering method and visualization of the clusters. Additionally using these visualization properties, the operator can easily analyze what kind of cell types he has in his network with respect to certain performance indicators and variables and combine the results with geographical relationships.

It is worth noting that in some cases it is beneficial to perform the clustering twice: first finding rough set of cell clusters and after this perform more detailed analysis in a selected cluster. This is due to the fact that strongly differently behaving cell will dominate in the analysis and thus reduce the sensitivity of the method.

III. TREND ANALYSIS

The complexity of the radio networks is growing as well as the networks themselves. Operators will need means to analyze changes in the network rapidly, given a high number of cells, several services with different QoS criteria and a huge amount of collected performance data. Trend analysis can be performed using data averaged over various time periods, ranging from tens of seconds to days. One could for example follow one cell movement in PS during peak traffic hours, assuming that networks are able to report cell performance frequently enough. Another possibility is to analyze the network's behavior using data collected during a whole year. Figure 3 shows a trend analysis for 32 cells, all in separate displays. There are three main groups or clusters in Figure 3 colored using red, green and blue. For some cells the group membership varies during the monitored period.

The advantage of this method is highly visual representation of changes. Furthermore, the behavior of the cells can be visualized as a function of time, for

example over 24 hours. Depending on the traffic mix and traffic density in the network the performance is different. On the performance spectrum the areas of bad performance are known and it can easily be visualized weather the monitored performance stays out of the unwanted areas. Compared to traditional analysis methods, it is easier and faster to understand the characteristics of cell behavior if this kind of function is used. Another application for trend analysis is related to the network optimization phase. When the network element configuration is changed, the operator normally wishes to see the effect of the change to the performance. The procedure to improve network performance with the PS basically goes:

1. Collect performance data
2. Teach the PS with the data
3. Analyze (This step can be done several times using different time periods)
4. Adjust parameters if needed to correct the possible problem
5. Verify adjustment effect on performance using PS.

One possible way of doing trend analysis is that the PS is taught several times after which the RGB method (described earlier) is used to analyze the changes in dominating colors of the PS.

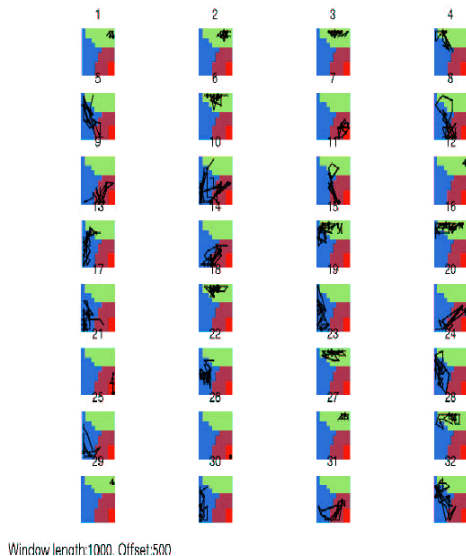


Figure 3. A trend analysis for 32 cells, with all cells in separate displays.

IV. ANOMALY DETECTION

Clustering methods such as the SOM that were used in previous sections to group and monitor cells can also form part of an anomaly detection method. Such a method can be used to detect anomalous or abnormal performance of network elements (e.g. base stations

and radio network controllers). An anomaly detection method makes it much easier to monitor a large amount of cells in a network, i.e. only abnormal observations have to be examined.

The principle of the method is as follows:

1. Select a network element type to be monitored
2. Select variables or performance indicators to monitor, one observation of these variables form a data vector
3. For **each** element to be monitored:
4. Store n data vectors that describe the functioning (normal behavior) of the element during a certain time period
5. Use the vectors as input data to a SOM (or e.g. K-means algorithm) to train a profile for each element, consisting of nodes
6. Calculate distances for each data vector used in training of profile to closest node in profile using a distance measure (e.g. Euclidean distance) to obtain a distance distribution
7. To test if a new data vector is abnormal, calculate a distance to closest node in profile
8. The new observation can be considered abnormal if its' distance is bigger than certain percentage (e.g. 0.5%) of the distances in the distance distribution of profile.
9. The most abnormal variables or performance indicators can be calculated by examining their contribution to the deviating distance. The biggest contribution means the most abnormal variable etc.

For details on the anomaly detection method see [1]. Owing to the fact that there is a very limited amount of measured data available from UMTS networks, the anomaly detection is here demonstrated with GSM speech service related data. Figure 4 shows an example of anomaly detection using hourly data for a GSM base station. The training period in this case was quite short, the previous 14 days with respect to the observation to be tested. The profile was retrained daily with previous 14 days. Eight performance indicators were monitored in addition to two time components (two needed for a daily repetitive pattern). The indicators describe dropping, blocking, traffic, success and requests on TCH (Traffic CHannel) and SDCCH (Standalone Dedicated Control CHannel). The indicators were normalized before analysis and plotting in Figure 4. The first set of anomalies on day 14 seems to be due to high SDCCH dropping. The anomaly on day 16 seems to have been caused by relatively high SDCCH blocking, and dropping while the SDCCH traffic was relatively low. The anomaly on day 23 seems to have been caused by high SDCCH blocking under heavy SDDCH traffic.

The main advantage of the anomaly detection method is that it detects abnormal variable or indicator combinations in addition to abnormal values of individual variables or indicators. The method is therefore very useful in network monitoring and much

easier to use than manual setting and updating of thresholds.

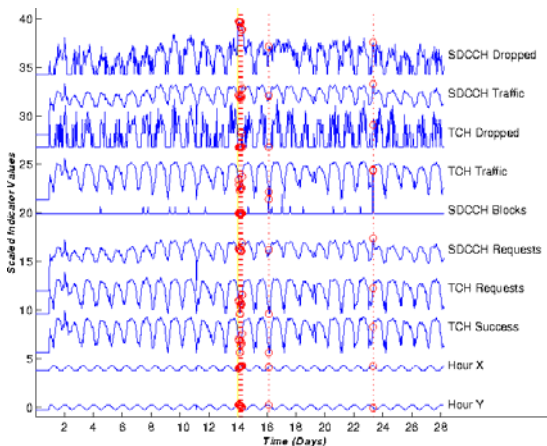


Figure 4 Scaled performance indicators analyzed using moving 14-day profile training. The anomalies are marked with red dotted vertical lines. The first 14 days were only used for profile training and were therefore not analyzed.

V. CONCLUSION

In this paper advanced analysis and visualization methods to support the operator optimization and troubleshooting tasks were introduced. The example cases were based on the Self-organizing Map (SOM), but also other neural analysis and statistical methods can be applied. It was shown that with the introduction of neural algorithms to the network analysis and optimization, the output is highly visual and these advanced methods make it possible to handle much more key performance indicators (KPIs) simultaneously than would be possible by traditional means. Furthermore, when utilizing the SOM-based clustering the behavior of the cells can be classified and the optimization task can be performed per cluster, rather than on a per cell basis. A trend analysis method based on the PS was presented that eases the follow-up of parameter optimization. Additionally, an automatic anomaly detection method, that provides a good way to detect abnormal performance of individual cells was introduced.

Acknowledgements:

The authors wish to thank their colleagues in Nokia for valuable support and comments. We are especially grateful to Pasi Lehtimäki and Kimmo Raivio, both from Helsinki University of Technology.

REFERENCES

- [1] Höglund, A.J., Hatonen, K., Sorvari, A.S: "Computer host-based user anomaly detection system using the self-organizing map", Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, Vol.5, pp.411-416.
- [2] Simon Haykin: "Neural Networks", 2nd edition, 1999 Prentice-Hall, Inc.
- [3] Kohonen, T., Self-Organizing Maps. Second Edition. Springer-Verlag, Heidelberg 1997.
- [4] J.A. Flanagan: "Topologically Ordered Neural Networks", Handbook of Biological Physics, Vol. 4, Ch.16, Eds. F.Moss and S.Gielen, Elsevier, 2001.
- [5] Kimmo Raivio, Olli Simula, Jaana Laiho: " Neural Analysis of Mobile Radio Access Network", Proceedings of IEEE International Conference on Data Mining, ICDM 2001.